

Community Health Workers: Efficacy, Taxonomy, and Performance

Madeleine Ballard

Trinity College, University of Oxford

Trinity Term, 2016

Thesis submitted in partial fulfilment of the requirements for
the degree of DPhil in Social Intervention in the Social
Sciences Division at the University of Oxford

Word count: approx. 84 500

Table of Contents

List of Figures.....	5
List of Tables	7
Abstract.....	9
Chapter 1: Community Health Workers and Evidence Based Practice.....	11
1.1 Introduction	11
1.1.1 Definition of Terms	11
1.2 A Brief History of CHW Programmes	13
1.2.1 First Wave CHW Programmes	14
1.2.2 Second Wave CHW Programmes	17
1.2.3 Community Health Workers Today.....	18
1.3 Defining Evidence Based Practice	22
1.3.1 Evidence Based Medicine.....	22
1.3.2 Evidence Based Public Health	28
1.3.3 Critiques of Evidence Based Practice	32
1.3.4 Evidence Based Practice and CHWs	36
1.4 Complexity and Frameworks for Intervention Development	36
1.4.1 Defining Complex Interventions	36
1.4.2 Frameworks for Complex Intervention Development and Testing	42
1.5 Justification for Research	48
1.6 Research Questions and Structure of the Thesis	50
1.6.1 Research Questions	50
1.6.2 Structure of the Thesis	51
Chapter 2: A Scoping Review of Methodological Guidance for Overviews of	
Reviews.....	53
2.1 Introduction	53
2.2 Foundations and Logic of Overviews of Reviews	54
2.3 Methods	55
2.3.1 Inclusion and Exclusion Criteria.....	56
2.3.2 Systematic Search Strategy.....	56
2.3.3 Selection of Studies	57
2.3.4 Data Extraction	58
2.3.5 Data Synthesis.....	58
2.4 Results	58
2.4.1 Characteristics of the Sample	60
2.4.2 Methodological Guidance and Debates	62
2.5 Discussion	75
2.5.1 Overview Functions.....	76
2.5.2 Overview Decision Tool	81
2.5.3 Strengths and Weaknesses	84
2.6 Conclusion.....	85
Chapter 3: Applying the Overview Decision Tool to the Community Health	
Worker Literature	87
3.1 Introduction	87
3.2 Methods for the Proposed Overview	88
3.2.1 Criteria for Including Studies in the Overview	88
3.2.2 Search Methods for Identification of Reviews	91
3.2.3 Data Collection and Analysis	94

3.3 Results	96
3.3.1 Characteristics of Included Reviews	96
3.3.2 Applying the Checklist	98
3.4 Conclusion	103
Chapter 4: A Systematic Review of Community Health Worker Efficacy	104
4.1 Introduction	104
4.1.1 How the Intervention Might Work	104
4.1.2 Why it is Important to do this Review	105
4.2 Methods	106
4.2.1 Criteria for Including Studies in the Review	106
4.2.2 Types of Participants, Interventions, Comparators, And Outcomes	107
4.2.3 Search Methods for Identification of Studies	108
4.2.4 Data Collection and Analysis	112
4.3 Results	123
4.3.1 Description of Studies	123
4.3.2 Risk of Bias of Included Studies	126
4.3.3 Effects of Interventions	134
4.4 Discussion	201
4.4.1 Summary of Main Results	201
4.4.2 Internal Validity	208
4.4.3 Overall Completeness and Applicability (External Validity) of Evidence	210
4.4.4 Quality of the Evidence	217
4.4.5 Potential Biases in the Review Process	218
4.4.6 Agreements and Disagreements With Other Studies or Reviews	220
4.5 Conclusion	222
4.5.1 Implications for Practice	222
4.5.2 Implications for Research	225
Chapter 5: Toward a Taxonomy of Community Health Worker Characteristics and Programme Features	228
5.1 Introduction	228
5.1.1 Conceptualising Community Health Workers	228
5.1.2 Reporting Guidelines and Intervention Description	231
5.1.3 An Introduction to Taxonomies	234
5.2 Methods	242
5.2.1 Theoretical Underpinnings	242
5.2.2 Data Sources	244
5.2.3 Data Selection and Extraction	247
5.2.4 Data Synthesis	248
5.3 Results	249
5.3.1 Results of the Search: Primary Studies and Influential Papers	249
5.3.2 Results of the Search: Existing Attempts to Classify CHWs	250
5.3.3 Construction of the Taxonomy	255
5.3.4 Explanation and Elaboration of The Taxonomy	259
5.4 Discussion	267
5.4.1 Summary of Main Results	267
5.4.2 Strengths and Contributions of the Taxonomy	268
5.4.3 Weaknesses of the Taxonomy and Future Work	272
5.5 Conclusion	273
Chapter 6: A Systematic Review of Interventions for Improving the Performance of Community Health Workers	275
6.1 Introduction	275
6.1.1 The Enigma of CHW Performance	275
6.1.2 How the Intervention Might Work	278

6.1.3 Why it is Important to do this Review	282
6.2 Methods	282
6.2.1 Criteria for Considering Studies for this Review	282
6.2.2 Search Methods for Identification of Studies	285
6.2.3 Data Collection and Analysis	288
6.3 Results	293
6.3.1 Description of Studies	293
6.3.2 Risk of Bias in Included Studies	308
6.3.3 Effects of Interventions	313
6.4 Discussion	327
6.4.1 Summary of Main Results	327
6.4.2 Internal validity	328
6.4.3 Overall Completeness and Applicability (External Validity) of Evidence.....	330
6.4.4 Quality of the Evidence	335
6.4.5 Potential Biases in the Review Process	336
6.4.6 Agreements and Disagreements With Other Studies or Reviews.....	337
6.5 Conclusion.....	339
6.5.1 Implications for Practice	340
6.5.2 Implications for Research	342
Chapter 7: Conclusion	345
7.1 Overview of the Thesis	345
7.1.1 Reflections	346
7.2 Summary of Findings by Thesis Aim	347
7.2.1 Aim One: Summarise the Effects CHW delivered interventions for Primary Health Outcomes in LMICs.....	348
7.2.2 Aim Two: Achieve Definitional Clarity Around CHW Characteristics and Programme Features	351
7.2.3 Aim Three: Identify the Effect of CHW Performance Improvement Interventions for Primary Health Outcomes in LMICs	352
7.3 Discussion	355
7.3.1 The Nature of Evidence Synthesis.....	355
7.3.2 Methodological Issues in Evaluating Complex Interventions	357
7.3.3 Conceptualising and Measuring CHW Efficacy and Performance	364
7.3.4 Considering the Results Against the Thesis Aims	367
7.4 Implications and Directions for Future Research.....	372
7.4.1 Implications for Practitioners And Policymakers	372
7.4.2 Implications for Researchers	373
7.4.3 Directions for Future Inquiry	374
7.5 Strengths and Weaknesses	375
7.5.1 Strengths	375
7.5.2 Limitations	377
7.6 Conclusion.....	378
References	381
Appendices.....	429

List of Figures

Figure 1.1 Visual representation of the CHW density in sub- Saharan Africa (One Million CHWs Campaign, 2016)	21
Figure 1.2 Evidence Based Medicine (Haynes et al. 1996)	25
Figure 1.3 Evidence Based Practice (Haynes, Devereaux, & Guyatt, 2002).....	25
Figure 1.4 Evidence Hierarchy	26
Figure 1.5 Circle of Methods (Walach et al., 2006)	27
Figure 1.6 Transdisciplinary model of EBP.....	29
Figure 1.7 The Process of EBPH	30
Figure 1.8 Flowchart of one cycle of the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011).....	46
Figure 1.9 MRC Framework for the development and testing of complex interventions (Craig et al., 2008)	48
Figure 2.1 PRISMA Study Flow Diagram.....	59
Figure 2.2 Scope mismatch vs. overlap	66
Figure 2.3 Algorithm to identify and resolve causes of discordance between systematic reviews (reproduced from Jadad et al., 1997).....	74
Figure 2.4 Overview in the absence of a systematic review on every relevant comparison.....	78
Figure 2.5 When to conduct an overview, a pilot decision tool.....	81
Figure 3.1 Pilot decision tool (reproduced from Chapter 2)	88
Figure 3.2 Overview PRISMA flow diagram	98
Figure 4.1 PRISMA Flow Diagram	123
Figure 4.2 Reviewer judgements about each methodological quality item as a percentage across all included studies.	126
Figure 4.3 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.1 Prevalence of depression – adjusted for clustering	140
Figure 4.4 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.2 Severity of common mental disorder symptoms (includes anxiety and depression) – adjusted for clustering	141
Figure 4.5 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.3 Functional impairment of adults with common mental disorders – adjusted for clustering	141
Figure 4.6 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.4 Severity of post-traumatic stress disorder (PTSD).....	142
Figure 4.7 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.5 Severity of depression in adolescents	143
Figure 4.8 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.1 Mortality among children under five years (0-59 months) – adjusted for clustering	147
Figure 4.9 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.2 child morbidity – adjusted for clustering	148
Figure 4.10 Comparison 2: CHW delivered interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.3 Infant mortality – adjusted for clustering	148

Figure 4.11 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.4 Infant morbidity – adjusted for clustering	149
Figure 4.12 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.1 Stillbirths – adjusted for clustering.....	155
Figure 4.13 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.2. Neonatal mortality – adjusted for clustering.....	156
Figure 4.14 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.5. Perinatal mortality – adjusted for clustering.....	157
Figure 4.15 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.6. Maternal mortality – adjusted for clustering.....	157
Figure 4.16 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.7. Maternal morbidity – adjusted for clustering.....	158
Figure 4.17 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.8. Facility-based delivery – adjusted for clustering.....	159
Figure 4.18 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – adjusted for clustering.....	160
Figure 4.19 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – adjusted for clustering.....	161
Figure 4.20 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.11. Initiation of breastfeeding within 1 h of birth – adjusted for clustering	162
Figure 4.21 Comparison 4: CHW delivered interventions to promote cognitive development in infants, Outcome: 4.1 Maternal knowledge of infant development – adjusted for clustering	165
Figure 4.22 Comparison 4: CHW delivered interventions to promote cognitive development in infants, Outcome: 4.2 Cognitive development – adjusted for clustering.....	166
Figure 4.23 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.1 Weight-for-age Z score – adjusted for clustering.....	169
Figure 4.24 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.2 Height-for-age Z score – adjusted for clustering.....	170
Figure 4.25 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.3 Weight-for-height Z score – adjusted for clustering.....	171
Figure 4.26 Comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, Outcome: 6.1 Cure for smear positive TB patients (new and retreatment) – adjusted for clustering	174

Figure 4.27 Forest plot of comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, outcome: 6.2 Treatment completion rate for all pulmonary TB patients (vs. facility-based care) – adjusted for clustering.....	174
Figure 4.28 Forest plot of comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, outcome: 6.2 Treatment completion rate for all pulmonary TB patients (vs. family DOTS) – adjusted for clustering.....	175
Figure 4.29 Forest plot of comparison 7: Usual care versus community health worker for maintenance of ART, outcome 7.1 Death (12-24 months) – adjusted for clustering.....	178
Figure 4.30 Forest plot of comparison 7: Usual care versus community health worker for maintenance of ART, outcome 7.2 Loss to follow-up (12-24 months) – adjusted for clustering.....	178
Figure 4.31 Forest plot of comparison 7: CHW delivered interventions to promote injury prevention in children, outcome: 7.1 hazard score – adjusted for clustering.....	181
Figure 4.32 Forest plot of comparison 9: CHWs to promote immunisation uptake and mass drug administration, outcome: 9.2 Mass therapy coverage – adjusted for clustering.....	185
Figure 5.1 Indian caste system.....	234
Figure 5.2 Bloom's Taxonomy (1956).....	236
Figure 5.3 Identification of records for taxonomy construction.....	249
Figure 5.4 Identification of existing information classification systems.....	251
Figure 5.5 Refining of codes.....	257
Figure 5.6 Final list of codes.....	258
Figure 5.7 Dimensions along which CHW characteristics and programme features are categorised in the taxonomy.....	259
Figure 5.8 Faceted CHW Taxonomy.....	279
Figure 6.1 CHW taxonomy categories from Chapter 5.....	279
Figure 6.2 PRISMA study flow diagram.....	294
Figure 6.3 Review author's judgement about each methodological quality item presented as percentages across all included studies.....	309

List of Tables

Table 1.1 The Process of EBM.....	24
Table 1.2 Characteristics of simple vs. complex interventions (Rehfuess & Akl, 2013).....	37
Table 1.3 Two Critical Questions about Behavioural Interventions.....	44
Table 2.1 Characteristics of Included Studies.....	61
Table 3.1 Corrected Coverage Area.....	99
Table 3.2 Overview criteria 2-4.....	100
Table 4.1 Review reviewers' judgements about each methodological quality item for all included studies.....	127
Table 4.2 Economic analysis of included studies.....	198

Table 5.1 Characteristics of classification systems from the natural, medical and social sciences (Adapted from Stavri & Michie, 2012).....	240
Table 5.2 Existing CHW information classifications systems.....	252
Table 6.1 Brief Study Characteristics	295
Table 6.2 Methodological quality summary: review author's judgement about each methodological quality item for each included study	309
Table 6.3 Primary outcomes.....	315
Table 6.4 Secondary outcomes.....	317
Table 6.5 Summary of cost data.....	326

Community Health Workers: Efficacy, Taxonomy, and Performance

Madeleine Ballard

Trinity College, University of Oxford

DPhil Social Intervention

Trinity Term, 2016

Abstract

Background: This thesis presents an empirical investigation into the efficacy, types, and performance of community health workers (CHWs)—trained lay people to whom simple medical procedures can be “task shifted” from doctors. It has three objectives: (1) assess the effects of CHW delivered interventions for primary health outcomes in low-and middle-income countries (LMICs), (2) develop a comprehensive taxonomy of CHW characteristics and programme design features, and (3) assess the relative efficacy of different types of CHW programme designs and how they can be used to optimise CHW performance.

Structure and Methods: Following the logic of early stage intervention development, this thesis has an iterative and developmental structure in which each section flows out of and builds on the previous section. Objective one is addressed in Chapters 2-4: scoping review, systematic review, and meta-analytic methods are applied to establish the efficacy and effectiveness of CHW-led interventions in LMICs. Objective two is addressed in Chapter 5: inductive, thematic analysis of systematically identified trials, influential papers, and existing information classification systems is used to develop a formal CHW taxonomy for intervention reporting and coding. Objective three is addressed in Chapter 6: systematic review

methods are employed to identify interventions for improving the performance of community health workers in LMICs.

Results: Objective one: a systematic review of 155 papers reporting 86 trials found high quality evidence that CHW delivered interventions reduce perinatal mortality, improve child nutritional status, and improve tuberculosis completion rates versus facility-based care. There is also moderate quality evidence that CHW delivered interventions improve certain mental, infectious disease, paediatric, and maternal health outcomes. In undertaking this process, an additional, methodological contribution was made in the form of a tool to reduce risk of bias in overviews of reviews. This tool may facilitate early stage intervention development in the future. Objective two: 253 records were used to establish, in a faceted taxonomy, the definitional clarity required for theory building and knowledge accumulation. Two categories (CHW Characteristics and CHW Programme Features) and six dimensions (Integration, Recruitment, Training, Supervision, Incentives, and Equipment) emerged. Objective three: a systematic review of 14 trials identified moderate quality evidence of the efficacy of CHW performance interventions in improving certain behavioural outcomes for patients, utilisation of services, and CHW quality of care. There was no effect on the biological outcomes of interest.

Conclusion: In bringing the tools of evidence based practice to bear on community health worker interventions, this dissertation has contributed to the theoretical, methodological, and empirical evidence base from which the field can continue to advance.

Chapter 1: Community Health Workers and Evidence Based Practice

This thesis presents an empirical investigation into the efficacy, types, and performance of community health workers (CHWs)—trained lay people to whom simple medical procedures can be “task shifted” from doctors. It is motivated by the need to determine what types of care CHWs can deliver safely and effectively, the lack of definitional clarity around the features of this class of health workers, and the need for better understanding of the relative efficacy of different types of CHW programme designs and how they can be used to optimise CHW performance.

1.1 Introduction

This chapter will introduce CHWs in the context of evidence-based practice, review the evidence that motivates this thesis, and outline the gaps the thesis aims to fill. The chapter begins with a brief history of CHWs, continues with an introduction to the field of evidence based practice and frameworks for intervention development, and concludes with a statement of why this research is needed and what contribution is made by the qualitative and quantitative findings contained herein.

1.1.1 Definition of Terms

Several definitions for and variants of the term community health worker have been employed in the literature. Two recent systematic reviews of CHW delivered interventions (Bhutta, Lassi, Pariyo, & Huicho, 2010; Gilmore & McAuliffe, 2013) identified nearly 70 unique terms (ranging from “barefoot doctor” to “lady health

worker”) used worldwide for community health workers (see Appendix 1a for a listing). The roles and responsibilities of CHWs vary greatly depending on patients’ access to facility-based care and the existence of other cadres of health workers (Love, Gardner, & Legion, 1997). As Perry & Zulliger (2012) have noted, some CHWs have only a few days of training, while others have six months or more; some CHW cadres are paid salaries, others are volunteers; some are generalists, while others perform a narrowly defined set of interventions specific to one disease.

For the purpose of this thesis, the definition adopted by the World Health Organization (1987) and the most frequently cited contemporary definition (Lewin et al., 2010) were drawn upon to define community health workers as any lay health workers who:

- a. Live in the catchment they serve,
- b. Are primarily based in the community (as opposed to a health facility),
- c. Belong to the formal health system (i.e. are managed by the government or an implementing NGO),
- d. Perform tasks related to healthcare delivery, and
- e. Have received organised training but have no formal or paraprofessional certification or tertiary education degree.

With the aim of developing more effective and efficient CHW programmes, this thesis will, at various points, disaggregate these interventions into components. Critically, this thesis will differentiate between the care being provided (e.g. malaria medicine) and the way it is being delivered (e.g. via a nurse, CHW, doctor etc.). In order to avoid confusion, the following terms will be used: (a) *Task* – this refers to the work being undertaken to provide care (e.g. providing health education, prescribing

diarrhoea medication, leading mental health counselling sessions). *Tasks* can be performed by a number of different health cadres, a critical question for CHW programmes is what tasks can be “task-shifted” from higher-skilled health workers to health workers who require shorter training periods, like CHWs. (b) *Delivery agent* – this refers to the health cadre undertaking the task (e.g. the CHW him or herself). One of the aims of this thesis is to delineate the characteristics of these workers. (c) *CHW programme feature* – this refers to components of an intervention related to the CHW (*delivery agent*), as opposed to the *task* being undertaken or the context in which it is delivered. Such features include how a CHW is selected and supported in their work (e.g. training, supervision).

1.2 A Brief History of CHW Programmes

While the history of individual CHW programmes is to some extent, specific to the country in which they operate, this section will trace the broad themes in the emergence of this cadre of health workers. The concept of lay health workers—local people with limited training authorised by the state to provide primary care in the absence of physicians—first emerged in Eastern Europe in the late 1800s (Perry & Zulliger, 2012). These lay health workers, known as *feldshers*, were a forerunner of the modern CHW movement and partial inspiration for the first large-scale CHW program, China’s 1960s-era barefoot doctors (Hsiao, 1984; Lehmann & Sanders, 2007; Perry & Zulliger, 2012; Roemer, 1977; Shi, 1993; Sidel, 1972; Terris, 1977; World Health Organization, 1974; Zhu, Ling, Shen, Lane, & Hu, 1989). The process by which the barefoot doctor concept gained traction internationally is explained in the next section.

1.2.1 First Wave CHW Programmes

First wave CHW programmes gained rapid support from national governments after their inception in the 60's, 70's and early 80's, before facing a decline in support and funding during the structural adjustment reforms of the late 1980's. This section outlines the early history of these pioneering programmes.

Following a crisis of vertical health programming (i.e. stand-alone programmes established to tackle individual diseases) sparked by the failure to eradicate malaria in the 1960s (Cueto, 2004), influential health theorists argued that technocentric, western, facility-based approaches to healthcare weren't addressing illness among rural, impoverished populations (e.g. Bryant, 1969; McKeown, 1976; Newell, 1975; Taylor et al., 1976). During this time, the idea of engaging local non-professional “barefoot doctors” in the provision of health gained global currency, and CHW programmes proliferated in Asia, Africa, and South America (Campbell & Scott, 2011; King, 1966; Lehmann, Friedman, & Sanders, 2004; Perry & Zulliger, 2012). Rather than re-creating doctor-driven, top-down systems focused on tertiary care—systems that had been accused of bias toward urban elites—these countries emphasised training health workers who could provide the preventive and routine curative interventions that were most needed in rural communities (Standing & Chowdhury, 2008). Consistent with the ethos of the era's decolonisation and democratisation movements, these programmes emphasised the role of CHWs not only as providers of care, but “liberators”—agents of social change who could tackle the environmental, cultural, and political factors that structure health (Bhattacharyya, 2001; Campbell & Scott, 2011; Frankel, 1992; Lehmann et al., 2004; Perry & Zulliger, 2012; Werner, 1981).

The 1978 International Conference on Primary Health Care at Alma-Ata, Kazakhstan—“the first truly global health conference” (Perry & Zulliger, 2012, p. 3)—was committed to achieving comprehensive primary healthcare and cemented CHWs as a cornerstone of this effort (Bhattacharyya, 2001; Cueto, 2004; Lehmann & Sanders, 2007; Newell, 1975; World Health Organization, 1978). This vision of primary health care incorporated two agendas (Standing & Chowdhury, 2008). One was driven by pragmatism and focused on how to create lower cost alternatives to expensive metropolitan health systems and address the growing shortage of health professionals (Walt, 1988). The other, far more transformative, agenda stemmed from the belief that ill health was rooted in poverty and inequality and that, as in several postcolonial socialist countries throughout Latin America and Sub-Saharan Africa in the period leading up to the conference, CHWs ought to work with communities to address these challenges via political means (Bennett, 1979; Hall & Taylor, 2003; Kan, 1992; Segall, 1977; Stark, 1985; Werner & Bower, 2005).

Following Alma Ata, many countries initiated national CHW programmes or scaled-up local initiatives to the national level (see Box 1.1 for a case example) (Bhattacharyya, 2001). By the mid-to-late 1980s, however, enthusiasm for CHWs began to wane among donor organisations and ministries of health (Campbell & Scott, 2011). Programmes encountered challenges due to inadequate training, logistical support, incentivisation, supervision, and integration with the wider health system—concerns highlighted in a series of reviews in the late 1980s and early 1990s (Berman, Gwatkin, & Burger, 1987; Frankel, 1992; Gilson et al., 1989; Heggenhougen, Vaughan, Muhondwa, & Rutabanzibwa Ngaiza, 1987; Walt, 1988, 1990). CHW programmes often served as poorly resourced health service stand-ins, rather than vehicles of liberation (Schneider, Hlophe, & van Rensburg, 2008).

Box 1.1: CHWs in Indonesia (Adapted from Crigler et al., 2013)

Throughout the 1970s, volunteers from *Pembinaan Kesejahteraan Keluarga* (PKK), a national women's community development movement, conducted health and nutrition promotion activities in each village. Called *kaders*, these volunteers were formally recognised by the Ministry of Health in the mid-1980s (Frankel, 1992). After a decade, the nationally-run kader programme was providing basic nutrition, growth monitoring, and immunisation services in 86% of villages in Indonesia (Rohde, 1993).

Economic downturn in the mid-1990s, however, significantly affected kader performance: some reports indicate that up to 70% of the health posts (*posyandus*) run by the kader stopped functioning during this time (Suryakusuma, 2013).

In 2001, the Indonesian Ministry of Home Affairs, through a ministerial letter, called for a revitalisation of the *Posyandu* Programme (Ministry of Health Indonesia, 2011). Today, kader remain almost exclusively women chosen by and from the community. Kaders receive one week of training and over time accumulate the skills and equipment necessary to carry out tasks such as growth monitoring treating common illnesses (e.g. diarrhoea), and preventing disease and malnutrition. Kaders are accountable to the village committee, the body that appointed them and supports them in their work, and continue to provide services without financial compensation.

In addition to conceptual and implementation problems, large-scale CHW programmes were hampered by context. Following the global oil crisis of the 1970s, the 1980s saw a global recession and a debt crisis for many developing countries (Perry & Zulliger, 2012). In order to access credit from international organisations, notably the World Bank, developing country governments were required to embrace the free market reforms of structural adjustment (Perry & Zulliger, 2012; Lehmann & Sanders, 2007). The reduction of public sector financing prescribed by neoliberalism meant that by the end of the 1990s, most national CHW programmes had been disbanded (see Box 1.2 for a case example) (Abbatt, 2005; Lehmann & Sanders, 2007; Schneider et al., 2008; Segall, 2003; Standing & Chowdhury, 2008).

Box 1.2: CHWs in Zimbabwe (Adapted from Crigler et al., 2013)

Following its independence from Britain in 1980, Zimbabwe launched its Village Health Worker Programme (VHW) as part of its transition toward primary health care (PHC). From 1982 to 1987, the government trained 900–1000 VHWs annually; by 1987, Zimbabwe had 7,000 VHWs (Frankel, 1992). The aim of the programme was to emphasise health promotion and prevention and provide “some acceptable level of health care to the majority rural population” (Woelk, 1994, p. 1027).

Over the course of the 1980s, the share of the health budget dedicated to preventive services rose from 6.7% in 1980 to 14.4% in 1989 (Sanders, 1993). Economic deterioration in the mid-1990s tied to Zimbabwe's Economic Structural Adjustment Programme (ESAP), however, led to a rapid decline in the health system. Public expenditure on health care declined by 39% in 1994–95 and the VHW programme collapsed soon afterward (Osika et al., 2011; Saunders, 1996).

The federal government reinstated the VHW programme in 2000 (MOHCW Zimbabwe, 2010). Today, VHWs are provided an initial 8-week training and equipped with a medical supply kit. They are responsible for conducting health promotion, treating common conditions (e.g. malaria and diarrhoea) and identifying and referring complicated cases to higher levels of the health system. VHWs receive a quarterly allowance, though it is often irregular, and are directly supervised by the nurse-in-charge at the health centre in their area (MOHCW Zimbabwe, 2010).

1.2.2 Second Wave CHW Programmes

The early 2000s saw a renewed interest in the use of CHWs (Christopher et al., 2011). Prompted by the HIV/AIDS pandemic, the slow pace of progress in reaching the health-related Millennium Development Goals, and resulting questions about how to improve coverage and equity of care in the face of a global health workforce shortage, the WHO began to promote task shifting—the rational redistribution of tasks from higher-skilled health workers to health workers who require shorter training periods—as a way to extend services to populations with limited access to health facilities (Chen et al., 2004; Hadley & Maher, 2000; Maher, Van Gorkom, Gondrie, & Raviglion, 1999; Perry & Zulliger, 2012; World Health Organization,

2006, 2007; World Health Organization, PEPFAR, & UNAIDS, 2008). Task shifting requires a rapidly expanding cadre of lay health workers, and it is in the context of this strategy that the concept of using CHWs to deliver basic health services regained currency and second wave programmes emerged (Crisp, Gawanas, & Sharp, 2008).

As a result of these developments, several countries again began to invest in large-scale CHW programmes (Bhutta et al., 2010; Liberian Ministry of Health and Social Welfare, 2008). As the “liberation, decolonisation, democratisation and self-reliance” zeitgeist of the 1960s faded during the economic crisis of the 1980s, however, so to did the framing of CHWs as agents of social change (Lehmann et al., 2004, p. 3). For example, while Latin American CHWs (typically, *promotores*) initially emerged at the intersection of Catholic liberation theology and the labour rights movement (Pérez & Martinez, 2008), a review of “second wave” CHW programmes reveals that CHWs were now described primarily as health extension workers, their role as organisers replaced by a largely technical function (see Box 1.3 for a case example) (Lehmann et al., 2004).

1.2.3 Community Health Workers Today

Data on the current number of CHWs worldwide are generally poor. The WHO’s Global Health Workforce Statistics 2014 estimate that there are currently 1 316 600 CHWs globally (World Health Organization, 2014). The Global Health Workforce Statistics, however, are compiled from diverse sources such as national population censuses, labour force and employment surveys, national statistical products, and routine administrative information systems. As a result, there is considerable variability across countries in the coverage, quality and reference year of the original data. The CHW estimate reflects this trend: it is based on information from only 46 countries with reference years ranging from 2000 to 2012.

Box 1.3: CHWs in Brazil (Adapted from Crigler et al., 2013)

Brazil's 1988 constitution established the Unified System of Health (*Sistema Único de Saúde* - SUS) and defined health as encompassing both social and political dimensions (Starfield, 1992). This development was associated with a movement to provide social protection, encourage mobilisation, secure the expansion of social rights, and provide free access to services (Fleury, 2011; Macinko, Guanais, & de Souza, 2006). Drawing on the experience of CHW programmes that had existed in Brazil for over a decade (e.g. the Community Health Agents Programme and Visitadora Sanitarias programmes, both initiated in the 1980s), the federal government launched the Family Health Strategy (*Programa Saúde da Família* - PSF) in 1994 to expand health care access to the poorest Brazilians.

Today, Community Health Agents (CHAs) are closely integrated into the formal health service (Svitone, Garfield, Vasconcelos, & Craveiro, 2000). They are trained by the Ministry of Health and serve on family health care teams (*Equipes de Saúde Familiar*) comprised of a doctor, nurse, auxiliary nurse, and a minimum of four CHAs (Macinko et al., 2006). Following pressure from nurses and medical societies CHAs do not administer injections or provide prescription antibiotics as part of integrated management of childhood illness (Saunders, 1996; Victora et al., 2011). Rather, CHAs register the households in the areas where they work and are expected to link them to the formal health system. In 2002 CHAs were officially recognized as professionals by Law No. 10.507/2002 and are now salaried, full-time employees of the federal government (Noronha, 2010; Presidência da República Brasil, 2002).

To counter this trend, the One Million CHWs campaign has set up The Operations Room, an information dashboard to collect country specific information on CHW scale-up activities across sub-Saharan Africa (One Million CHWs Campaign, 2016). While their dashboard is also missing data from several countries (Fig 1.1) it features information from 10 countries not currently featured in the Global Health Workforce Statistics. Perhaps as a consequence, their estimate for the number of CHWs in sub-Saharan Africa (322 200 CHWs in 34 countries) is three times higher than that of the Global Health Workforce Statistics (98 800 in 24 countries). That said, numbers in the One Million CHW dataset are by practitioner self-report and no reference year is

provided for any of the data. As a result, estimates may not be exact, coverage may be inconsistent, and information may not be up-to-date. For example, the author of this DPhil thesis personally inputted the data for Konobo District, Grand Gedeh County, Liberia while working as a practitioner in the area in summer 2013. Three years later (summer 2016) the data have not been updated—despite the number of CHWs in the region having grown by an order of magnitude.

As noted earlier, the roles and responsibilities of CHWs vary greatly depending on patients' access to facility-based care and the existence of other cadres of health workers (Love, Gardner, & Legion, 1997). While CHWs share the characteristics outlined in section 1.1.1—working outside of health facilities in their communities, belonging to the health system and performing tasks related to healthcare delivery but having limited training—CHWs are a diverse category of health workers (Perry, Zulliger, & Rogers, 2014). The tasks performed by CHWs nonetheless tend to fall into four broad categories that span both public health and medicine: (1) assisting individuals and communities to adopt healthy behaviours, (2) conducting outreach to ensure access to care, (3) providing primary care, and (4) advocating structural changes related to community health needs (Perry, Zulliger, & Rogers, 2014). The share of tasks in each category varies across locations and, as suggested in a narrative review of CHW programmes (Lehmann et al., 2004), the proportion of tasks in the latter category has declined over time.

Given this variation, it is difficult to describe a CHW's routine practice. That said, qualitative descriptions of large-scale programmes indicate that CHWs commonly perform their tasks in the context of routine home visits, home visits in response to an acute issue, and community-wide gatherings at which they speak (Alam, Khan, & Walker, 2012a; De La Torre et al., 2012; Medhanyie, Spigt, Dinant,

& Blanco, 2012; Olang'o, Nyamongo, & Aagaard-Hansen, 2010; Pongvongsa et al., 2011; Puchalski Ritchie et al., 2012). These activities can be complemented by a range of other duties that may include: re-stocking medical supplies, traveling to the nearest clinic for regional meetings or training, and undergoing performance-based audits or supportive supervision (World Health Organization et al., 2008).

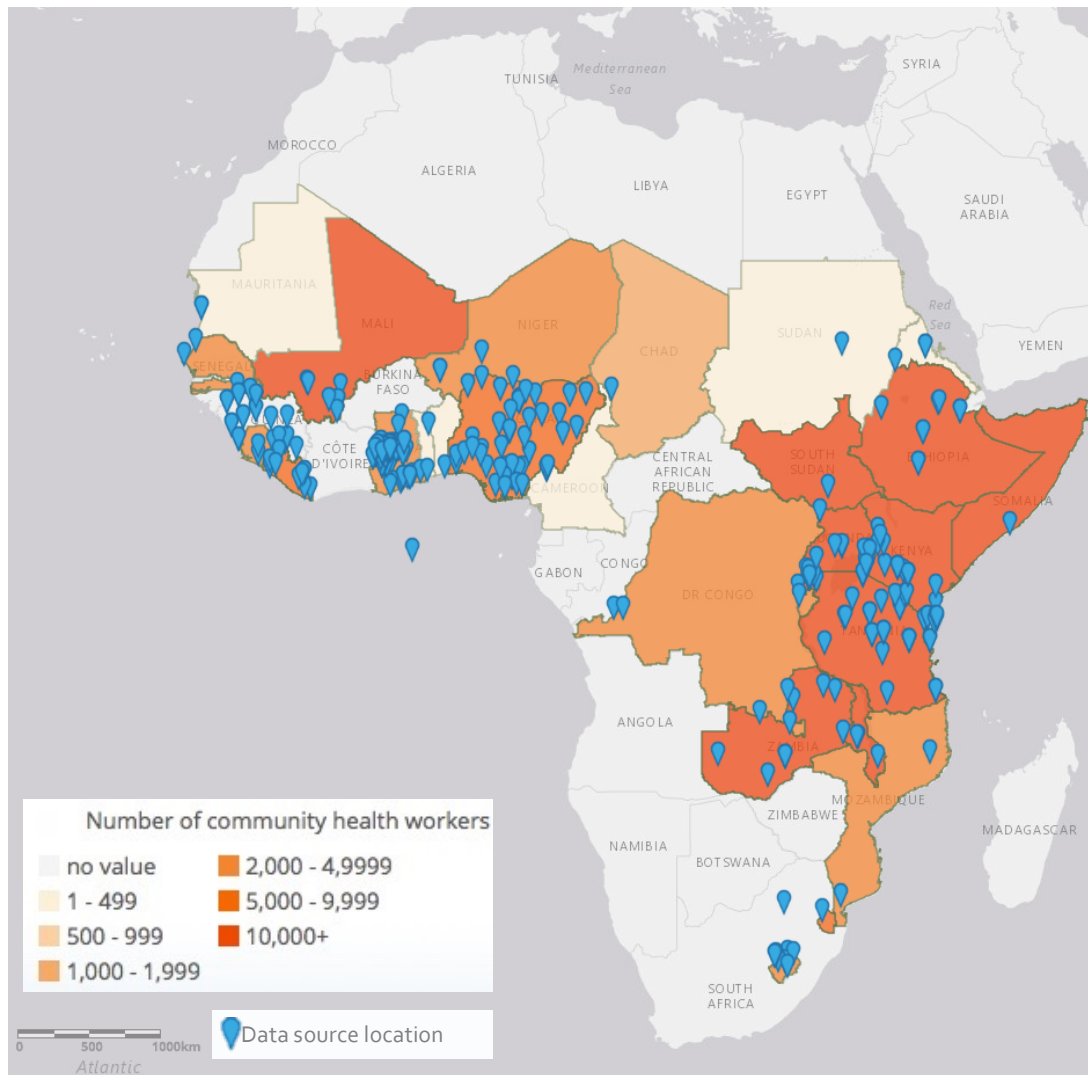


Figure 1.1 Visual representation of the CHW density in sub- Saharan Africa (One Million CHWs Campaign, 2016)

1.3 Defining Evidence Based Practice

CHWs operate at the intersection of individual patient care and health promotion for populations—some of their tasks fall under the umbrella of medicine while others are better classified as public health. Correspondingly, to appropriately situate CHWs in the context of evidence-based practice, this section will review the history, substance, and critiques of evidence-based medicine and evidence-based public health.

1.3.1 Evidence Based Medicine

The origins of evidence based medicine can be traced to a series of lectures by the epidemiologist Archie Cochrane published in 1972 as *Effectiveness and Efficiency: Random Reflections on Health Services* (Cochrane, 1972). While social scientists of the era critiqued the institution of medicine—e.g. professional dominance (Freidson, 1974), iatrogenesis (Illich, 1976), and the fact that non-medical interventions (e.g. housing) were responsible for major health improvement (McKeown, 1976)—Cochrane targeted its scientific underpinnings (Kelly et al., 2010). Cochrane argued that too much of medical care consisted of treatments whose safety and effectiveness was dubious or unknown, causing harm at both the individual and population levels via iatrogenic injury, wasted resources, and failure to employ more effective treatments (Ashcroft, 2004). Because resources would always be limited, he contended that they only be used to provide health interventions shown to be effective in evaluations that seek to minimise bias—like randomised controlled trials (RCTs).

The methods to which Cochrane appealed were pioneered by R.A. Fisher and Bradford Hill in the early 20th century (Kelly et al., 2010). In the 1920s, Fisher demonstrated the requirement of randomisation in experimental design, noting that

randomisation eliminates bias and permits a valid test of significance (Fisher, 1925). In the 1940s, Bradford Hill promoted randomisation for use in the assessment of medical treatments (Hill, 1951). Nonetheless, these methods were not the usual basis for medical decision-making at the time of Cochrane's lectures (Kelly et al., 2010).

While some of Cochrane's contemporaries considered him "a bit of a maverick" (Kelly et al., 2010, p. 1058), his arguments for the ethical imperative to use resources efficiently and do no harm were ultimately prescient. Later work has demonstrated how even complex interventions can be optimised using efficient randomised experiments (Collins et al., 2011) and that interventions—performed with the best of intentions—often do more harm than good (Chalmers, 2003; McCord, 2003). "Scared straight" programmes for teenage delinquents that increase offending rather than reduce it, group interaction programmes that increase misbehaviour among delinquents rather than improve social skills, and recreation programmes for teenagers that increase criminal offenses rather than keep teenagers out of trouble, are but a few examples of how care does not always produce the intended effect and how rigorous evaluation has helped temper these potentially harmful forces (Dufty & Richards, 1978; Gottfredson, 1987; Petrosino, Turpin-Petrosino, & Buehler, 2002). As Chalmers (2003, p. 23) has since observed, "practitioners are often certain about things that are a matter of opinion."

The term evidence-based medicine (EBM) was first coined and formally introduced in 1992 as a "new paradigm" for the teaching and practice of clinical medicine (Evidence-Based Medicine Working Group, 1992, p. 2420). Following critical exchanges within the medical community (Satterfield et al., 2009), EBM was later more explicitly defined as "the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients"

(Sackett & Rosenberg, 1996, p. 71). Sackett and Rosenberg (1996) described the five steps this would involve in practice (see Table 1.1 for the original and most recent formulation of these steps):

Table 1.1 The Process of EBM

Sackett and Rosenberg (1995)	Straus, Richardson, Glasziou, and Haynes (2005)
1. Convert information needs into answerable questions	1. Converting the need for information (about prevention, diagnosis, prognosis, therapy, causation, etc.) into an answerable question
2. Track down, with maximum efficiency, the best evidence with which to answer them (whether from the clinical examination, the diagnostic laboratory, the published literature, or other sources)	2. Tracking down the best evidence with which to answer that question
3. Critically appraise that evidence performance for its validity (closeness to the truth) and usefulness (clinical applicability)	3. Critically appraising that evidence for its validity (closeness to the truth), impact (size of the effect), and applicability (usefulness in our clinical practice)
4. Apply the results of this appraisal in our clinical practice	4. Integrating the critical appraisal with our clinical expertise and with our patient's unique biology, values and circumstances
5. Evaluate our performance	5. Evaluating our effectiveness and efficiency in executing steps 1-4 and seeking ways to improve them both for next time

This conceptualisation of EBM was represented visually by Haynes, Sackett, Gray, Cook, and Guyatt (1996) in the conceptual model reproduced in Fig 1.2, with the circles illustrating the distinct but overlapping sources of data to be used when making clinical decisions (Satterfield et al., 2009).

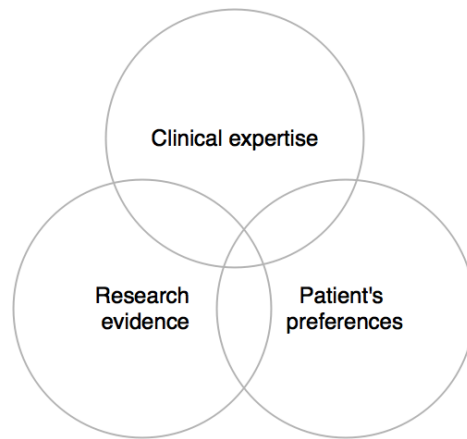


Figure 1.2 Evidence Based Medicine (Haynes et al. 1996)

This original model, however, lacked explicit guidance on how the circles (i.e. sources of data) were to be integrated when making decisions, particularly when they conflicted (Satterfield et al., 2009). An updated model attempted to address these concerns by inserting "clinical state and circumstances" as a new circle and moving clinical expertise to the intersection points of the circles (Fig 1.3). Clinical expertise now being the ability to assess, appraise, and integrate these potentially heterogeneous sources of data (Haynes, Devereaux, & Guyatt, 2002).



Figure 1.3 Evidence Based Practice (Haynes, Devereaux, & Guyatt, 2002)

More recent versions of this model have delineated an evidence hierarchy to help EBM users appraise and integrate multiple types of evidence (Guyatt & Rennie, 2007; Straus et al., 2005). The hierarchy, a key feature that distinguishes evidence-based medicine from other approaches to decision-making, is underpinned by a commitment to internal validity—i.e. the correct assessment of the association between an exposure and an effect in the target population (Delgado-Rodríguez & Llorca, 2004). As such, the relative ranking of different types of study designs is based on their ability minimise selection and information bias, to control confounding, and to attempt to rule out chance (Fig 1.4) (Mills, Montori, & Guyatt, 2004; Porta, Greenland, Hernán, dos Santos Silva, & Last, 2014).

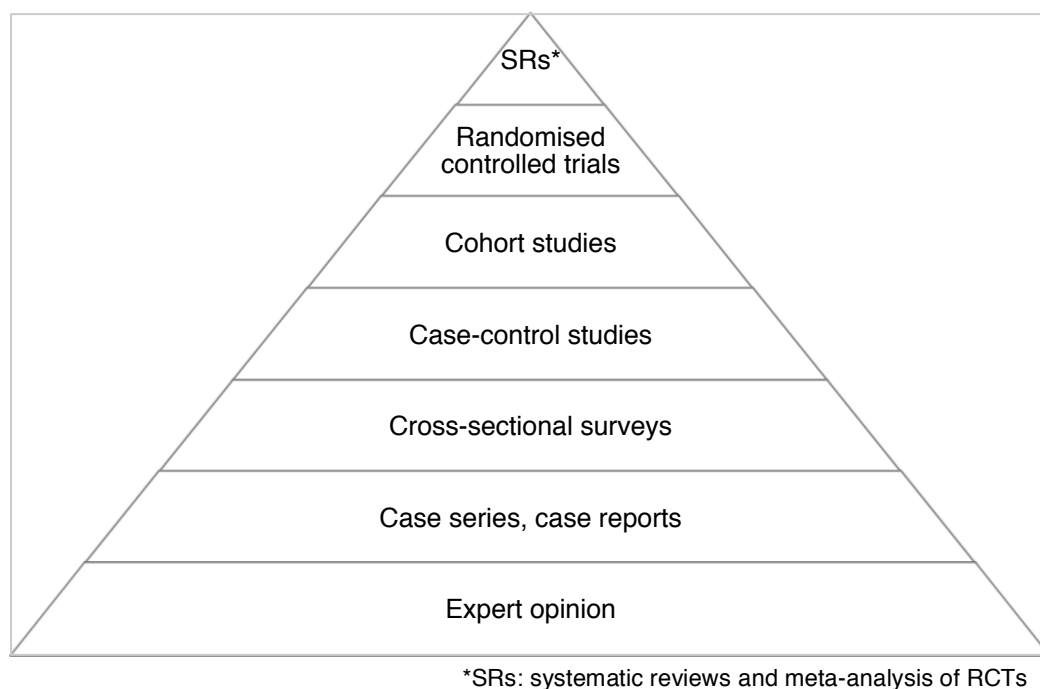


Figure 1.4 Evidence Hierarchy

While external validity—the degree to which results observed in one population can be generalised to others—is likewise an essential part of intervention research and is not reflected in this diagram, as pointed out by Delgado-Rodríguez

and Llorca (2004), it cannot exist in the absence of internal validity. The use of randomisation—in which participants are randomly allocated to intervention or comparison group—helps ensure that any systematic differences between participants (including those that may be unknown or unmeasured) are distributed consistently across groups and thus the influence of confounding factors of prognostic importance is eliminated (Field & Hole, 2002). Nonetheless, the privileging of RCTs over other methods has resulted in considerable debate, even within the medical community (Ashcroft, 2004; Bothwell, Greene, Podolsky, & Jones, 2016). Indeed, Walach, Falkenberg, Fonnebo, Lewith, and Jonas (2006) have recently proposed a Circular Model of evidence that illustrates how employing a multiplicity of methods while counterbalancing their individual strengths and weaknesses can result in pragmatic but rigorous evidence (Fig 1.5). Shading indicates the complementarity of (i) experimental and quasi-experimental or observational methods and (ii) internal and external validity.

These debates, and similar critiques of evidence-based public health, will be explored at the end of the next section.

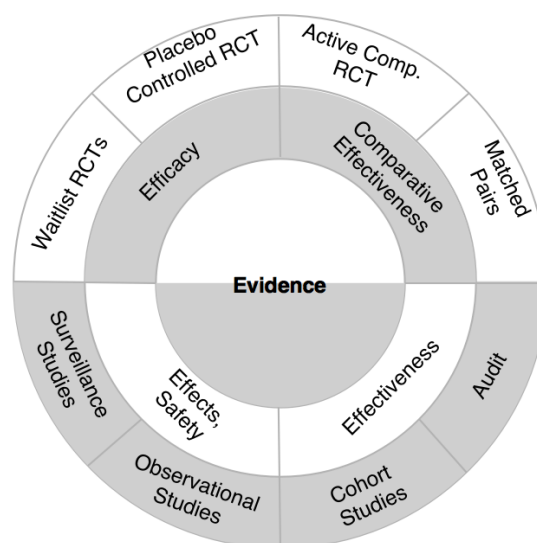


Figure 1.5 Circle of Methods (Walach et al., 2006)

1.3.2 Evidence Based Public Health

Early enthusiasm for evidence-based medicine led to a proliferation of evidence base practice in several disciplines including social work (Sheldon, 2001), ethics (Goldenberg, 2005), nursing (Ingersoll, 2000), and librarianship (Eldredge, 2000). This section traces the evolution of evidence-based public health (EBPH) (Glasziou & Longbottom, 1999).

Formal discourse on the nature and scope of evidence-based public health (EBPH) originated in the late 1990s (Brownson, Fielding, & Maylahn, 2009). Several authors have attempted to define EBPH. In 1997, Jenicek defined EBPH as the “conscientious, explicit, and judicious use of current best evidence in making decisions about the care of communities and populations in the domain of health protection, disease prevention, health maintenance and improvement (health promotion)” (Jenicek, 1997, p. 190). In 1999, Brownson and colleagues crafted their own definition: “EBPH is the development, implementation, and evaluation of effective programmes and policies in public health through application of principles of scientific reasoning, including systematic uses of data and information systems and appropriate use of programme planning models” (Brownson, Baker, Leet, & Gillespie, 2003; Brownson, Gurney, & Land, 1999, p. 87). Kohatsu, Robinson, & Torner (2004) subsequently broadened these two definitions, proposing a definition that places a stronger emphasis on the perspectives of community members: “EBPH is the process of integrating science-based interventions with community preferences to improve the health of populations” (p. 418-419). In this way, Kohatsu, Robinson, & Torner’s (2004) EBPH definition is similar to that of EBM (“the integration of best research evidence with clinical expertise and patient values”) and distinct in that it

explicitly highlights the objective of the process (“to improve the health of populations”).

Ultimately, there appears to be a consensus among public health practitioners and researchers that key characteristics of EBPH include: (1) employing the basis of the best available evidence, (2) engaging the community in decision making, (3) conducting sound evaluation, (4) taking resources and context into account, and (5) disseminating what is learned (Brownson et al., 2003; Brownson et al., 2009; Muir Gray, 2000; Rychetnik, Hawe, Waters, Barratt, & Frommer, 2004). A recent transdisciplinary model grounds evidence-based practice in an ecological framework in order to emphasise the role of context (Fig 1.6) (Satterfield et al., 2009).

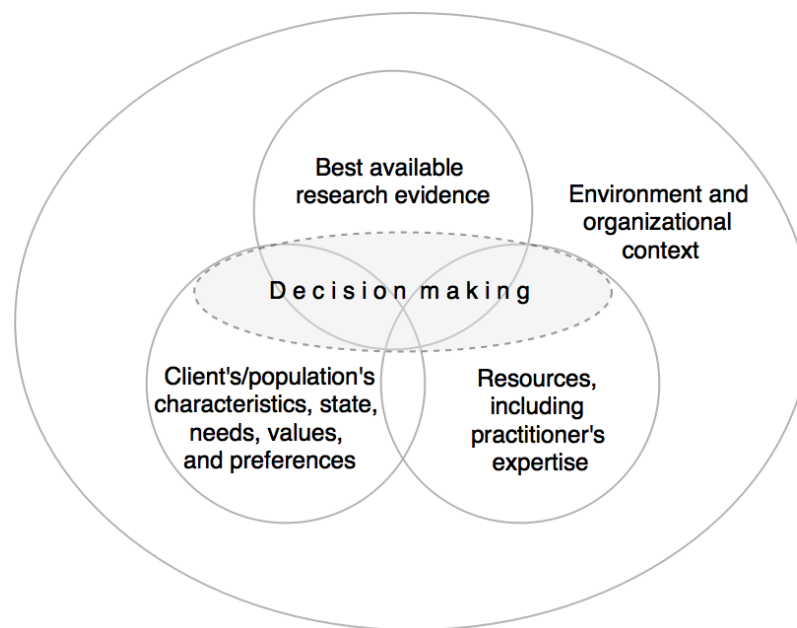


Figure 1.6 Transdisciplinary model of EBP

Brownson et al. (2003) have represented the elements of EBPH in a seven-stage process for putting EBPH into practice by public health practitioners (Fig 1.7). Analogous to the 5 “steps” of EBM, it is the most widely employed framework in EBPH (Satterfield et al., 2009).



Figure 1.7 The Process of EBPH

Important distinctions between EBPH and EBM include (i) the volume of evidence, (ii) a more epistemologically and methodologically diverse evidence base, (iii) multiple levels of analysis and operation, and (iv) the degree of variation in professional training and certification (Brownson, Baker, Leet, Gillespie, & True, 2010).

(i) Compared to the more robust field of EBM, there is a relative dearth of well-designed studies assessing the impacts of public health interventions (Brownson et al., 2010). As a result, attempts to apply criteria from the evidence hierarchy in public health have led, in several areas, to the conclusion that there is no quality evidence on the question of interest (Kemmer, 2006).

(ii) Moreover, the two assumptions underlying the evidence hierarchy—that interventions delivered through RCTs can be replicated under routine conditions and that methodologically sound design is sufficient to maximise internal validity—have been questioned.

EBPH has borrowed the term “intervention” from clinical disciplines, insinuating specificity and discreteness (Brownson et al., 2009). Public health practitioners have recognised, however, that due to the complex causal chains in large-scale public health interventions and often greater time from intervention to outcome, they are difficult to replicate under routine conditions (Kelly et al., 2010). Furthermore, as Victora, Habicht, and Bryce (2004) have argued, probability results alone often fail to provide answers to many of the relevant external validity questions posed by evaluations of such programmes (e.g. cost, sustainability, transferability). Observational studies using adequacy statements (based on trends in process indicators, impact indicators, or both) or plausibility statements (derived from nonrandomised, controlled evaluations) can greatly enhance the internal and external validity of RCT findings by, for instance, accounting for diversions from the RCT protocol or helping to assess the clinical significance of the change. As a consequence, most public health practitioners now accept that different questions require different methods and, as a result, tend to draw from a more epistemologically and methodologically diverse evidence base than EBM (Klein, 2000; Sackett & Wennberg, 1997).

(iii) Further, EBPH represents a shift from individual-clinical level to population-policy level decision-making (Black, 2001; Bonell, Hargreaves, Strange, Pronyk, & Porter, 2006; Dobrow, Goel, & Upshur, 2004; Kelly et al., 2010). Because the operational focus of public health is much more diffuse than biomedicine, it works at several different analytic and operational levels simultaneously. For example, interventions at individual level (e.g. education) might have outcomes for the population levels (e.g. social norms) and vice versa (e.g. mass media campaigns, national legislation) (Kelly et al., 2010).

(iv) Finally, unlike medicine with its clear hierarchies, public health involves a broad range of professionals with nonstandardised training and certification (Brownson et al., 2009).

1.3.3 Critiques of Evidence Based Practice

The merits of EBP have been vigorously debated in both medicine and public health. While some objections betray a fundamental misunderstanding of what EBP entails, others raise broader ethical or epistemological questions. Each will be considered in turn.

Misconceptions

Misconceptions about EBP in social intervention focus on what EBP purportedly ignores or ostensibly implies. Common claims of detractors include that evidence based practice:

- (1) Ignores practitioner expertise,
- (2) Ignores client values and preferences,
- (3) Ignores judgment, i.e. is a "cookbook approach,"
- (4) Ignores research that does not take the form of randomised control trials or systematic reviews,
- (5) Leads to cost-cutting, or
- (6) Leads to therapeutic nihilism in the absence of evidence from randomised trials

(Gambrill, 2003; Gibbs & Gambrill, 2002; Hammersley, 2005)

Though such critiques are important to address in that they are well-established—indeed, Straus and McAlister (2000) somewhat cheekily chronicled them in a systematic review—they are unconvincing when considered within a broad, systemic understanding of what evidence based practice in social work entails.

Arguments 1-3 are similar in that they concern the actors involved in decision-making and EBP's perceived minimisation of their roles. Yet, when considered against the broadly-endorsed definitions of EBM and EBPH offered in sections 1.2.1 and 1.2.2, respectively, it is clear that EBP explicitly takes clinical expertise and

client/community values into account and that judgment is the central feature of the model—the cognitive action that turns evidence into contextualised evidence-based practices (Satterfield et al., 2009). Because it is, by definition, a process that involves evolving external evidence, practitioner expertise, and client values EBP should never lead to formulaic "cookbook" practice (Haynes et al., 2002; Sackett & Rosenberg, 1996).

Critiques that the evidence consulted in EBP consists solely of randomised control trials, or systematic reviews thereof, are likewise untrue. The Cochrane and Campbell Libraries contain a number of EBM and EBPH reviews that include non-randomised intervention studies and the importance of including qualitative research in systematic reviews has long been recognised by those active in the field (Higgins & Green, 2011; Thomas et al., 2004). Though undertaking the latter has raised theoretical and practical issues (Daly et al., 2007; Thomas et al., 2004), Cochrane Research Methods Group focused on refining methods for the inclusion of qualitative research in reviews has been active for 15 years and its innovations have informed a number of qualitative systematic reviews (Chalmers, 2005; Cochrane Collaboration Qualitative & Implementation Methods Group, 2016).

Concerns that EBP will lead to cost cuts and therapeutic nihilism are likewise unwarranted given EBP's broad, inclusive process. EBP—because of its emphasis on patient/community values—may increase, not decrease, cost (Sackett & Rosenberg, 1996; Straus & McAlister, 2000). Likewise, because EBP combines best evidence, practitioner experience and patient/community circumstances, clients are able to benefit from practitioner judgement and contribute—as informed participants—to decision making, even if no relevant empirical data are found (Gibbs & Gambrill, 2002; Sackett, 2000).

Epistemology

Reflecting the age-old divide in social science between positivists and interpretivists, epistemological critiques of EBP contend that rational, mechanistic models of knowledge seeking and decision-making don't appropriately capture our complex and relative existence (Macintyre & Petticrew, 2000; Oakley, 2006; Ritchie, Lewis, Nicholls, & Ormston, 2013; Webb, 2001). Though, as Mullen and Streiner (2004) point out, there is no easy resolution to such debates, EBP does not feed into this polarity as much as been assumed (Sheldon, 2001). While it will never reconcile the competing ontologies at play in such disputes, EBP is a fundamentally pragmatic approach, well placed to help bridge this gap.

Candid assessments of the limitations of research studies and a consideration of individual preferences and contextual factors are built into EBP by default. For instance, RCTs have been criticized for being conducted in highly controlled (viz. "artificial") contexts, prioritising internal validity at the expense of assessing how interventions would work in real world social agencies (Gellis & Reid, 2004; Mullen, Shlonsky, Bledsoe, & Bellamy, 2005). Public health practitioners employing the critical appraisal skills that form part of EBP, however, would—and have—come to the same conclusion (Klein, 2000; Sackett & Wennberg, 1997). In such an instance they might issue a call for effectiveness studies paired adequacy or plausibility designs while drawing on their professional experience and community preferences to issue the "reflexive, locally optimal" client advice detractors assert positivists can't provide (Webb, 2001, p. 57). Given that RCTs have long been used to evaluate social interventions and are not, in fact, a blindly borrowed and recklessly applied biomedical paradigm, further effectiveness studies are possible (Oakley, 1998). While it's true evidence-based practitioners employ a hierarchy of evidence that places a

premium on systematic reviews, they continue to consider client values, professional inclinations, cognitive heuristic devices, and the extent of the indeterminate. In its expansiveness, EBP transcends narrow epistemological paradigms in a pragmatic, patient/community-centred way.

Ethics

In rarer instances, EBP has been accused of violating ethical standards (Gibbs & Gambrill, 2002; Oakley et al., 2003). These objections play into an ethical double standard that has long plagued experimental work and fail to give credence to the potential of rigorous evaluation to minimise harm.

Oakley et al. (2003) chronicle objections to the use of randomisation: that it withholds “valuable” interventions from control groups and, what she calls the more “diffuse” view, that quantitative experimental research is inherently exploitative (Rose, 1994). Such claims ignore what social scientist Donald Campbell (1969, p. 30) pointed out over four decades ago: “designating some interventions as ‘experiments’—a term loaded with negative associations—ignores the reality that policy makers and practitioners are experimenting on other people most of the time.” The only difference being that such “unnamed” experiments are usually secretive or poorly controlled (Chalmers, 2003; Webb & Webb, 1932). It is in the context of this ethical double standard that paediatrician Richard Smithells (1974, p. 41) observed that he needed “permission to give a drug to half of my patients, but not to give it to them all.”

This reality should be particularly troubling from an ethical standpoint in light of the fact that interventions—as explained in section 1.3.1—often do more harm than good (Chalmers, 2003; McCord, 2003).

1.3.4 Evidence Based Practice and CHWs

Chalmers (2003, p. 23) once observed, “practitioners are often certain about things that are a matter of opinion.” EBP’s inherent expansiveness, pragmatism, and experimental rigour make it an ideal approach through which to consider ongoing uncertainty regarding the optimisation of large scale CHW programmes and *help* “adjudicate” among competing approaches (Chalmers, 2003, p 24). In bringing the tools of evidence based practice to bear on community health worker programmes, this thesis aims to use robust synthesis methods, including critical appraisal and comprehensive searching, to assemble and interpret evidence about CHW efficacy, types, and performance. In the spirit of EBP, it is hoped that this empirical contribution will be interpreted in a context where practitioner experience, contextual differences, and patient preference continue to be honoured. The reasoning behind the specific methodological approaches employed will be discussed in the next section.

1.4 Complexity and Frameworks for Intervention Development

The experiments and reviews in this thesis are conducted in a sequential way so that the results of one inform the planning and execution of the next. In this way, the chapters in this thesis reflect a phased approach to intervention development that has its conceptual roots in the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011) and the MRC Framework for Developing and Evaluating Complex Interventions (Craig, Dieppe, Macintyre, & Michie, 2008). A brief introduction contextualising these approaches to intervention development will be given here.

1.4.1 Defining Complex Interventions

Because of their long causal pathways and sometimes multiple levels of operation, community health worker interventions are likely to be classified as

“complex.” One of the first attempts to conceptualise complex interventions in health was made by the Medical Research Council (MRC) in 2000. The MRC now summarises the dimensions of complexity as follows (Craig et al., 2008):

- Number of and interactions between components within the experimental and control interventions
- Number and difficulty of behaviours required by those delivering or receiving the intervention
- Number of groups or organisational levels targeted by the intervention
- Number and variability of outcomes
- Degree of flexibility or tailoring of the intervention permitted

Rehfuess and Akl (2013) further define complexity, noting the key differences between “simple” and “complex” interventions in an extended PICO structure (reproduced in Table 1.2). “Simple interventions” tend to show more of the characteristics in the left column while “complex” interventions tend to show more of those in the right column. Importantly, Rehfuess and Akl (2013) do not distinguish between clinical and public health interventions, but rather simple and complex interventions, as both clinical and public health interventions tend to be located along a spectrum of complexity.

Table 1.2 Characteristics of simple vs. complex interventions (Rehfuess & Akl, 2013)

	“Simple” interventions	“Complex” interventions
Population	• Sick population seeking care	• Healthy general or at-risk population

Intervention	• Individual-level intervention • Single component • “Reactive” treatment through medication or surgery or clinical prevention • Implementation in healthcare setting	• Population- and/or individual-level intervention • Multiple interacting components • “Proactive” prevention through behaviour change and/or technical intervention and/or policy • Implementation in household, community or policy setting
Comparison	• No intervention or alternative intervention through treatment/surgery	• “Business as usual” in several sectors
Outcome	• Shorter causal pathway • One or a small number of health outcomes • Usually impact after short lag period	• Longer causal pathway • Multiple health outcomes and broader societal consequences • Usually impact after long lag period
Delivery of intervention	• Delivery through health sector	• Delivery through multiple sectors
Contextual effects	• Variation between healthcare providers (individuals, institutions) • Patient preference and compliance	• Variation between providers of different intervention components in multiple sectors • Large cultural and behavioural variation

A recent content analysis of published papers (Datta & Petticrew, 2013) found that when defining complexity, nearly one third (31%) cited either the 2000 MRC guidance (Medical Research Council, 2000), the 2008 update (Medical Research Council, 2008), or both. A further 23% cited the paper that accompanied the 2000 guidance (Campbell, Fitzpatrick, Haines, & Kinmonth, 2000a), 12.5% the paper that accompanied the 2008 guidance (Craig et al., 2008), and 7.5% a paper that described the MRC Framework (Campbell et al., 2007). The most highly cited papers not related to the MRC framework were Hawe, Shiell, and Riley (2004), cited by 7% of the sample, and Oakley, Strange, Bonell, Allen, and Stephenson (2006), 6%. The former argues that standardising intervention components treats potentially complex interventions like simple ones and that it is rather the function and process of the intervention that ought to be standardised. The latter argues that randomised

controlled trials (RCTs) are the most rigorous way to evaluate the effectiveness of interventions, regardless of their complexity, but that process evaluations should be an integral element of these randomised assessments.

Underlying these definitions is the assumption that “complexity” and “simplicity” are inherent characteristics of interventions. However, as Petticrew (2011) has noted, most interventions reveal themselves to be complex upon closer inspection. For example, the delivery, participant uptake, and/or context of interventions can vary across trial arms and time or between different evaluations—adding several layers of complexity to even a “straightforward” biomedical intervention (Montgomery, Underhill, Gardner, Operario, & Mayo-Wilson, 2013b). Though the MRC Framework identifies implementation as an aspect of complex interventions, Petticrew (2011) draws out the implication of this concept. In so doing, he sidesteps the above definitions and asks whether simplicity and complexity are instead pragmatic perspectives adopted by researchers that may vary depending on the question asked.

On the opposite end of the spectrum are those who would extend the commonly employed MRC definition. Drawing on complexity science, Hawe, Shiell, and Riley (2009) argue that rather than defining complexity based on an intervention’s constituent parts, the focus ought to be on the dynamic properties of the complex ecological systems into which the intervention is introduced. In this view, an ecological approach is not simply defined, as it is so often by public health programmes, as an intervention having multiple strategies directed at multiple levels (Hawe et al., 2009). Rather, it is meant to emphasise that ecological systems follow the principles of system dynamics (Hawe et al., 2009; Trickett & Birman, 1989; Trickett, Kelly, & Vincent, 1985). While the MRC framework notes the importance

of context in delineating dimensions of complexity, Hawe et al. (2009) suggest that this does not go far enough, proposing instead that an intervention may be seen as “a critical event in the history of a system” (p. 267). In other words, complexity is not just a property of the intervention but a property of the system in which the intervention is implemented. As a consequence, intervention effects might not be the sum of their components but emergent trajectories that result from more dynamic processes (e.g. non-linear relationships in which change in outcome is not proportional to change in input, feedback loops, and interactions among the system’s parts) (Hawe et al., 2009; Shiell, Hawe, & Gold, 2008).

A helpful semantic distinction, blurred by the MRC Framework, is that between “complicated” and “complex.” While complicated interventions (like sending a rocket to the moon) typically contain many components requiring specialised expertise or difficult coordination, they are repeatable and generalisable (Glouberman & Zimmerman, 2002). Complex systems, on the other hand, are adaptive to changes in the local environment, composed of other complex systems (for example, the human body), and behave in a non-linear fashion (Holland, 1995; Lorenz, 1995; Shiell et al., 2008). These systems carry with them large elements of ambiguity and uncertainty (Glouberman & Zimmerman, 2002) and as a result, successfully implementing a complex intervention (like raising a child) does not assure subsequent success.

In summary, complexity is thus characterised not only by multiple, tailored, interacting components comprised of many difficult behaviours operating on multiple levels with multiple outcomes (Medical Research Council, 2008). Feedback loops, and other non-linear linkages and relationships define the longer causal pathways of complex interventions and the complex systems in which they unfold. Petticrew’s

(2011) question of whether this definition applies to all or only a subset of interventions remains open.

For the purposes of this thesis, complexity will be understood in line with Rehfuess and Akl (2013) and the Medical Research Council (Craig et al., 2008): as a set of characteristics that interventions display to varying degrees. CHW programmes, by virtue of their interacting components, multiple levels of operation, long causal chains, and potential for threshold effects and interactions with social context, can thus be understood as having “complex” characteristics.

This has two epistemological implications: (1) randomised controlled trials remain the most rigorous way to evaluate the effectiveness of interventions, regardless of their complexity (Oakley et al., 2006). By controlling for confounding, RCTs are able to cut through elements of complexity (e.g. long causal pathways, non-linearity) in order to isolate whether a particular intervention has an effect (Campbell et al., 2000a; Rogers, 2008). As Petticrew et al. (2013) have noted, even in the face of complexity it remains useful to answer the initial question “What is the average effect (or range of effects) of the complex intervention on the outcome(s) of interest?” This question, fully delineated in section 1.6, is the first of three connected queries considered in this thesis.

(2) As with all interventions, optimisation requires answering not just first-generation questions (i.e. does an intervention “work”) but second-generation questions (i.e. when, where, why, and how an intervention may work) (cf. Higgans, 1999). Given the multifaceted nature of complex CHW programmes, this requires attempting to isolate not the effect of the intervention as a whole, but the effects and interactions of intervention components. Systematically identifying and testing a

subset of these components related to CHW characteristics and programme features is the focus of the latter two questions posed in this thesis.

Justification for the sequence and structure of these related inquiries is provided in the next section.

1.4.2 Frameworks for Complex Intervention Development and Testing

Public health interventions, for example CHW delivery of primary care, tend to exhibit many of the characteristics described in the previous section. As a result, added challenges exist when attempting to assess their efficacy. Thorough development and testing of interventions within frameworks for complex intervention development can help mitigate these difficulties, and make the final product easier to evaluate, implement and understand. Though a number of implementation design models have been proposed in the literature, there exists broad consensus that early stage development and testing of interventions involves (1) identifying an evidence base, (2) identifying or developing a theory for the intervention, (3) modelling process and outcomes, and (4) using a pilot study to test procedures for acceptability, estimate likely rates of recruitment and retention of subjects, and determine sample size. These frameworks and processes will be reviewed here.

Both sequential and iterative models have been proposed for intervention design. Fraser, Richman, Galinsky, and Day (2009) described a sequence that involves describing the problem, defining a programme rationale, developing a programme theory, and delineating a programme format and session content, followed by expert review, pilot testing and efficacy trials. Rounsaville, Carroll, and Onken (2001) describes a similar, staged sequence in which, based on a theoretical rationale for a problem, an intervention manual is created in stage 1a, pilot tested in stage 1b, and tested for efficacy in stage II. The Medical Research Council (2000)

framework for complex interventions likewise sets out a sequential approach in which the theory of the intervention is addressed in phase 0, the intervention modelled in phase one, and tested in exploratory or pilot trials in phase two. The Medical Research Council (2000) acknowledged that stages can sometimes be addressed simultaneously and that modelling can help refine theory. For example, if an exploratory trial found an intervention to be unacceptable to its intended population of interest, the theoretical basis and design of the intervention might warrant revision (Campbell et al., 2000a). More recent consideration of the MRC framework (Campbell et al., 2007) and indeed, published new guidance (Craig et al., 2008) have built on this observation of non-linearity and propose that all phases of development and testing of an intervention be considered as a larger iterative activity rather than sequential stages.

The MOST framework is similarly nonlinear. MOST is a comprehensive framework for optimising and evaluating behavioural interventions (Collins et al., 2011). While a multicomponent intervention is typically evaluated as a package to determine whether it has a statistically significant effect, MOST seeks, in addition, to determine whether such an intervention is maximising its public health impact (where impact is defined in terms of efficacy, effectiveness, cost-effectiveness, or some other relevant criterion).

Methods in intervention science have largely been developed for intervention evaluation rather than optimisation (Collins et al., 2011). As Table 1.3 illustrates, statistical significance and optimisation are distinct and independent concepts. If an intervention does not achieve a statistically significant effect and has not achieved its maximum public health impact, it is possible that a statistically significant effect is possible if the intervention is optimised. However, if the intervention is already

yielding its maximum public health impact and this effect is not statistically significant, a different approach is warranted. Likewise, a statistically significant behavioural intervention may or may not have achieved its maximum public health impact.

Table 1.3 Two Critical Questions about Behavioural Interventions

Is this intervention achieving its maximum public health impact?	Does this intervention have a statistically significant effect?	
	No	Yes
No	Improvements to improve public health impact possible, may lead to statistically significant effect	Improvements to improve public health impact possible
Yes	Different intervention approach required	Objective of Multiphase Optimisation Strategy (MOST)

The MOST approach is guided by two principles: (i) continuous optimisation—that an intervention can always be made more potent or more efficient and so optimisation is an ongoing process—and (ii) the resource management principle—that research resources (money, experimental subjects, time, equipment, personnel, etc.) must be managed strategically to move intervention science forward fastest. These principles underlie a series of steps aimed at developing interventions with both statistically significant effects and maximum public health impact (Fig 1.8).

The suggested sequence begins with the drafting of a theoretical model, which informs the identification and selection of intervention components that could form part of the intervention. Collins et al., (2011) define a component as “any aspect of an intervention that is of interest and can be separated out for study” (p. 221). Components may (a) be part of the intervention, (b) pertain to how a particular component or the entire intervention is administered or who administers it, (c) be some aspect of the environment in which the intervention takes place, (d) pertain to

adherence to the intervention, or (e) represent any aspect of the intervention pertaining to translation, efficacy, effectiveness, or cost. Because the intervention components suggested by theory could be combined to form scores of different packages, the next step is to examine each intervention component individually. Typically randomised experiments, these component selection experiments are designed to gather information necessary to make decisions about which intervention components and component levels will be included in the intervention. That said, in keeping with the principles of EBP, it is necessary to draw on practitioner judgment when the components of an intervention are being specified (e.g. session length, setting, sequence, content) as their tally likely exceeds the number that can be systematically varied within a reasonable time frame (Rounsaville et al., 2001).

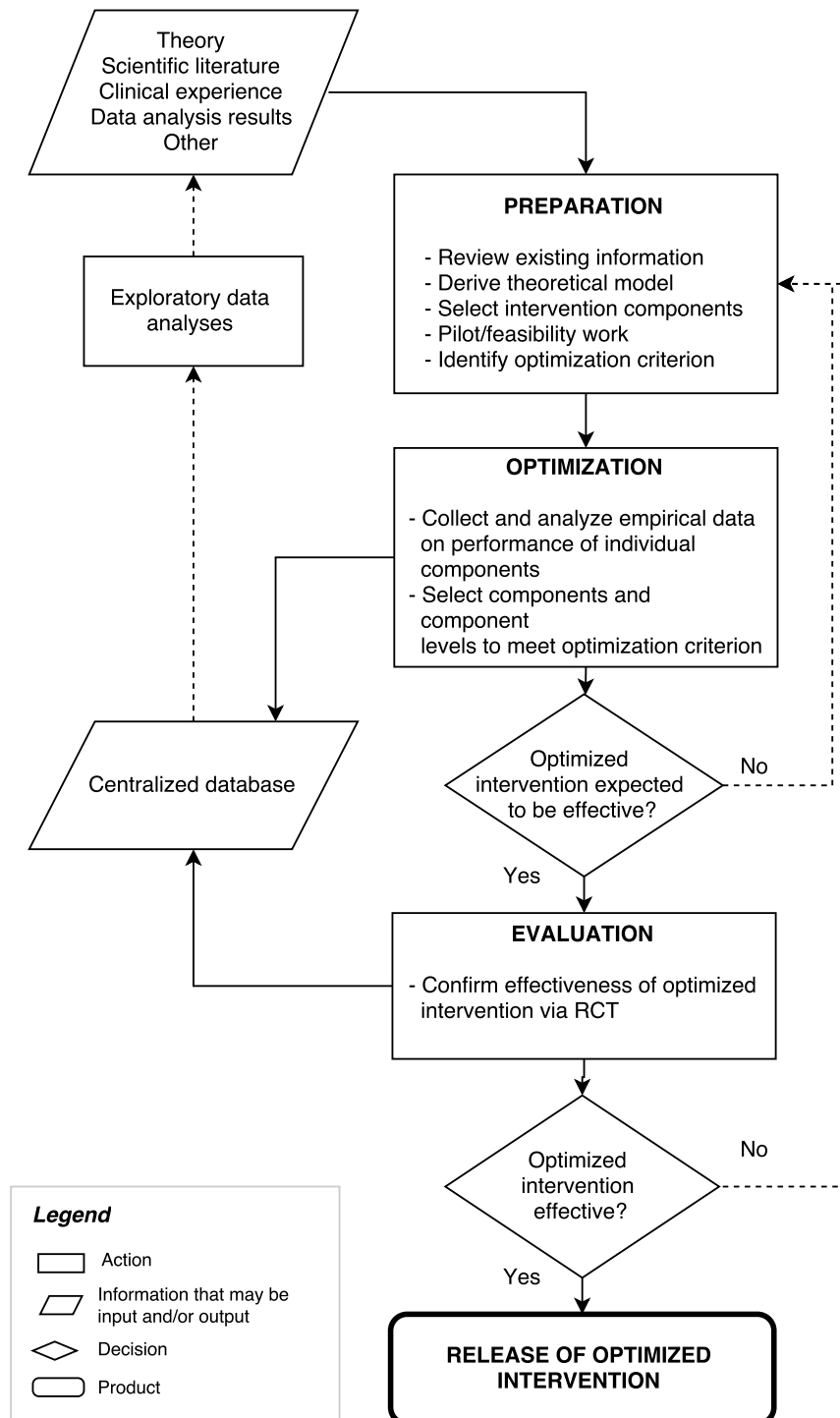


Figure 1.8 Flowchart of one cycle of the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011)

Once the components have been fine-tuned and estimates of the individual and combined effectiveness of a set of intervention components has been established, this

information is applied to assemble a beta version of the intervention. If, based on the effect size estimates obtained in the component selection trials, it appears that the optimised intervention will have a sufficiently large effect to justify proceeding to the evaluation phase, the efficacy or effectiveness of this draft intervention is then assessed in the standard way by means of a randomised controlled trial (RCT). If efficacy/effectiveness is confirmed, the intervention is implemented and a new cycle of optimisation begins.

It is worth noting the role of theory within evaluation is contested (Fretheim, Flottorp, & Oxman, 2005). In a review of health behaviour theories, Munro, Lewin, Swart, and Volmink (2007) found that an explicit theoretical basis was not always necessary for a successful intervention. They argued that further study was needed to determine whether theory-based interventions were more effective than those without an explicit theoretical underpinning. This is, however, perhaps the wrong question. As Craig et al. (2008) have warned, without this theoretical basis it is difficult to interpret trial results—whether or not an intervention is effective. It remains unclear if a result is negative, for example, because the intervention is inherently ineffective, was poorly applied, or was conducted in an inappropriate context. Likewise if the result is positive, it is difficult to assess how the intervention would fare if applied in different contexts (Campbell et al., 2007). The use of theory and evidence-based component selection enables intervention designers to both use and contribute to the accumulation of scientific knowledge in intervention design, evaluation, and synthesis (Collins et al., 2011; De Silva et al., 2014; Moore et al., 2015; Rycroft-Malone, 2007).

MOST and MRC guidance

MOST is thus a particular, optimisation-oriented manifestation of the steps suggested in stages of intervention development as outlined by the MRC (Fig 1.9).

MOST's preparation phase is a particular way of identifying an evidence base, developing a theory for the intervention, and modelling process and outcomes. MOST's optimisation phase is a way of assessing feasibility and piloting the intervention, and the evaluation phase of both MOST and MRC are geared toward assessing effectiveness. Both models, given their iterative nature, are structured to include long-term follow up.

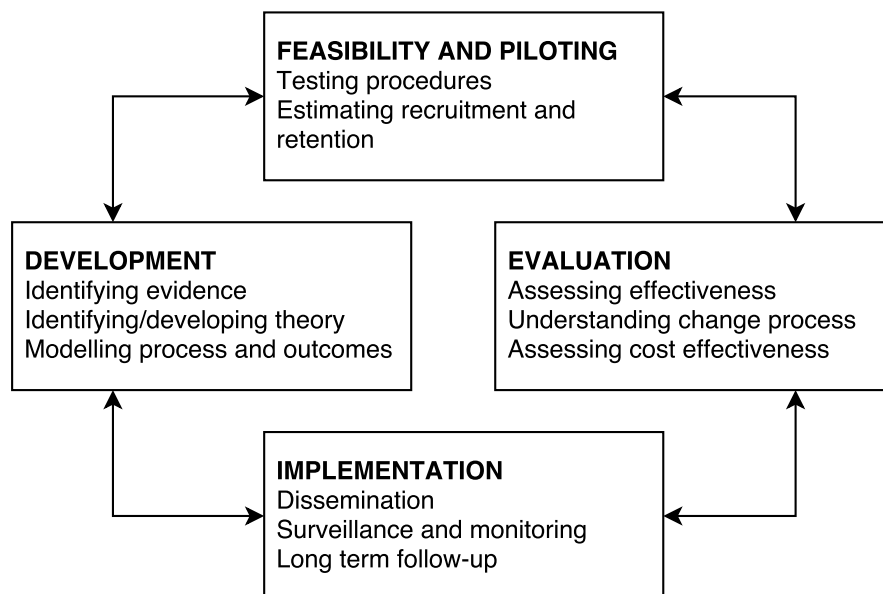


Figure 1.9 MRC Framework for the development and testing of complex interventions (Craig et al., 2008)

The steps outlined in these frameworks are important: unless time is taken in the early stages to iteratively incorporate evidence, theory, and pilot test data into an intervention, larger effectiveness trials often fail to yield useful information about the processes and mechanisms that might enable others to replicate the causal pathways responsible for their results (Craig et al., 2008).

1.5 Justification for Research

Though widely promoted as a means to provide primary healthcare in resource poor settings as early as the 1978 Alma-Ata Declaration, CHWs have recently

become the subject of renewed interest and debate in low- and middle-income countries (LMICs) facing a growing human resource crisis (Gilmore & McAuliffe, 2013; Lehmann & Sanders, 2007).

The World Health Organization estimates the world is currently short 7.2 million healthcare workers and that this deficiency is expected to grow to 12.9 million by the year 2035 (Campbell et al., 2013). This phenomenon is particularly severe in sub-Saharan Africa, which bears an estimated 24% of the world's burden of illness, yet has only 3% of the health care workers and 1% of the worldwide financial resources for health care (Anyangwe & Mtonga, 2007). Most acute in rural areas, personnel shortages mean over one billion people around the world go their entire lives without seeing a health worker (Ghani, 2011; Moszynski, 2011). This lack of human resources for health (HRH) significantly impeded progress toward the realisation of health-related Millennium Development Goals, namely to reduce child mortality (Goal 4), improve maternal health (Goal 5), and combat HIV/AIDS, malaria and other diseases (Goal 6) (Nullis-Kapp, 2005; United Nations, 2015); WHO officials have already noted that the broad health agenda proposed in the new Sustainable Development Goals has yet to be matched by likewise ambitious investment in the global health workforce (Open Working Group of the General Assembly on Sustainable Development Goals, 2014; Sidibe & Campbell, 2015).

CHWs have been proposed as a way to fill this gap by extending services to hard-to-reach populations in remote areas. The recent high-level call by a global Technical Task Force to train one million CHWs in Africa is but one example of the renewed interest in CHWs (Lewin et al., 2010; Singh & Sachs, 2013; World Health Organization et al., 2008). Yet recent ethnographies and cross-sectional surveys of programmes in several countries have noted high rates of absenteeism, substandard

health knowledge, poor diagnostic capacity, and inconsistent reporting among CHWs—reminiscent of observations made during the first wave of CHW scale-up in the 1970s and 1980s (Alam et al., 2012a; De La Torre et al., 2012; Medhanyie et al., 2012; Olang'o et al., 2010; Pongvongsa et al., 2011; Puchalski Ritchie et al., 2012).

Mounting pressure to meet ambitious new international health goals and avoid repeating mistakes of the past, indicates an urgent need to the need to better understand the efficacy, types, and performance of community health workers (CHWs). Despite international enthusiasm for a more pronounced investment in CHW programmes as a strategy for accelerating health progress, it remains the case that, as Frankel (1992) noted two decades ago, our knowledge of the optimisation of CHW delivered interventions remains limited (Strachan et al., 2012).

In order to identify what CHWs actually should or should not do, for whom and under which circumstances, this thesis takes a phased approach to intervention development that seeks to answer both first- and second-generation questions. These are detailed in the next section.

1.6 Research Questions and Structure of the Thesis

1.6.1 Research Questions

Following the logic of early stage intervention development, this thesis has an iterative and developmental structure in which each section flows out of and builds on the previous one. In order to develop more effective and efficient CHW programmes, this research has three objectives:

1. Assess the effects of community health worker interventions for primary health outcomes in low and middle-income countries (LMICs). Such a user-friendly front end is particularly useful for central planning purposes (Ministry of Health policy etc.) in an area where evidence has been changing rapidly over the past several

years and is scattered across many sources. The lack of definitional clarity in CHW characteristics and programme features observed in undertaking this work gave rise to objective number two.

2. Develop a comprehensive taxonomy of CHW characteristics and programme design features. This would aid trialists in reporting relevant data and reviewers in extracting and comparing outcomes across primary trials. Systematically disaggregating and conceptualising CHW types would provide the necessary language for theory building and optimisation of CHW programmes through component selection experiments. The need to gather information on the relationship between outcomes and individual intervention components so as to support the design and implementation of CHW programmes gave rise to objective three.

3. Identify what interventions for improving the performance of community health workers, if any, may result in better primary healthcare outcomes in LMICs. This is a critical step in optimising CHW-led primary healthcare delivery.

1.6.2 Structure of the Thesis

The thesis is structured as follows: question one is addressed in Chapters 2-4, question two is addressed in Chapter 5, and question three is addressed in Chapter 6.

Chapter 2 introduces overviews of reviews. The systematic identification and synthesis of existing guidance concerning overview practice, however, highlights outstanding methodological challenges in overview practice and the degree to which methodological recommendations conflict. As a result of these findings, the chapter proposes a pilot decision tool to assess the conditions under which employing an overview of systematic reviews is likely to lead to a high risk of bias. Chapter 3 reports the protocol of an overview of reviews to determine what types of care CHWs can deliver safely and effectively (i.e. for what populations with what illness). The

decision tool is then applied in order to determine whether the effects of community health worker interventions for primary health outcomes in LMICs are best assessed in the context of an overview or a systematic review. Given that the conditions to proceed with an overview outlined by the decision tool were not met, an additional systematic search was carried out for primary trials on CHW effects. Chapter 4 reports the amended review protocol and the results of an 85 trial systematic review of CHW efficacy and effectiveness. While the methodological contribution of the overview decision tool has ramifications beyond the development of CHW programmes and was not a primary aim of the thesis, it was of critical importance in setting the methodological direction for answering substantive question one.

Chapter 5 builds on the preceding chapters to generate a formal community health worker taxonomy that enables comprehensive reporting and standardised coding of CHW characteristics and programme features. This taxonomy is intended to facilitate testing and optimisation of programme components and ultimately, better uptake and replication of efficacious CHW programme models and interventions. In order to better understand what CHW or programme characteristics enable or mediate effects, Chapter 6 identifies, critically appraises, and summarises head-to-head comparison trials of the same intervention delivered by CHWs with different characteristics or support in a systematic review of interventions to improve CHW performance. Chapter 7 summarises the results of Chapters 2-6, outlines their implications, and concludes the thesis.

Chapter 2: A Scoping Review of Methodological Guidance for Overviews of Reviews

2.1 Introduction

Over the last five years, there has been a proliferation of reviews on CHW effectiveness. It is important to summarise the effects of these to discern (a) what types of care CHWs can deliver safely and effectively (i.e. for what populations with what illness) and (b) what type of CHWs programme designs produce the greatest effects.

Identifying and reviewing existing information is the first aim of this thesis and the first step in a phased approach to intervention development that has its conceptual roots in the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011) and the MRC Framework for Developing and Evaluating Complex Interventions (Craig et al., 2008). While CHWs typically deliver medical interventions in combination (Paintain et al., 2014; Puett, Coates, Alderman, Sadruddin, & Sadler, 2012), first ascertaining whether and how they can do each safely and efficaciously is important for intervention design and optimisation. The logic of overviews and outstanding methodological debates regarding their conduct are reviewed here.

2.2 Foundations and Logic of Overviews of Reviews

Overviews of reviews compile information from multiple systematic reviews to provide a comprehensive synthesis of evidence (Higgins & Green, 2011). While systematic reviews bring together evidence from multiple studies in a rigorous fashion for a single intervention or group of interventions, a limitation of such systematic reviews as they have evolved in the healthcare field is that they are often narrow in scope (Hartling, Vandermeer, & Fernandes, 2014). They typically focus on direct pairwise comparisons and lack formal comparisons of competing interventions for a given condition or evidence on multiple applications of the same intervention (Hartling et al., 2014). Overviews facilitate these broad comparisons and are similar to systematic reviews in that they use “systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data” in order to answer clearly formulated questions that facilitate evidence-informed decision-making, but differ in their unit of synthesis: systematic reviews rather than primary studies (Hartling et al., 2014; Higgins & Green, 2011; The Cochrane Collaboration, 2016).

Overviews examine: (a) multiple interventions for the same condition or population (e.g. interventions to manage pain in labour: Jones et al., 2012); (b) the same intervention for different conditions or populations (e.g. the efficacy of occupational therapy for different patient groups: Steultjens, Dekker, Bouter, Leemrijse, & van den Ende, 2005); (c) multiple outcomes of the same intervention for the same condition or population (e.g. financial incentives for various behavior changes among healthcare professionals: Flodgren et al., 2011); or (d) adverse effects from the same intervention across multiple conditions (e.g. the adverse effects of biologics: Singh Jasvinder et al., 2011). Nomenclature for this review type varies

(Hartling, Chisholm, Thomson, & Dryden, 2012a); overviews are referred to as ‘umbrella reviews,’ ‘meta-reviews,’ ‘overviews of systematic reviews,’ ‘reviews of reviews’ and ‘systematic review of systematic reviews,’ among others; each of these terms has its adherents and has been used in recent publications (e.g. Hill & Ryan, 2014; Lamming et al., 2014; Martínez-González, Berchtold, Ullman, Busato, & Egger, 2014; O'Donnell et al., 2014; Theodoratou, Tzoulaki, Zgaga, & Ioannidis, 2014). This thesis will use the term overview.

While overviews are not new and exist in fields other than healthcare (e.g. Meltzer, 1972), there remains a lack of clarity around overview methodology (Cochrane Comparing Multiple Interventions Methods Group, 2012; Pieper, Buechter, Jerinic, & Eikermann, 2012). Given that the number of overviews published per year between 2004 and 2011 has risen fivefold, from six to thirty (Pieper et al., 2012), and that overviews of reviews are the most downloaded content in certain evidence synthesis journals (e.g. Thomson, 2014), it is critical that this methodological haziness be addressed so as to avoid dissemination of potentially misleading overview results.

In order to identify and synthesise existing guidance concerning overview practice, a scoping review was conducted (PROSPERO 2015:CRD42015027592). This is the first attempt to systematically summarise such guidance.

2.3 Methods

All stages of this scoping review were undertaken by the DPhil candidate alone, except where indicated. While scoping reviews are typically conducted in teams, this approach was taken to demonstrate for the purposes of university examination the DPhil candidate's own and sole expertise in scoping reviews.

Additional expertise will be sought to review the analysis ex post and prior to publication.

2.3.1 Inclusion and Exclusion Criteria

Studies in any language were included if their primary focus was to describe or develop methodology for overviews (i.e. “systematic reviews of systematic reviews”). Studies were excluded if they were published before Jan 1, 2000 as (a) very few overviews were published before this and (b) any guidelines that were formulated prior to that time are very likely to be out of date.

2.3.2 Systematic Search Strategy

Electronic searches

Searches were run on Oct 27 2015 on MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations, PsycINFO, and Embase, all on the Ovid platform. The following strategy was developed to achieve a high level of precision and specificity without sacrificing sensitivity.

1. (overviews adj3 review*).m_titl.
2. (umbrella adj5 reviews).m_titl.
3. (meta-reviews or (meta adj reviews)).m_titl.
4. (review adj reviews).m_titl.
5. "systematic reviews of systematic reviews".m_titl.
6. Or/1-5
7. (overview adj3 review*).m_titl.
8. (umbrella adj5 review).m_titl.
9. (meta-review or (meta adj review)).m_titl.
10. "systematic review of systematic reviews".m_titl.

11. Or/7-10

12. (11 adj7 (challenges or method* or how or approach or conduct*)).m_titl

13. 6 or 12

14. limit 13 yr="2000-2015"

Searching other sources

In addition to a systematic search of electronic databases using a pre-determined search strategy, studies were identified through:

1. Contacting study authors to solicit relevant unpublished papers, on-going research, and suggestions for other contacts
2. Examining reference lists of included studies for other studies that meet the inclusion criteria
3. Searching webpages of leading evidence-synthesis experts: Agency for Healthcare Research and Quality (AHRQ), Cochrane Collaboration, Campbell Collaboration, Canadian Agency for Drugs and Technology in Health (CADTH), Danish Centre for Health Technology Assessment (DACEHTA), Institute for Quality and Efficiency in Health Care (IQWiG), National Institute for Health and Care Excellence (NICE), & York Centre for Reviews and Dissemination: Systematic Reviews (CRD)

2.3.3 Selection of Studies

Titles and abstracts were independently screened by two authors in order to identify studies that met the eligibility criteria. Both authors then retrieved the full text of studies selected as meeting or possibly meeting the criteria and verified them again in order to produce a final list of included studies. Any discrepancies were discussed until agreement was reached.

2.3.4 Data Extraction

For each paper, the following information was extracted by one reviewer into Microsoft Excel: author, year, title, overview methodological challenges delineated, methodological recommendations, and how recommendations were derived. Results were then narratively synthesised by examining commonalities and differences across included studies. In order to improve data quality, a second reviewer audited extracted data.

2.3.5 Data Synthesis

This review delineates (a) areas highlighted as outstanding challenges by overview methodology papers and (b) areas that have not been identified as challenges but where recommendations conflict.

Further to systematically summarising overview methods guidance, this chapter operationalises relevant methodological debates into a decision making tool that will help improve the quality of overviews produced. This was an addition to the methods described in the review protocol; such a tool was believed to be the most practical way to summarise outstanding issues with included guidance.

2.4 Results

Excluding duplicates, a total of 48 records were screened for inclusion. Of these, 36 were recovered from electronic database searches, 2 from handsearching, and 9 from references. A total of 31 full text papers were considered for inclusion in the scoping review; 24 papers reporting 22 guidelines or descriptive studies met the inclusion criteria (Fig 2.1) (Aromataris et al., 2014; Aromataris et al., 2015; Baker, Costello, Dobbins, & E, 2014; Becker & Oxman, 2011; Caird, Sutcliffe, Kwan, Dickson, & Thomas, 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Conn & Coon Sells, 2014; Cooper & Koenka, 2012; Hartling et

al., 2012a; Hartling et al., 2014; Ioannidis, 2009; Li et al., 2012; Pieper, Antoine, Mathes, Neugebauer, & Eikermann, 2014a; Pieper, Antoine, Morfeld, Mathes, & Eikermann, 2014b; Pieper, Antoine, Neugebauer, & Eikermann, 2014c; Pieper et al., 2012; Piso et al., 2015; Ryan, Kaufman, & Hill, 2009; Santaguida et al., 2013; Smith, Devane, Begley, & Clarke, 2011; Thomson et al., 2013; Thomson, Russell, Becker, Klassen, & Hartling, 2010; Yuan, Liu, Yang, Chen, & Yang, 2012).

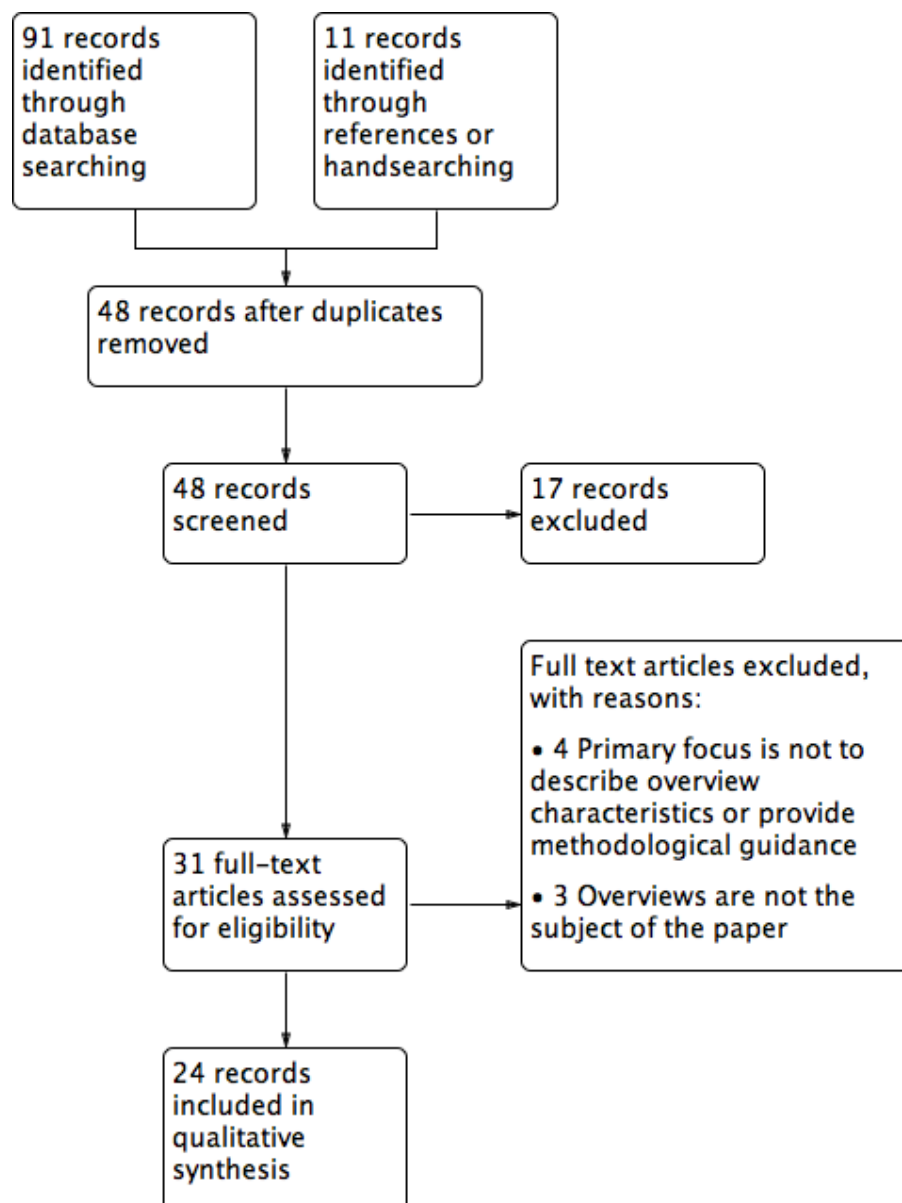


Figure 2.1 PRISMA Study Flow Diagram

2.4.1 Characteristics of the Sample

Of the papers included in the scoping review, six described four sets of methodological guidelines assembled by a working group or editorial team (Aromataris et al., 2014; Aromataris et al., 2015; Becker & Oxman, 2011; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Conn & Coon Sells, 2014); ten were narrative reviews or descriptions of authors' anecdotal experiences conducting overviews (Baker et al., 2014; Caird et al., 2015; Cooper & Koenka, 2012; Hartling et al., 2014; Ioannidis, 2009; Piso et al., 2015; Ryan et al., 2009; Santaguida et al., 2013; Smith et al., 2011; Thomson et al., 2010); seven were systematic reviews in which authors used a pre-specified, comprehensive search strategy to assemble all overviews published during a given date range in order to describe their methodological characteristics and reporting (Hartling et al., 2012a; Li et al., 2012; Pieper et al., 2014a; Pieper et al., 2014c; Pieper et al., 2012; Thomson et al., 2013; Yuan et al., 2012); and one was a systematic review of HTA guidelines (Pieper et al., 2014b). Papers are listed by type in Table 2.1 and the objective and methodological issues addressed by each summarised. Excluded papers (Becker, 2006; Delgado-Rodríguez, 2006; Pieper, Mathes, & Eikermann, 2014d; Robinson et al., 2014; Sarrami-Foroushani, Travaglia, Debono, Clay-Williams, & Braithwaite, 2015; Thomson, 2014; Whitlock, Lin, Chou, Shekelle, & Robinson, 2008), along with reasons for exclusion, are listed in Appendix 2a.

Table 2.1 Characteristics of Included Studies

Reference	Objective	Inclusion criteria	Search strategy	Data extraction and analysis	Quality: individual studies	Quality: systematic reviews	Quality of evidence	Updating	Overlap	Scope
Methodological Guidelines										
Aromataris 2014, JBI 2014	Provide methodological guidance for the conduct of JBI overviews	✓	✓	✓	✓	✓	✓	✓	✓	
Becker 2011	Provide guidance to authors for the preparation of Cochrane Overviews	✓	✓	✓		✓	✓	✓		
CMIMG 2013, CMIMG 2012	Present a set of recommendations for comparing multiple interventions	✓	✓	✓	✓	✓		✓	✓	✓
Conn 2014	Provide guidance for authors contemplating preparation of overviews for WJNR	✓	✓	✓	✓	✓			✓	✓
Narrative Reviews and Descriptions of Authors' Anecdotal Experiences										
Baker 2014	Describe the first overview of the Cochrane Public Health Group (CPHG) and propose a way forward for the methodologies under consideration	✓	✓	✓		✓	✓			
Caird 2015	Highlight the challenges inherent in the overview method and present, where possible, potential solutions to these challenges	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cooper 2012	Describe what methods overviews have developed or have adopted from other forms of scholarship.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hartling 2014	Describe and compare the contributions made by, and the challenges of producing three types of evidence synthesis	✓	✓	✓	✓	✓	✓	✓	✓	
Ioannidis 2009	Review the main features, strengths and limitations of methods that integrate evidence across multiple meta-analyses.	✓		✓	✓			✓		
Piso 2015	Assess methods and strengths/ limitations of overviews and discuss their applications using examples from two Austrian institutions.	✓	✓	✓	✓	✓			✓	
Ryan 2009	Outline a method for preparing data integration tables to enable review-level evidence synthesis			✓	✓	✓				
Santaguida 2013	Describe the methodology used to conduct an overview	✓	✓	✓		✓		✓	✓	
Smith 2011	Provide a guide to clinicians and researchers who wish to conduct an overview.	✓	✓	✓		✓	✓			
Thomson 2010	Discuss outstanding challenges for overview authors and the limitations of overviews.	✓	✓	✓	✓		✓	✓	✓	✓
Systematic Reviews of Overview Methodology and Reporting or HTA Guidelines										
Hartling 2012	Inform the methodology for overviews by 1) identifying all overviews of health care interventions published 2000-2011; 2) describing the overviews with respect to their methodological approaches.	✓	✓	✓	✓	✓	✓			
Li 2012	To evaluate the reporting and methodological quality of overviews of systematic reviews.	✓	✓			✓	✓			
Pieper 2012	Examine published overviews of systematic reviews in terms of descriptive and methodological characteristics.	✓	✓			✓		✓	✓	
Pieper 2014A	Summarize Health Technology Assessment (HTA) agency guidance for and applied methodology in overviews		✓	✓		✓	✓		✓	
Pieper 2014B	Examine published overviews of systematic reviews in terms of updating techniques, publication lag in overviews and to develop recommendations for authors of overviews.	✓	✓				✓	✓	✓	
Pieper 2014C	Determine if the authors mention overlapping reviews in overviews and calculate the actual overlap in published overviews using newly introduced, validated measures.								✓	
Thomson 2013	Describe some aspects of the approaches taken in EBCH overviews	✓	✓					✓		✓
Yuan 2012	To survey and analyze the quality assessment of the included studies in overviews				✓	✓	✓			

Included records specified that overviews serve the following three functions:

1. To identify gaps in the literature where multiple comparable studies may exist but a research synthesis has not been performed (Caird et al., 2015; Cooper & Koenka, 2012; Piso et al., 2015; Ryan et al., 2009; Santaguida et al., 2013)
2. To compare and contrast existing systematic reviews (Aromataris et al., 2014; Aromataris et al., 2015; Baker et al., 2014; Conn & Coon Sells, 2014; Cooper & Koenka, 2012; Pieper et al., 2014a; Pieper et al., 2012; Santaguida et al., 2013; Smith et al., 2011).
3. To provide a summary of evidence from existing systematic reviews, with or without synthesis (Becker & Oxman, 2011; Caird et al., 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Cooper & Koenka, 2012; Hartling, Chisholm, Thomson, & Dryden, 2012b; Hartling et al., 2014; Li et al., 2012; Pieper et al., 2014a; Pieper et al., 2012; Piso et al., 2015; Ryan et al., 2009; Smith et al., 2011; Thomson et al., 2013; Thomson et al., 2010).

2.4.2 Methodological Guidance and Debates

The majority of included papers noted, either explicitly or implicitly, that many of the methodological standards for systematic reviews could also be applied to overviews (Becker & Oxman, 2011; Pieper et al., 2014a; Pieper et al., 2012; Santaguida et al., 2013). For example, there was a consensus that, as with systematic reviews, overviews ought to set a-priori inclusion and exclusion criteria that specify the purpose of the overview (Cochrane Comparing Multiple Interventions Methods Group, 2013; Smith et al., 2011) and describe features of systematic reviews to be

included in enough detail so as to enable a reader to repeat the selection (Ioannidis, 2009; Li et al., 2012; Santaguida et al., 2013; Smith et al., 2011; Thomson et al., 2010).

Similarly, overview guidelines indicated that search strategies ought to be pre-specified (Aromataris et al., 2014; Becker & Oxman, 2011; Conn & Coon Sells, 2014), incorporate manual searches of key journals and of reference lists of included systematic reviews (Conn & Coon Sells, 2014), and—with the exception of the Cochrane Collaboration (Becker & Oxman, 2011) which notes that overview authors should “typically” search only for relevant Cochrane reviews—include multiple online databases (Aromataris et al., 2014; Conn & Coon Sells, 2014). These recommendations differ from systematic review methodology only in that, rather than cover many decades as is typical in a systematic review search, overview searches rarely need to extend prior to 1990 as all but a small minority of systematic reviews were conducted after that time (Aromataris et al., 2014; Baker et al., 2014; Smith et al., 2011).

There is, however, emerging debate related to (i) overlapping systematic reviews, (ii) the scope of systematic reviews, (iii) evaluating the quality and reporting of included research, (iv) updating included systematic reviews, and (v) synthesising and reporting the results of included systematic reviews.

The questions of whether overviews will save time, when they are likely to produce valid results, and when they are best avoided, will be explored in the following narrative summary.

(i) Overlapping systematic reviews

An included systematic review of overviews found that nearly half (46.7%) of overviews published between 2009-2011 did not mention overlaps (Pieper et al.,

2014a). When included systematic reviews are meta-analysed, using data from individual studies more than once without accounting for overlap will overestimate statistical power and thus risk producing a misleading, overly precise estimate (Pieper et al., 2014c; Smith et al., 2011). When systematic reviews are narratively summarised, similar problem occurs if steps are not taken to ensure certain primary studies are not overrepresented (Caird et al., 2015; Pieper et al., 2014c; Thomson et al., 2010). Several possible approaches to this issue have been proposed, though no consensus has been reached.

A first approach, put forth by Pieper et al. (2014c), is the corrected covered area (CCA)—a validated method to calculate the degree of overlap in an overview. The CCA calculates overlap by dividing the frequency of repeated occurrences of the index publication (the first occurrence of a primary publication) in other systematic reviews by the product of index publications and systematic reviews less the number of index publications (see Box 2.1). The CCA prevents potential pitfalls of simply reporting primary studies in a grid (Cooper & Koenka, 2012): because the numerator only counts primary publications included in more than one systematic review it diminishes the impact of large systematic reviews. While two guidelines (Aromataris et al., 2014, p. 13; Conn & Coon Sells, 2014) note that a “clear indication” of overlap should be presented, reporting overlap does not solve the problem of producing misleadingly precise estimates as it does not eliminate the overlap that can influence the precision of the results.

Box 2.1 Calculation of the Corrected Covered Area (Pieper et al., 2014c)

$$CCA = (N - r) / (rc - r)$$

N = total publications (includes double counting)

r = index publications

c = # of reviews

Rather than simply including all syntheses regardless of overlap, a second approach is to develop a priori criteria for choosing a single systematic review for inclusion when multiple potential candidates are available (Cochrane Comparing Multiple Interventions Methods Group, 2012). Cooper and Koenka (2012) suggest selecting the synthesis that (a) provides the most complete description, (b) is most recent, (c) contains the most evidence, (d) is methodologically most rigorous, or (e) is published (vs. in the grey literature). It is possible, however, that this approach would still result in excluding a systematic review that had relevant studies.

The Cochrane Collaboration's stipulation that their overviews should typically only include Cochrane systematic reviews avoids the issue of overlap and so can be counted as a third approach (Becker & Oxman, 2011). Pieper et al. (2012), however, point out that while surveys of these systematic reviews have demonstrated that Cochrane reviews tend to be of higher quality than non-Cochrane reviews (Jørgensen, Hilden, & Gøtzsche, 2006; Moseley, Elkins, Herbert, Maher, & Sherrington, 2009; Olsen et al., 2001; Shea, Moher, Graham, Pham, & Tugwell, 2002), it is an ecological fallacy to suggest that for this reason they merit being preferentially included. Moreover, as both narrative (Caird et al., 2015) and systematic papers (Pieper et al., 2012) have argued, other systematic reviews may be more up-to-date, have a different focus, include additional data (e.g. on adverse effects), or incorporate new analyses (e.g. on various subgroups) that could increase the value of an overview. Excluding these systematic reviews from an overview would omit potentially important information. Thus, the problem of overlap in overviews currently remains unresolved (Cochrane Comparing Multiple Interventions Methods Group, 2013).

(ii) Scope of systematic reviews

Where the scope of the individual systematic reviews is more heterogeneous and doesn't align as closely with the overview question, several papers note that authors of overviews may have the opposite problem: scope mismatch, in which the scope of the overview does not match that of individual systematic reviews (Fig 2.2) (Caird et al., 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012; Conn & Coon Sells, 2014; Cooper & Koenka, 2012). Thomson et al. (2010) provide the example of overviews about child health; if a systematic review includes both adults and children, it is unclear whether it ought to be included. Where the scope of some systematic reviews lies outside the scope of the overview, authors of such overviews face a dilemma: either they (a) include the systematic review, undertaking the time-consuming process of extracting relevant data from primary studies (Caird et al., 2015; Thomson et al., 2010) or (b) in cases where guidelines do not recommended that authors retrieve primary studies (Aromataris et al., 2014), exclude the systematic review, risking omission of relevant results and thereby potentially introducing bias into the outcome (Caird et al., 2015).

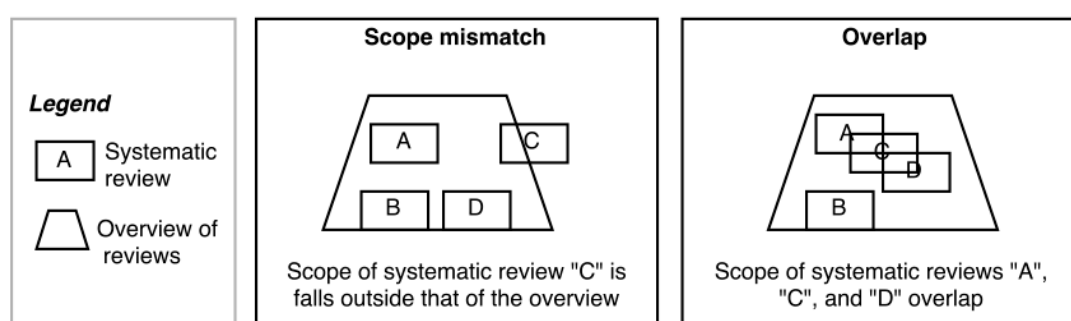


Figure 2.2 Scope mismatch vs. overlap

Exacerbating this difficulty, as Caird et al. (2015) note, systematic reviews do not always report all data from included primary research. This is, of course, impossible to tell in the absence of examining primary trials, yet retrieving and

examining primary trials negates the anticipated efficiency gain of conducting an overview and challenges the definition of an overview as a “systematic review of systematic reviews” (Cooper & Koenka, 2012; Smith et al., 2011, p. 1).

(iii) Evaluating the quality and reporting of included research

Included studies recommend overview authors address three quality assessments: methodological quality of trials included in systematic reviews, included systematic reviews themselves, and evidence across included systematic reviews.

Primary trials

Several papers note that considering the quality-assessment system used in included syntheses and evaluating whether it is appropriate is an important step in conducting an overview (Aromataris et al., 2015; Caird et al., 2015; Conn & Coon Sells, 2014; Cooper & Koenka, 2012; Yuan et al., 2012). The Cochrane Handbook recommends overview authors should report the quality of evidence as assessed by the systematic review authors applying GRADE (Becker & Oxman, 2011).

In the event that an appropriate quality assessment was not used in the original systematic review, overview authors will need to either (a) assess the risk of bias themselves (Becker & Oxman, 2011)—a time-consuming process in which data are extracted at trial level—or (b) include only systematic reviews that have detailed reporting of the quality of primary studies—an approach that increases efficiency by possibly excluding relevant evidence from the resulting evidence base (Caird et al., 2015). This may explain why information on quality of individual studies was extracted from fewer than 40% of overviews over the last decade (Hartling et al., 2012a).

Systematic reviews

While included guidelines agree that included systematic reviews must be assessed for methodological quality (Aromataris et al., 2014; Aromataris et al., 2015; Becker & Oxman, 2011; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Conn & Coon Sells, 2014); there is no consensus on what instrument should be used (Yuan et al., 2012). A descriptive analysis of overviews (Hartling et al., 2012a) found that among overviews published between 2000 to July 2011, only 37% assessed the quality of included systematic reviews at all; the following tools were employed: Oxman and Guyatt Overview Quality Assessment Questionnaire (OQAQ) (n = 9, 27%) (Oxman & Guyatt, 1991), AMSTAR (n = 7, 21%) (Shea et al., 2007a), QUORUM (a precursor to PRISMA) guidelines (n = 3, 9%) (Moher et al., 1999), Jadad decision algorithm for interpretation of discordant reviews (n = 1, 3%) (Jadad, Cook, & Browman, 1997), Methodological Quality Checklist (n = 1, 3%), the Index of the Scientific Quality of Research Overviews (n = 1, 3%) and the Systematic Review Appraisal Sheet of the Centre for Evidence Based Medicine, University of Oxford (n = 1, 3%). Though the use of AMSTAR may be underestimated in these findings, as it was only introduced in 2007, this diversity of instruments and low level of quality reporting was echoed in the other descriptive analyses (Hartling et al., 2012a; Hartling et al., 2014; Li et al., 2012).

There is a vast difference in the purpose and comprehensiveness of the listed instruments. As Pieper et al. (2012) point out, it is important to differentiate between methodological quality, which considers how well the systematic review was conducted, and reporting quality, which considers how well systematic reviewers have reported their methodology and findings. PRISMA/QUORUM are guidelines for reporting systematic reviews; they do not involve judgments regarding risk of bias. While quality of reporting is important—only well-reported systematic reviews allow

users to faithfully recreate the interventions in question and to judge whether the results are applicable to their own settings—it is likewise important to be conceptually clear about the difference between reporting and methodological quality. For this reason, it has been recommended that PRISMA be used in conjunction with a critical appraisal tool (Pieper et al., 2012). Similarly, some items on the list (e.g. Jadad et al., 1997) measure only one dimension of quality, and so also ought to be used in conjunction with more comprehensive instruments. Of comprehensive instruments, only AMSTAR (Shea et al., 2007a) and OQAQ (Oxman & Guyatt, 1991) have been validated.

The question of what to do if systematic reviews are of low quality remains similarly unresolved (Pieper et al., 2014b). Excluding low-quality systematic reviews will likely give an incomplete picture of the data (Caird et al., 2015) and attempting to compensate by performing new searches (Pieper et al., 2014c) has its own difficulties that will be discussed in the next section.

Quality of the body of evidence across included systematic reviews

No method exists to evaluate the strength of evidence for findings across different systematic reviews (Santaguida et al., 2013). While GRADE has been suggested for this task, some GRADE criteria are only applicable to primary studies (Santaguida et al., 2013). For example, GRADE stipulates that quality of evidence can be upgraded in the case of large effect sizes—a criterion not applicable to narratively synthesised systematic reviews. Although they describe an approach to apply GRADE to systematic reviews (Santaguida et al., 2013), as Hartling et al. (2014) note, guidance is needed to ensure appropriate use and interpretation of GRADE when not applied to the primary studies for which the tool was designed. This has yet to be done.

(iv) Updating included systematic reviews

For an overview to be useful for decision-making, underlying systematic reviews must be up-to-date (Cooper & Koenka, 2012; Hartling et al., 2014; Pieper et al., 2014c; Thomson et al., 2010). In a study of how quickly systematic reviews go out of date cited by several included papers (Caird et al., 2015; Pieper et al., 2014c; Pieper et al., 2012), the median duration of survival free of a signal for updating¹ is 5.5 years (CI, 4.6 to 7.6 years), with a signal having occurred for 23% of systematic reviews after 2 years and for 15% of systematic reviews within 1 year (Shojania et al., 2007).

These figures are striking because in their systematic review of overview characteristics, Pieper et al. (2014c) found that the mean publication lag in overviews (defined as publication date of the overview minus publication date of the systematic review) was more than 5 years; meaning that nearly half of overviews are at risk of being out of date on the day they are published. This statistic is made even more troubling by the fact that only one in four overviews considered up-to-dateness and methods for updating systematic reviews had high heterogeneity (Pieper et al., 2014c).

While some guidelines ignore this issue, recommending that authors not undertake an update of included systematic reviews or a new systematic review within the overview (Aromataris et al., 2015; Becker & Oxman, 2011), others propose asking subject specialists to provide relevant primary research (Caird et al., 2015) or considering whether a new search for recent primary studies may be necessary (Pieper et al., 2014b; Pieper et al., 2012). Pieper et al. (2014c) distinguished two ways of

¹ Quantitative signals for updating were defined as changes in statistical significance or relative changes in effect magnitude of at least 50% involving a primary outcome or any mortality outcome; qualitative signals were defined as substantial differences in characterisations of effectiveness, new information about harm, and caveats about the previously reported findings that would affect clinical decision making (Shojania et al., 2007).

searching for primary studies in overviews: (a) updating included systematic reviews and (b) searching for secondary and primary literature simultaneously. Both approaches are also relevant in instances where there is not a systematic review for all interventions relevant to a particular overview (Ioannidis, 2009; Thomson et al., 2010).

Updating included systematic reviews in the context of an overview, however, creates the added complexity of deciding how to apply inclusion/exclusion criteria of the systematic reviews to new trials: it is unclear how to proceed if a trial examines a comparison clinically meaningful to the overview but does not clearly match the systematic review criteria; for example, the population specified (Thomson et al., 2013). Moreover, updating low-quality systematic reviews by searching only the time span not covered by the systematic review might be pointless if it did not initially include a comprehensive search strategy (Pieper et al., 2014c).

Although option (b), performing a search for secondary and primary literature simultaneously solves both problems, it undermines the purported time-savings of conducting an overview and raises the question of why one would undertake an overview, rather than a broad systematic review, in the first place (Pieper et al., 2014c). This is particularly true given that trials introduce complexity into the overview synthesis process (discussed in the next section) that they do not in the context of systematic reviews (Baker et al., 2014). Indeed, it was as a result of confusion generated by the Cochrane Comparing Multiple Interventions Methods Group (2012) attempt to broaden the Cochrane definition of an overview such that individual studies could be included that the CMIMG reversed their decision a year later (Cochrane Comparing Multiple Interventions Methods Group, 2013).

(v) Synthesizing and reporting the results of included systematic reviews

While debates i-iv deal with the trial selection stage, methods for synthesising the results of multiple systematic reviews are also evolving.

Some included guidelines indicate that overview findings should not be re-synthesised, and should be presented in tables or figures (Aromataris et al., 2014; Aromataris et al., 2015; Becker & Oxman, 2011; Ryan et al., 2009). While Caird et al. (2015) argue that such overviews are a useful tool in comparing the consistency of scope and findings across systematic reviews, others have called the qualitative juxtaposition of data from separate meta-analyses “subjective and suboptimal” (Ioannidis, 2009, p. 489) as it facilitates “informal” indirect comparison that may be misleading (Cochrane Comparing Multiple Interventions Methods Group, 2013, p. 6).

The CMIMG (Cochrane Comparing Multiple Interventions Methods Group, 2013) suggests this simple reformatting of summary estimates into a table or figure should only be used if included trials in these systematic reviews meet the transitivity assumption: that two or more interventions can validly be compared using trials that did not compare them directly but which compared one or other of them against a common comparator (Cochrane Comparing Multiple Interventions Methods Group, 2013). If the trials do not meet this assumption, the CMIMG (Cochrane Comparing Multiple Interventions Methods Group, 2013) suggests overview authors present results of included systematic reviews on a review-by-review basis—despite their being no reporting guidelines for systematically presenting data from narrative reviews (Pieper et al., 2014a) and no user testing on the method of presenting data in overviews (Smith et al., 2011).

A second strategy for integrating the results of research syntheses is to employ a formal statistical analysis to make comparisons across interventions (Hartling et al.,

2014). These analyses can incorporate direct (head-to-head) comparisons, indirect comparisons or both (i.e. network meta-analysis - NMA) (Thomson et al., 2010). For example, a re-analysis of a published overview of treatments for childhood nocturnal enuresis cited by Thomson et al. (2010) used NMA to establish internally consistent estimates for all possible pairwise treatment comparisons, thereby correcting the results of the original overview which had been based on a summary table (Caldwell, Welton, & Ades, 2010).

While this proves that an NMA can be performed based on pooled summaries from systematic reviews, this shortcut is rarely available (Cochrane Comparing Multiple Interventions Methods Group, 2013). Firstly, while the enuresis systematic reviews used the same two key outcomes, systematic reviews included in other overviews have not demonstrated the same consistency in their outcome selection (Wang, Lindsley, & Li, 2012). Secondly, a strong form of the transitivity assumption is employed in indirect comparison (i.e. $(A - B) = (A - C) + (C - B)$ and vice-versa) (Cochrane Comparing Multiple Interventions Methods Group, 2013). Even if included systematic reviews use the same outcome measures, appropriate use of NMA requires that authors ensure that the method's underlying assumptions have been met, which is best achieved by examining the primary trials in detail. Ioannidis (2009) and Hartling et al. (2014) caution that merging results from pre-existing systematic reviews is limited by differences in eligibility criteria; outcome selection, reporting, and definition; methods of combining estimates (random vs. fixed effects models); and approaches to missing data, updating, and overlap. Moreover, as Cooper and Koenka (2012) point out, going back to primary trials—or attempting to analyse systematic reviews and primary trials simultaneously—challenges the definition of an

overview as a systematic review of reviews, as overview authors would simply be using the research syntheses as a source of references for their new effort.

Another related issue concerns the methods used to explain heterogeneity (Conn & Coon Sells, 2014). While overviews are uniquely placed to address the issue of discordance, defined as conflicting results of systematic reviews on the same question, Pieper et al. (2014a) found that strategies to deal with discordant systematic reviews were reported in only 5% of overviews. This despite Jadad et al. (1997) having developed a relevant algorithm (summarised in Fig 2.3) nearly two decades ago (Conn & Coon Sells, 2014).

While the way to improve exploration of heterogeneity in overviews is clear, how to appropriately synthesise and report their findings remains debated.

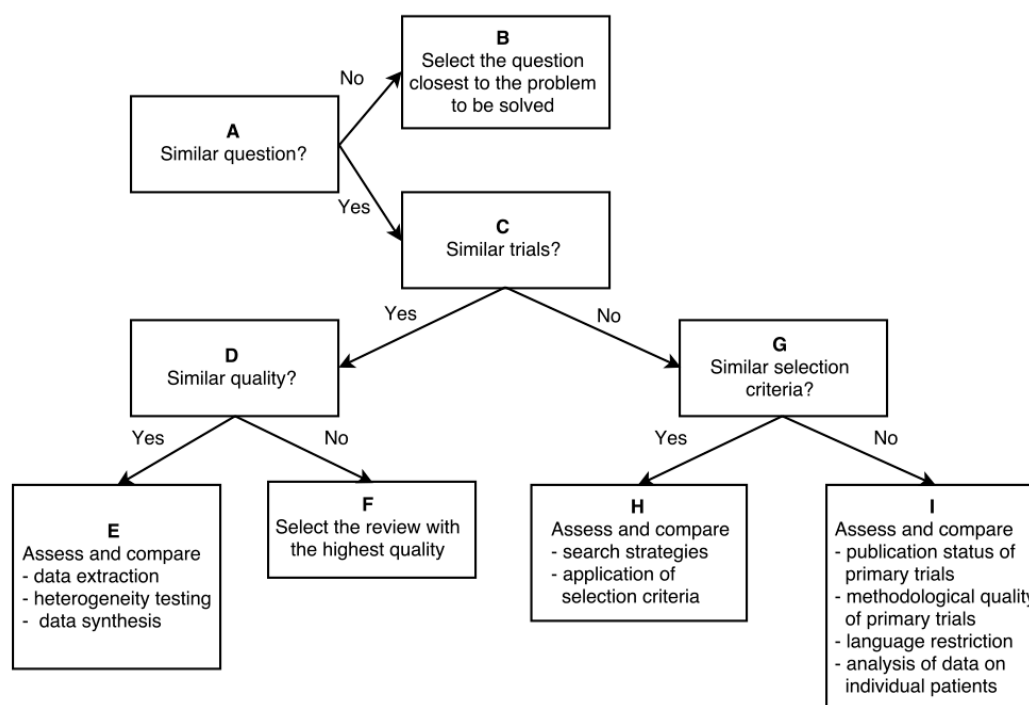


Figure 2.3 Algorithm to identify and resolve causes of discordance between systematic reviews (reproduced from Jadad et al., 1997)

2.5 Discussion

In this chapter, a scoping review was conducted in order to synthesise existing guidance concerning overview practice and delineate areas highlighted as outstanding challenges or where methodological recommendations conflict. Systematic searches identified 102 records, of which 24 met the inclusion criteria. Six described four sets of methodological guidelines assembled by a working group or editorial team (Aromataris et al., 2014; Aromataris et al., 2015; Becker & Oxman, 2011; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Conn & Coon Sells, 2014); ten were narrative reviews or descriptions of authors' anecdotal experiences conducting overviews (Baker et al., 2014; Caird et al., 2015; Cooper & Koenka, 2012; Hartling et al., 2014; Ioannidis, 2009; Piso et al., 2015; Ryan et al., 2009; Santaguida et al., 2013; Smith et al., 2011; Thomson et al., 2010); seven were systematic reviews in which authors used a pre-specified, comprehensive search strategy to assemble all overviews published during a given date range in order to describe their methodological characteristics and reporting (Hartling et al., 2012a; Li et al., 2012; Pieper et al., 2014a; Pieper et al., 2014c; Pieper et al., 2012; Thomson et al., 2013; Yuan et al., 2012); and one was a systematic review of HTA guidelines (Pieper et al., 2014b).

The scoping review demonstrates that, while many of the methodological standards for systematic reviews could also be applied to overviews, the problems of overlap, scope, evaluating the quality and reporting of included research, updating included systematic reviews, and how to synthesise and report findings remain unresolved.

These debates have implications for the potential risk of bias of prospective overviews. Given these outstanding issues, the question remains: when should

overviews be undertaken? Caird et al. (2015) answer the latter question by arguing that overviews are a “solution” when (a) the systematic review addresses a broad research question, (b) results are required within a short timescale, and (c) there are not sufficient resources to conduct a full systematic review in the time available. While Caird et al. (2015) acknowledge the trade-off between comprehensiveness and reliability vs. speed and note that this may be “uncomfortable for reviewers,” they nonetheless conclude that overviews are a useful means of synthesising evidence across broad topic areas (p. 95).

The question of when and how to make such trade-offs, however, requires additional specificity. Current guidance is heavily informed by consensus and personal experience rather than empirical evidence (Smith et al., 2011). Enough is known, however, to differentiate between circumstances in which the efficiency of an overview does not compromise the validity of its outcome and those in which such expediency is likely to introduce serious bias in the form of overly precise estimates and the improper exclusion of relevant data. Clarifying a prospective overview’s intended purpose can elucidate these conditions.

2.5.1 Overview Functions

As explained in the results section, three distinct, not entirely consistent functions have been proposed for overviews: (1) identify gaps in the literature where multiple comparable studies may exist but a research synthesis has not been performed (Caird et al., 2015; Cooper & Koenka, 2012; Piso et al., 2015; Ryan et al., 2009; Santaguida et al., 2013); (2) compare and contrast existing systematic reviews (Aromataris et al., 2014; Aromataris et al., 2015; Baker et al., 2014; Conn & Coon Sells, 2014; Cooper & Koenka, 2012; Pieper et al., 2014a; Pieper et al., 2012; Santaguida et al., 2013; Smith et al., 2011); and (3) provide a summary of evidence

from existing systematic reviews, with or without synthesis (Becker & Oxman, 2011; Caird et al., 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012, 2013; Cooper & Koenka, 2012; Hartling et al., 2012b; Hartling et al., 2014; Li et al., 2012; Pieper et al., 2014a; Pieper et al., 2012; Piso et al., 2015; Ryan et al., 2009; Smith et al., 2011; Thomson et al., 2013; Thomson et al., 2010).

That these purposes are typically listed consecutively (Cooper & Koenka, 2012; Pieper et al., 2012) belies the fact that a different methodological approach is required to achieve each one. This conflation of functions and equivocation of the term overview is the source of considerable methodological confusion.

The methodological conditions of each proposed aim will be considered in turn. A summary of this discussion is presented in Fig 2.5.

(1) Identify gaps in the literature

Contrary to the claim made by some included papers (Caird et al., 2015; Cooper & Koenka, 2012; Piso et al., 2015; Ryan et al., 2009; Santaguida et al., 2013), overviews that fail to find a systematic review for every relevant comparison will not, by default, detect evidence gaps. In contrast to empty systematic reviews, such overviews will not serve an agenda-setting function by directing future research and are likely to be misleading (Cochrane Empty Reviews Project, 2013). For example, consider a situation where there are three possible interventions (A, B, and C) for a particular condition. Though there exist primary trials evaluating each intervention, only intervention A and intervention B have been systematically reviewed (Fig 2.4).

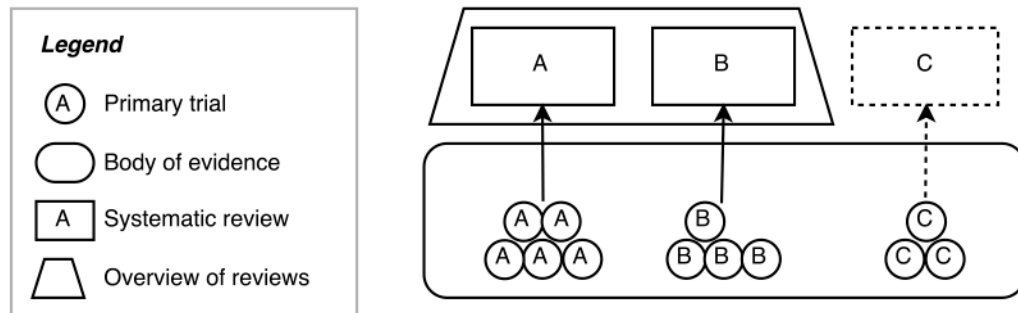


Figure 2.4 Overview in the absence of a systematic review on every relevant comparison

Because conducting an overview does not involve searching for primary trials, an overview of this particular condition would include the systematic review of intervention A and the systematic review of intervention B, but no information on intervention C. While it is possible that authors well-versed in possible intervention options would note the absence of a systematic review for option C, such an observation would be ad hoc, rather than systematic. More critically, the resulting overview would still be incomplete and potentially misleading in that it does not reflect current evidence: though evidence about option C exists, the overview will not capture it.

Identifying gaps in the literature is best done using a systematic review and should not be considered a function of overviews.

(2) Compare and contrast existing systematic reviews

The second proposed function for overviews is to assess whether systematic reviews addressing similar questions independently report similar results and arrive at similar conclusions (Aromataris et al., 2014; Aromataris et al., 2015; Baker et al., 2014; Conn & Coon Sells, 2014; Cooper & Koenka, 2012; Pieper et al., 2014a; Pieper et al., 2012; Santaguida et al., 2013; Smith et al., 2011). As noted above, this process

can be guided by Jadad et al. (1997) algorithm to identify and resolve causes of discordance between systematic reviews.

While overviews are often framed as tools to facilitate broad comparisons (e.g. different interventions for the same condition, the same intervention for different conditions etc.) this function highlights that they can also be used to assess “differences between reviews on the *same* topic” (Baker et al., 2014, p. 518).

This aim is, however, in tension with the third proposed function summarising current evidence. When examining research syntheses focused on the same research problem or hypothesis, a high degree of overlap is likely. As explained above, when included systematic reviews are meta-analysed, using data from individual studies more than once without accounting for overlap will overestimate statistical power and thus risk producing a misleading, overly precise estimate (Pieper et al., 2014c; Smith et al., 2011). When systematic reviews are narratively summarised, a similar problem occurs if steps are not taken to ensure certain primary studies are not overrepresented (Caird et al., 2015; Pieper et al., 2014c; Thomson et al., 2010).

While it is possible to explore heterogeneity and summarise results in the same overview (by, for example, pre-specifying how overlapping systematic reviews explored for heterogeneity will be handled in the synthesis), it is worth distinguishing between overviews intended to summarise current evidence and those intended to resolve discrepancies between systematic reviews on the same research problem.

(3) Summarise existing systematic reviews

Two separate but related ways of achieving the summary function have been proposed. In one, overviews serve as a “friendly front end,” compiling an exhaustive list of systematic reviews relevant to a specific decision and presenting results of included systematic reviews on a review-by-review basis (Cochrane Comparing

Multiple Interventions Methods Group, 2013, p. 2). In the second, overviews integrate or synthesise evidence from existing systematic reviews narratively and/or quantitatively. Regardless of how the results are presented, the aim of such a summary or synthesis is to provide “an overall examination” of “the evidence base on an issue” (Aromataris et al., 2014, p. 6; Piso et al., 2015, p. 301). As explained above, however, overviews may fail to find a systematic review for every relevant comparison and in such cases will not accurately reflect the current evidence base. This raises the question of whether overviews should ever be conducted in the absence of such a search, and, if not, what would differentiate such an overview from a broad systematic review with appropriate subgroupings.

The ability of the overview to provide valuable decision support thus depends on the research question and realistic expectations towards the method. If authors of overviews are confident that a systematic review has been produced for all relevant comparisons they should consider an overview, with the understanding that the only way to test this assumption is to do a systematic search for primary trials.

2.5.2 Overview Decision Tool

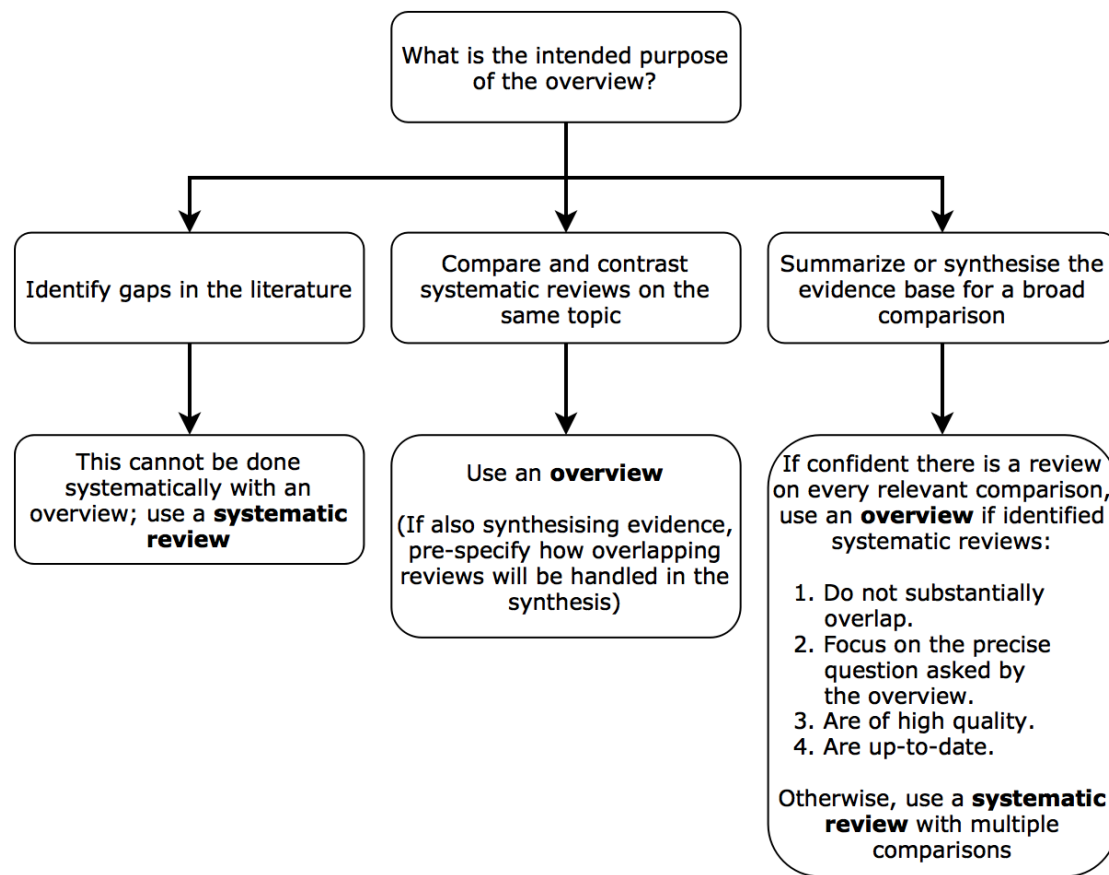


Figure 2.5 When to conduct an overview, a pilot decision tool

When considering whether summarising a body of evidence using an overview is likely to produce useful and accurate results, this scoping review suggest that four further conditions ought to be met. After screening search results against eligibility criteria but before undertaking data extraction or analysis, prospective overviews should ensure identified systematic reviews:

1. Do not substantially overlap.
2. Focus on the precise question asked by the overview.
3. Are of high quality.
4. Are up-to-date.

While operationalising this checklist would require undertaking the full process of guideline development (c.f. Grant, Mayo-Wilson, Melendez-Torres, &

Montgomery, 2013), an explanation of how each was derived from the literature and some suggested cut-offs are offered below. Whereas items 1 and 4 are mentioned in the Comparing Multiple Interventions Methods Group (2013) editorial decision tree for overviews, this discussion adds further detail, offers two additional conditions, and contextualises the checklist in terms of the intended purpose of the overview.

1. The systematic reviews do not substantially overlap

When multiple systematic reviews focus on the same question, there will likely be overlap. Yet there is currently only a way to measure the degree of overlap in an overview—the corrected covered area (CCA) (Pieper et al., 2014c)—not a way to correct for it if it is found to be high. Problems arise if systematic reviews substantially overlap (i.e. get moderate, high, or very high CCA scores) because either, as suggested by Cooper & Koenka (2012), some systematic reviews are excluded—resulting in potential loss of valuable information—or the CCA is noted but all systematic reviews are included—resulting in a misleading, overly precise estimate.

2. The systematic reviews focus on the precise question asked by the overview

Problems likewise arise with systematic reviews that do not focus on the precise question asked by the overview, but contain relevant trials. While the systematic reviews may sometimes run a separate meta-analysis with studies relevant to the overviews, they may sometimes lump all studies together. In the latter case, extracting which studies are relevant will likely require going back to trial level. When systematic reviews that do not focus on the precise question asked by the overview but contain relevant trials not found in other included systematic reviews are found, an overview may not be the best way of synthesising the relevant evidence.

3. The systematic reviews are of high quality

While there is a reliable and valid measure of systematic review quality (Moher, Schulz, Simera, & Altman, 2010) the way in which overview findings should be presented, interpreted, and applied if systematic review quality is moderate or poor (i.e. most systematic reviews do not achieve a score of 8 or higher on AMSTAR) is less clear. While poor quality trials can be excluded from systematic reviews, poor quality systematic reviews may include high quality trials and so ought not to be excluded from overviews. That said, poor quality reviews may not be comprehensive or may have used inappropriate methods to combine study findings and so, when included, produce biased results.

Moreover, as explained above, there is currently no guidance on appropriate use and interpretation of the GRADE tool when assessing the quality of evidence based on systematic reviews rather than primary studies for which the tool was designed.

4. The systematic reviews are up-to-date

If the systematic review is out of date, important evidence may be excluded. While the likelihood of a systematic review being out of date was calculated based on signals of the need to update (defined on p. 68) and not definitive judgments about actual changes in evidence, such signals are important for two reasons. As outlined Shojania et al. (2007), either a formal update that incorporates new evidence might yield conclusions that differ substantially from those of the previous systematic review or in the case that it does not, such an update would explicitly address the new evidence by placing it in context. The latter outcome is particularly important if, for example, a previous systematic review concluded that a treatment was effective and a trial in a high-impact journal concluded that the treatment had no benefit. Because

practitioners might preferentially act on the trial's conclusions because it was more recent or appeared in a high-impact journal, it would be important to reassert the findings of the original systematic review. Unfortunately, there is currently little empirical evidence on when and how to conduct updates of systematic reviews in the context of overviews; as such, if a substantial portion of the systematic reviews appear to be out of date, it may be misleading to proceed with the overview.

Given that the “speed” with which overviews can be conducted is cited as one of their primary advantages (Aromataris et al., 2014; Caird et al., 2015) it should not be onerous to assess these conditions after conducting searches but before full data extraction and analysis. An example of how these conditions might be applied is presented in Chapter 3. In the event that the identified systematic reviews do not meet these criteria the overview should not proceed and a broad systematic review with appropriate subgroupings—in which identified systematic reviews could be used as a source of primary trials—ought to be conducted instead.

2.5.3 Strengths and Weaknesses

Given the rising number of overviews produced (Pieper et al., 2012) and their role as a highly-accessed tool for decision support (Thomson, 2014), addressing risk of bias in this publication type is a priority. This paper is the first to use rigorous, systematic methods to comprehensively summarise current methodological guidance for overviews and delineate areas highlighted as outstanding challenges or where methodological recommendations conflict. It is also the first to consider these methodological debates in terms of the proposed functions of overviews. The result is a decision tool that unpicks the methodological implications of these functions, demonstrating the ways in which they may be unachievable, in tension with each other, or possible only under certain conditions. This tool provides clarity on when an

overview ought to be considered and its conditions can quickly be applied before data extraction and synthesis to ascertain whether a proposed overview ought to proceed.

That said, this work has some limitations that should be taken into account. This tool was constructed solely on the basis of existing literature. Subjecting this decision tool to revision through a consensus process involving experts in the field of overviews together with producers and consumers of overviews could (a) help identify important methodological issues not mentioned in the literature (e.g. aggregation bias) (Pieper et al., 2014b), (b) refine the conditions and cut-offs, and (c) ensure its relevance to and uptake by potential authors of overviews and those who commission them (Grant et al., 2013; Moher et al., 2010).

2.6 Conclusion

This scoping review highlights the lack of guidance regarding many critical facets of overview conduct and delineates five emerging debates. These debates have implications for the potential risk of bias of prospective overviews and can be elucidated by clarifying the intended purpose of overviews. While methods papers routinely list three functions for overviews—identify gaps, explore heterogeneity, summarise evidence—overviews can only perform the latter two. They are unlikely to achieve the second and third simultaneously, and can only perform the third with a low risk of bias when four conditions are met.

Overviews are always susceptible to the interpretive, conceptual, and reporting biases of previous reviewers (Hartling et al., 2014); however, when the four conditions are not met, the overview outcomes risk being biased and misleading. In these cases, a broad systematic review with multiple comparisons would likely produce a more precise result than an overview. Though there may be substantial heterogeneity in the included studies of broad systematic reviews—a reflection of the

diversity of issues considered by the overview and the myriad ways they have been implemented—this can be addressed by careful grouping of studies and selecting for meta-analysis only those subgroups in which the intervention, context and outcomes were sufficiently similar. Such a broad systematic review, while more time consuming than an overview, serves the same essential purpose—summarising a broad evidence base and/or indicating the relative efficacy of different interventions—in a way that is less likely to produce a biased outcome.

While future work is required to fully develop consensus-based guidelines (Moher et al., 2010) these preliminary thoughts are offered for consideration by the field. Attention to the intended function of proposed overviews along with the corresponding methodological conditions may improve the quality of this burgeoning publication type.

This chapter has made a methodological contribution to the first objective of this thesis by developing a tool that will help facilitate the synthesis of evidence and thus early stage intervention development. The next two chapters will also address the first objective by directly assessing the effects of CHW delivered interventions for primary health outcomes in low and middle-income countries.

Chapter 3: Applying the Overview Decision Tool to the Community Health Worker Literature

3.1 Introduction

The aims of this chapter are twofold: (1) to demonstrate the applicability of the overview decision tool on a proposed overview of CHW efficacy and (2) to assess whether the effects of community health worker interventions for primary health outcomes in LMICs are best assessed in the context of an overview or a systematic review.

The proposed overview (PROSPERO 2015:CRD42015016638) was designed to examine the same category of intervention (CHW care) for primary health outcomes in LMICs for different conditions and populations. The purpose of the overview is to summarise or synthesise the evidence base for a broad comparison (effects of lay health worker interventions). The decision tool derived in Chapter 2 was applied and the four conditions under which a “summary/synthesis” overview should proceed were assessed after screening search results against eligibility criteria but before undertaking data extraction or analysis (see Fig 3.1). The chapter first outlines the protocol for the overview and then assesses identified systematic reviews against the four conditions.

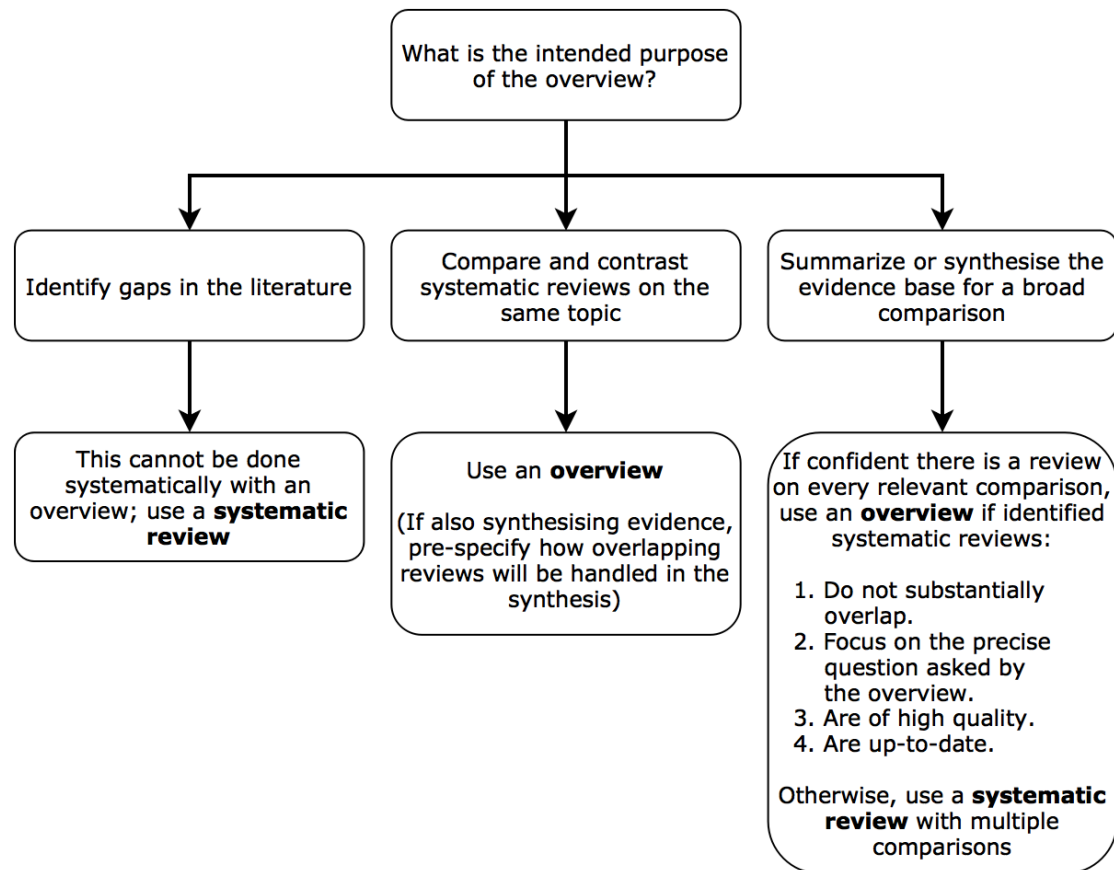


Figure 3.1 Pilot decision tool (reproduced from Chapter 2)

3.2 Methods for the Proposed Overview

3.2.1 Criteria for Including Studies in the Overview

Types of studies

Based on the Cochrane Effective Practice and Organisation of Care (EPOC) guidelines, systematic reviews of randomised controlled trials (RCTs), controlled clinical trials (CCTs), controlled before and after studies (CBAs), and interrupted time series (ITSs) evaluating the effect of lay health worker interventions on the quality of healthcare outcomes were included (Effective Practice and Organisation of Care Group, 2015).

Reviews were only included if they met the following criteria defined by the Cochrane Collaboration:

- a clearly stated set of objectives with pre-defined eligibility criteria for studies
- an explicit, reproducible methodology
- a systematic search that attempts to identify all studies that would meet the eligibility criteria
- a systematic presentation, and synthesis, of the characteristics and findings of the included studies

Reviews were not excluded for failing to assess the validity of the findings of the included studies (e.g. through the assessment of risk of bias), though this was noted in the assessment of their risk of bias. Results within a review covering questions other than effectiveness were not included in the review. Reviews that only searched for and included studies from a specific high-income country were excluded as were reviews published 10 or more years ago, as these are highly unlikely to be up-to-date (Pieper et al., 2014c; Shojania et al., 2007).

Types of participants

There were no restrictions on the types of patients for whom data were extracted aside from residence in an LMIC (as defined by The World Bank, 2014. See Appendix 3a for the country categorisations).

Types of interventions

Systematic reviews of studies evaluating the effects of any type of LMIC-based CHW delivered intervention—including neonatal, child, maternal, and adult healthcare. The definition of CHW employed in this review is the same as that set out in chapter one. CHWs:

- Are primarily based in the community (as opposed to a health facility)

- Belong to the formal health system (i.e. are managed by the government or an implementing NGO)
- Perform tasks related to healthcare delivery
- Have received organised training but have no formal or paraprofessional certification or tertiary education degree

Although CHWs have been used in high-income countries, differences in healthcare budgets, education levels, access to care, and disease profile make it unproductive to include such actors in the comparison.

Types of outcome measures

A number of processes and outcomes might be affected by CHW delivered intervention. Reported outcomes were extracted and categorised as follows; listed outcomes of interest were not used as criteria for including studies.

- Primary:
 - Objective physical outcomes for patients
 - Behavioural health outcomes for patients
 - Harms or adverse effects
- Secondary
 - Utilisation of services indicators (e.g. mean number of consultations/month)
 - Recipient satisfaction with care
 - Costs

3.2.2 Search Methods for Identification of Reviews

Electronic searches

A comprehensive search strategy was used to identify all relevant reviews regardless of language, year, or publication status (published, unpublished, in press, and in progress). We searched the following electronic databases for reviews:

1. 3ie Systematic Reviews and Policy Briefs
2. BiblioMap: EPPI-Centre database of health promotion research
3. Campbell Library of Systematic Reviews
4. CINAHL, Ebsco
5. Cochrane Database of Systematic Reviews (CDSR)
6. Database of Abstracts of Reviews of Effectiveness (DARE)
7. Database of Promoting Health Effectiveness Reviews (DoPHER)
8. Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index – Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities)
9. EMBASE, Ovid
10. Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) Evidence Library
11. Global Health, Ovid
12. Health Technology Assessment
13. MEDLINE In-Process & Other Non-Indexed Citations, Ovid
14. MEDLINE, Ovid
15. NHS Economic Evaluation Database (EED)
16. POPLINE
17. Trip Database

Based on an analysis of thirty-one search filters conducted in MEDLINE, EMBASE and CINAHL (Lee et al., 2012), the most sensitive systematic review filter that best balanced specificity and number needed to read was chosen for each

database. In MEDLINE and CINAHL filters from the Centre for Reviews and Dissemination (Ciliska, Dobbins, & Thomas, 2007) were used, in EMBASE the Wilczynski, Haynes, and Hedges (2007) best optimisation query was used.

The following strategy was used to search MEDLINE and was adapted for other databases:

1. ((community or village) adj2 (agent\$ or aid\$ or promot\$ or mobili?er\$ or distribut?r\$ or worker\$)).tw.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary\$ or extension) adj2 (worker\$ or volunteer\$)).tw.
3. (accompanier\$ or accompagnateur\$ or activista\$ OR animatrice\$ OR brigadista\$ or kader\$ or monitora\$ or promotora\$ or sevika\$ or fhw\$ or chw\$ or lhw\$ or vhw\$ OR chv\$ or "shastho shebika" or "shasto karmis" or anganwadi\$ or "barefoot doctor" OR "agente comunitario de salud" or "agente comunitario de saude").tw.
4. Or/1-3
5. review.ab.
6. review.pt.
7. meta-analysis.ab.
8. meta-analysis.pt.
9. meta-analysis.ti.
10. or/5-9
11. letter.pt.
12. comment.pt.
13. editorial.pt.
14. or/11-13
15. 10 not 14
16. 4 and 15

Terms 1-4 search for the intervention of interest, terms 5-15 searched for the study design. Search terms for all databases can be found in Appendix 3b.

The reference lists of all included reviews were searched. Authors of included papers were contacted regarding any further published or unpublished work.

Searching other sources

In addition to a systematic search of electronic databases using a pre-determined search strategy, studies were identified through:

1. Handsearch of (a) seven journals from 1990 (or first online issue) to May 2014 (*Health Policy and Planning*, *Journal of Community Health*, *Social Science and Medicine*, *Human Resources for Health*, *BMC Public Health*, *Global Health Action*, and *Global Public Health*) and (b) conference proceedings from the American Society of Tropical Medicine and Hygiene.
2. Review of websites of nineteen leading global public health research institutions, donors, implementers, and technical agencies: International Child Health Review Collaboration; World Health Organization; United Nations Children's Fund; United Nations Development Programme; United Nations Population Fund; Joint United Nations Programme on HIV/AIDS; United States Agency for International Development; UK Department for International Human Development Resource Centre; Partners in Health; the Global Fund to Fight AIDS, Tuberculosis and Malaria; Bill & Melinda Gates Foundation; Clinton Foundation; Centre for Global Development; University Research Co-Center for Human Services (URC-CHS); CARE, Frontline Health Workers Coalition; One Million Community Health Workers Campaign; Global Health Workforce Alliance; and CHW Central.
3. Search of grey literature databases relevant to public health: Open GREY, New York Academy of Medicine Grey Literature Report, BIOSIS, and Proquest Dissertations & Theses.
4. Direct contact with study authors and leading experts in the field of community health to solicit potentially relevant unpublished papers, on-going research, and suggestions for other contacts.
5. Examination of the reference lists of related reviews and primary studies to identify other studies that meet the inclusion criteria.

3.2.3 Data Collection and Analysis

Selection of reviews

The author independently screened titles and abstracts in order to identify reviews that met the inclusion criteria. The full text of reviews selected as meeting or possibly meeting the criteria were retrieved and verified a second time in order to produce a final list of included reviews. In keeping with the four conditions of the overview decision tool, instances of scope mismatch were noted.

Data extraction and management

The author extracted the data from reviews into a data extraction form. Data would have been extracted on the stated aim of the review, characteristics of the intervention, CHWs, and the participants; the relevant outcomes; and considerations of applicability and equity. The proposed extraction sheet can be found in Appendix 3c.

Discordant reviews

Jadad et al.'s (1997) algorithm would have been applied to discordant reviews and discordance clearly explained in narrative form.

Overlapping reviews

In keeping with the four conditions of the overview decision tool, the degree of overlap was noted by reporting the corrected covered area (CCA). The CCA calculates overlap by dividing the frequency of repeated occurrences of the index publication (the first occurrence of a primary publication) in other systematic reviews by the product of index publications and systematic reviews less the number of index publications (Pieper et al., 2014a).

Updating included systematic reviews

Each review was assessed to determine if it was up-to-date. In keeping with the four conditions of the overview decision tool, the percentage of identified systematic reviews for which no search had been performed in the previous 24 months was calculated.

Assessment of methodological quality

Included primary research

The certainty of the evidence for the main comparisons would have been assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Guyatt et al., 2008a).

Included systematic reviews

In keeping with the four conditions of the overview decision tool, the methodological quality of each systematic review was assessed using the AMSTAR (A Measurement Tool to Assess Reviews) instrument (Shea 2007a; Shea 2007b; Shea 2009). AMSTAR assesses the degree to which review methods avoided bias by evaluating the methods against 11 distinct criteria: establishing the research question and inclusion criteria before the conduct of the review, data extraction by at least two independent data extractors, comprehensive literature review with searching of at least two databases, key word identification, expert consultation and limits applied, detailed list of included/excluded studies and study characteristics, quality assessment of included studies and consideration of quality assessments in analysis and conclusions, appropriate assessment of homogeneity, assessment of publication bias and a statement of any conflict of interest. Each item on AMSTAR is rated as yes (clearly done), no (clearly not done), cannot answer, or not applicable. The standard scoring system provided by AMSTAR was used to classify reviews into three

categories: those achieving scores of between 8 to 11 were classed as high quality; scores of 4 to 7 as medium quality; and scores of 0 to 3 as be low quality. The author performed quality assessment.

Data synthesis

Review data were organised according to the outcome (Thomson et al., 2010). Meta-analyses would have been attempted if the overview had proceeded.

3.3 Results

3.3.1 Characteristics of Included Reviews

Electronic searches retrieved a total of 3412 references, 1606 references excluding duplicates. Including 16 records identified via other sources, a total of 1622 records were screened for inclusion (Fig. 3.2). Sixty-one records reporting fifty-eight systematic reviews were eligible for inclusion (Ahmed et al., 2014; Amouzou, Morris, Moulton, & Mukanga, 2014b; Arrowsmith, Aicken, Saxena, & Majeed, 2012; Arshad et al., 2014; Bee et al., 2014; Bhutta et al., 2010; Bhutta et al., 2008b; Bigirwa, 2009; Bruxvoort, Goodman, Kachur, & Schellenberg, 2014; Byrne, Hodge, Jimenez-Soto, & Morgan, 2014; Callaghan, Ford, & Schneider, 2010; Christopher, May, Lewin, & Ross, 2011; Corluka, Walker, Lewin, Glenton, & Scheel, 2009; Darmstadt et al., 2009; Das, Lassi, Salam, & Bhutta, 2013; Decroo et al., 2013; Druetz, Siekmans, Goossens, Ridde, & Haddad, 2013; Gilmore & McAuliffe, 2013; Glenton, Khanna, Morgan, & Nilsen, 2013b; Glenton, Scheel, Lewin, & Swingler, 2011b; Gogia et al., 2011; Gogia & Sachdev, 2010; Hall, 2011; Hopkins, Talisuna, Whitty, & Staedke, 2007; Hundley, Avan, Sullivan, & Graham, 2013; Joshi et al., 2014; Kamal-Yanni, Potet, & Saunders, 2012; Kredo, Adeniyi, Bateganya, & Pienaar, 2014; Lassi & Bhutta, 2015; Lassi, Haider, & Bhutta, 2010; Lassi, Haider, & Bhutta, 2012; Lee et

al., 2009; Lee et al., 2014; Lefeuvre, Dieng, Lamara, Raguin, & Michon, 2015; Lewin et al., 2010; Malarcher et al., 2011; Mangham-Jefferies, Pitt, Cousens, Mills, & Schellenberg, 2014; Mannan et al., 2012; Mdege, Chindove, & Ali, 2013; Musa, Iliyasu, Yusuf, & Uloko, 2014; Mutamba, Ginneken, Paintain, Wandiembe, & Schellenberg, 2013; Mwai et al., 2013; Okwundu, Nagpal, Musekiwa, & Sinclair, 2013; Oladapo, Fawole, Blum, & Abalos, 2012; Paintain et al., 2014; Pegurri, Fox-Rushby, & Damian, 2005; Rahman et al., 2013; Ruizendaal et al., 2014; Schiffman, Darmstadt, Agarwal, & Baqui, 2010; Smith, Gubin, Holston, Fullerton, & Prata, 2013; Spaulding et al., 2009; Sunguya et al., 2013; Toczek, Cox, Cros, Cooke, & Ford, 2013; van Ginneken et al., 2013; Velema, Ebenso, & Fuzikawa, 2008; Volmink & Garner, 2007; Vouking et al., 2014; Vouking, Tamo, & Mbuagbaw, 2013; Wiley-Exley, 2007; Wouters, Van Damme, van Rensburg, Masquillier, & Meulemans, 2012).

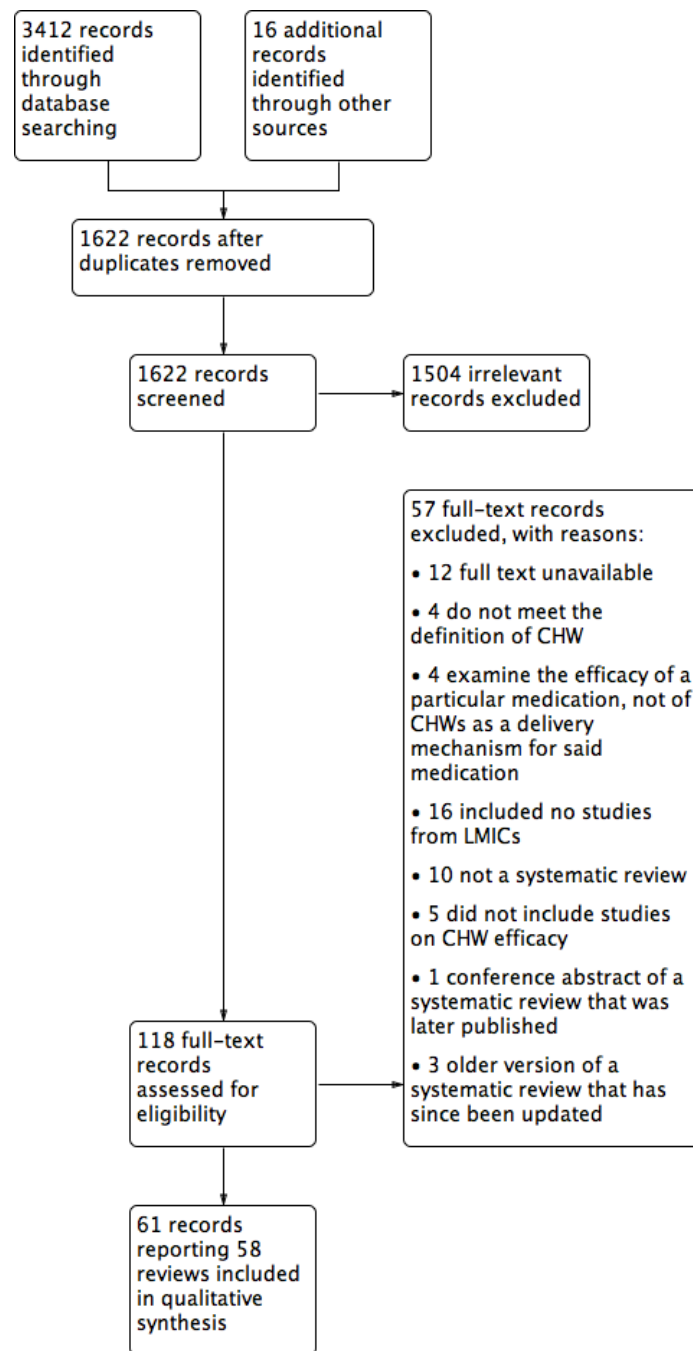


Figure 3.2 Overview PRISMA flow diagram

3.3.2 Applying the Checklist

The four conditions under which a “summary/synthesis” overview ought to proceed were considered in turn.

Criterion one, that the reviews do not substantially overlap, was not met. The corrected covered area (CCA), a validated method to calculate the degree of overlap

in an overview was employed as per the protocol (Pieper et al., 2014a). A CCA value lower than 5 can be considered as a slight overlap, values 6-10 moderate overlap, 11 to 15 high overlap, and greater than or equal to 15 very high overlap. It was determined that overlap within related groupings of reviews was moderate to very high (Table 3.1). The citation matrix on which these calculations are based is included in Appendix 3d.

Table 3.1 Corrected Coverage Area

Review Grouping	Corrected Coverage Area (CCA)
Child Health	Very high (15.9)
HIV	Very high (66.7)
Maternal and Newborn Health	High (11.4)
Mental, Neurological and Substance Abuse	Moderate (6.9)
Noncommunicable Diseases	Very high (50)
Nutrition	Very high (16.7)
Tuberculosis	Very high (42.2)

Criteria 2-4 are summarised in Table 3.2. Criterion two, that the reviews focus on the precise question asked by the overview, was not met: over half of the systematic reviews (32/58) contained studies with no relevance to the question alongside studies that were relevant.

Criterion three, that the reviews are of high quality, was not met. Of the 58 included reviews, less than a quarter (13/58, 22%) were rated as high quality (scoring 8 or more points on AMSTAR) (Shea et al., 2007b). The majority of reviews (34/58, 59%) were rated as moderate quality (scoring between 4 and 7 points), and the remaining reviews (11/58, 19%) were rated low quality (scoring 3 points or less), suggesting that these reviews may be at risk of bias that might influence the results. Common reasons for a review to lose points on AMSTAR were: not reporting conflicts of interest for both the review authors and in the review's included studies

(94%); not assessing publication bias (i.e. the possibility that results from the review are biased because the published literature is systematically unrepresentative of the population of completed studies; 75%); not listing both excluded and included studies (72%); not pre-specifying the question and exclusion criteria (57%); and not using study quality appropriately in formulating conclusions (50%). The complete scores for each checklist item by systematic review are presented in Appendix 3e.

Criterion four, that the reviews are up-to-date, was not met. Fewer than one in five reviews have been updated in the last two years (11/58, 19%).

Table 3.2 Overview criteria 2-4

Review ID	CRITERION #2: All included studies relevant to overview?	CRITERION #3: Of high quality? (AMSTAR score)	CRITERION #4: >2 years since last assessed as up-to date? (month/year)
Ahmed 2014	No (non CHWs)	7	No (09/2013)
Amouzou 2014	Yes	6	No (09/2013)
Arrowsmith 2012	No (non CHWs, non LMIC)	10	Yes (06/2011)
Arshad 2014	No (non CHWs, non LMIC)	4	No (09/2013)
Bee 2014	No (non CHWs)	10	Yes (N.S./2012)
Bhutta 2008	No (non CHWs)	8	Yes (07/2007)
Bhutta 2010	Yes	3	Yes (N.S.)
Bigirwa 2009	Yes	3	Yes (2009)
Bruxvoort 2014	Yes	1	Unclear (N.S)
Byrne 2014	No (non CHWs)	5	Yes (10/2012)
Callaghan 2010	No (non CHWs)	7	Yes (05/2009)
Christopher 2011	Yes	8	Yes (N.S./2007)
Corluka 2009	No (non LMICs)	7	Yes (02/2008)

Review ID	CRITERION #2: All included studies relevant to overview?	CRITERION #3: Of high quality? (AMSTAR score)	CRITERION #4: >2 years since last assessed as up-to date? (month/year)
<i>Continued from previous page</i>			
Darmstadt 2009	No (non CHWs)	4	Yes (N.S./2009)
Das 2013	No (non CHWs)	5	Yes (11/2009)
Decroo 2009	Yes	4	No (02/2013)
Druetz 2013	Yes	5	Yes (09/2012)
Gilmore 2013	Yes	6	Yes (06/2012)
Glenton 2011	No (non LMICs)	6	Yes (02/2009)
Glenton 2013	No (empty review)	6	Yes (05/2011)
Gogia 2010	Yes	7	Yes (10/2008)
Gogia 2011	Yes (empty review)	7	Yes (03/2011)
Hall 2011	No (non CHWs)	5	Yes (03/2010)
Hopkins 2007	Yes	4	Yes (31/03/2012)
Hundley 2013	Yes	5	Yes (03/2012)
Joshi 2014	No (non CHWs)	6	No (06/2013)
Kamal-Yanni 2012	No (non CHWs)	1	Yes (N.S.)
Kredo 2014	No (non CHWs)	10	No (03/2014)
Lassi 2010/2012/2015	No (non CHWs)	10	No (05/2015)
Lee 2009	No (non CHWs)	5	Yes (05/2009)
Lee 2014	No (non CHWs)	9	No (05/2014)
Lefevre 2015	Yes	3	Unclear (N.S./2013)
Lewin 2010	No (non LMICs)	10	Yes (02/2009)
Malarcher 2011	Yes (empty review)	5	Yes (N.S.)

Review ID	CRITERION #2: All included studies relevant to overview?	CRITERION #3: Of high quality? (AMSTAR score)	CRITERION #4: >2 years since last assessed as up-to date? (month/year)
<i>Continued from previous page</i>			
Mangham-Jefferies 2014	No (non CHWs)	8	No (10/2013)
Mannan 2012	Yes	2	Yes (N.S.)
Mdege 2013	No (non CHWs)	7	Yes (01/2011)
Musa 2014	Yes	5	Yes (N.S./2012)
Mutamba 2013	Yes	5	No (06/2013)
Mwai 2013	Yes	3	Yes (12/2012)
Okwundu 2013	Yes	10	Yes (09/2012)
Oladapo 2012	Yes (empty review)	4	Yes (01/2012)
Pantain 2014	Yes	5	Yes (05/2013)
Pegurri 2005	No (non CHWs))	4	Yes (12/2001)
Rahman 2013	No (non CHWs)	6	Yes (05/2012)
Ruizendaal 2014	Yes	6	No (10/2013)
Schiffman 2010	No (non CHWs)	1	Yes (N.S.)
Smith 2013	No (non CHWs)	6	Yes (12/2012)
Spaulding 2009	No (non CHWs)	7	Yes (12/2007)
Sunguya 2013	No (non CHWs)	8	Yes (10/2012)
Toczek 2012	No (non CHWs, non LMICs)	4	Yes (05/2012)
van Ginneken 2013	No (non CHWs)	11	Yes (06/2012)
Velema 2008	Yes	3	Yes (N.S.)
Volmink 2007	No (non CHWs, non LMICs)	9	Yes (05/2007)
Vouking 2013	Yes	4	Yes (07/2012)

Review ID	CRITERION #2: All included studies relevant to overview?	CRITERION #3: Of high quality? (AMSTAR score)	CRITERION #4: >2 years since last assessed as up-to date? (month/year)
<i>Continued from previous page</i>			
Vouking 2014	Yes	3	Yes (12/2012)
Wiley-Exley 2007	No (non CHWs)	1	Yes (N.S./2006)
Wouters 2012	Yes	4	Yes (12/2011)

Not one review met all three of criteria 2-4.

3.4 Conclusion

Overviews are always susceptible to the interpretive, conceptual, and reporting biases of previous reviewers (Hartling et al., 2014), however, when these four conditions are not met, the overview outcomes risk being biased and misleading. In these cases, a broad systematic review with multiple comparisons would likely produce a more precise result than an overview. Though there may be substantial heterogeneity in the included studies of broad systematic reviews—a reflection of the diversity of issues considered by the overview and the myriad ways they have been implemented—this can be addressed by careful grouping of studies and selecting for meta-analysis only those subgroups in which the intervention, context and outcomes were sufficiently similar. Such a broad review, while more time consuming than an overview, serves the same essential purpose—indicating the relative efficacy of different interventions, highlighting gaps in knowledge unlikely to have been identified by narrow systematic reviews—in a way that is less likely to produce a biased outcome.

Given that the four conditions were not met, precisely this type of broad systematic review of CHW efficacy was carried out. The protocol and results of this review are reported in the next chapter.

Chapter 4: A Systematic Review of Community Health Worker Efficacy

4.1 Introduction

The application of the overview decision tool in Chapter 3 demonstrated that the effects of community health worker interventions for primary health outcomes in LMICs are best assessed in the context of a systematic review, rather than an overview. Of the systematic reviews to be included in the proposed overview, there existed a moderate to high degree of overlap within related groupings of reviews, less than a third were of high quality, nearly half contained studies with no relevance to the question, and fewer than one in five reviews have been updated in the last two years. Given these characteristics of the identified systematic reviews, compiling them into an overview was likely to result in a high risk of bias and would not have provided a comprehensive synthesis of evidence as intended.

As such, the original overview protocol (PROSPERO 2015:CRD42015016638) was amended to allow for a systematic review to be carried out on the same topic: the effects of community health worker interventions for primary health outcomes in LMICs. This chapter reports the amended protocol and the results of this review. These results pertain directly to the first objective of this thesis.

4.1.1 How the Intervention Might Work

As explained in Chapter 1, the world is currently short 7.2 million healthcare workers and this deficiency is expected to grow to 12.9 million by the year 2035

(Campbell et al., 2013). This worsening human resource crisis is characterised not just by an absolute shortage but by a devastating geographical distribution: sub-Saharan Africa bears an estimated 24% of the world's burden of illness, yet has only 3% of the health care workers and 1% of the worldwide financial resources for health care (Anyangwe & Mtonga, 2007). Most acute in rural areas, personnel shortages mean over one billion people around the world go their entire lives without seeing a health worker (Ghani, 2011; Moszynski, 2011).

When appropriate, rationally redistributing specific interventions from highly qualified health workers to health workers with shorter training and fewer qualifications is theorised to make more efficient use of the available human resources for health (World Health Organization et al., 2008). “Task shifting” repetitive and relatively simple to administer—yet labour intensive—interventions to CHWs is believed to mitigate the effects of the health worker shortage by reducing costs and increasing access to care while providing equivalent biological outcomes (Callaghan, 2010). Moreover, because community health workers are recruited from and work in their own communities, it has been suggested that their unique cultural sensitivity and accessibility may *improve* access to care, patient satisfaction, and biological outcomes when compared to facility based care—particularly for “high touch” interventions like chronic and preventative care (Berman et al., 1987; Perry et al., 2016; Perry et al., 2014).

4.1.2 Why it is Important to do this Review

In an effort to determine what types of healthcare provision can be safely task shifted, a number of reviews have been carried out on CHW effectiveness. This systematic review, however, arose from the reality that an overview of these reviews would have been at risk of bias and would not have provided a comprehensive view

of the evidence. As explained in Chapter 3, existing reviews of CHW effectiveness are of generally poor quality and do not present a comprehensive understanding of the literature. As such, the reviews currently available are of limited use.

It is important to summarise the latest evidence of CHW efficacy to discern (a) what types of care (i.e. for what populations with what illness) CHWs can deliver safely and effectively and, if possible, (b) what type of CHW programmes produce the greatest effect. Such a user-friendly front end is particularly useful for central planning purposes (e.g. Ministry of Health) in an area where evidence is scattered across many sources.

4.2 Methods

All stages of this systematic review were undertaken by the DPhil candidate alone, except where indicated. While systematic reviews are typically conducted in teams, this approach was taken to demonstrate for the purposes of university examination the DPhil candidate's own and sole expertise in systematic reviews. Additional expertise will be sought to review the analysis *ex post* and prior to publication.

The reviewer followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Methodological Expectations of Cochrane Intervention Reviews (MECIR) recommendations when performing this systematic review (see Appendix 4e for a cross-referenced PRISMA Checklist).

4.2.1 Criteria for Including Studies in the Review

Because delivery agents are empirically testable and possible to randomise, randomised controlled trials, cluster-randomised trials, crossover trials, and factorial trials were eligible for inclusion.

4.2.2 Types of Participants, Interventions, Comparators, And Outcomes

These were as specified in the criteria for the overview and remain unchanged.

They are:

Types of participants

There were no restrictions on the types of patients for whom data were extracted aside from residence in an LMIC (as defined by The World Bank, 2014. See Appendix 3a for the country categorisations).

Types of interventions and comparators

We included primary studies evaluating the effects of any type of CHW delivered intervention that takes place in the community, including neonatal, child, maternal, and adult healthcare. The definition of CHW employed in this review is the same as that set out in chapter one. CHWs:

- Are primarily based in the community (as opposed to a health facility)
- Belong to the formal health system (i.e. are managed by the government or an implementing NGO)
- Perform tasks related to healthcare delivery
- Have received organised training but have no formal or paraprofessional certification or tertiary education degree

Unlike previous reviews that have excluded trials in which there are CHWs in both arms (Lewin et al., 2010), this review included trials where CHWs existing in the control arm did not deliver the intervention in question (i.e. the intervention was delivered by CHWs in the intervention arm and a health facility in the control arm, even though CHWs existed in the control areas). This decision was taken to allow for

the inclusion of effectiveness trials, as the programme settings for such trials were typically pre-existing, regional or national CHW programmes (e.g. Lady Health Workers of Pakistan) that covered the entire study area. The review differentiated between such studies, where CHWs in the control arm did not deliver the intervention in question, and those where (a) CHWs of same type with same support delivered different health interventions for the same ailment in different arms (i.e. trials testing the efficacy or effectiveness of different tasks, not the efficacy or effectiveness of CHW delivery versus an alternative) and (b) head-to-head comparisons of CHWs of different types or with different support delivering the same intervention (i.e. trials concerned with optimising CHW performance, not testing the efficacy or effectiveness of CHW delivery versus an alternative).

Although CHWs have been used in high-income countries, differences in healthcare budgets, education levels, access to care, and disease profile make it unproductive to include such actors in the comparison.

Types of outcome measures

A number of processes and outcomes might be affected by CHW delivered intervention; outcomes were not used as criteria for including studies.

4.2.3 Search Methods for Identification of Studies

Electronic searches

A comprehensive search strategy was used to identify all relevant studies regardless of language, year, or publication status (published, unpublished, in press, and in progress). All health-related databases searched in the overviews, plus additional relevant databases from global health and international development, were searched:

1. AMED, Ovid
2. BiblioMap, EPPI-Centre
3. CINAHL, EBSCO
4. Cochrane Central Register of Controlled Trials (CENTRAL)
5. Databases within ISI Web of Science
6. EMBASE, Ovid
7. Global Health, Ovid
8. Impact Evaluation Database (3ie)
9. IPA (Innovations for Poverty Action) Database
10. JOLIS
11. J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database
12. LILACS
13. MEDLINE In-Process & Other Non-Indexed Citations, Ovid
14. MEDLINE Ovid
15. PAHO
16. POPLINE, K4Health
17. PsycINFO, Ovid
18. TRoPHI, EPPI-Centre
19. WHO Global Health Library (AIM (AFRO), IMEMR (EMRO), IMSEAR (SEARO), WPRIM (WPRO), WHOLIS (KMS))
20. World Bank Independent Evaluation Group

The following search strategy was developed by combining relevant terms from all reviews included in the overview²; it was further refined in consultation with a Social Sciences Data Librarian/Information Specialist. The search below was used to search MEDLINE and was adapted for other databases using the appropriate controlled vocabulary, as applicable:

1. Community Health Workers/
2. ((community or village) adj2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)).tw.

² (Ahmed et al., 2014; Amouzou et al., 2014b; Arrowsmith et al., 2012; Arshad et al., 2014; Bee et al., 2014; Bhutta et al., 2010; Bhutta et al., 2008b; Bigirwa, 2009; Bruxvoort et al., 2014; Byrne et al., 2014; Callaghan et al., 2010; Christopher et al., 2011; Corluka et al., 2009; Darmstadt et al., 2009; Das et al., 2013; Decroo et al., 2013; Druetz et al., 2013; Gilmore & McAuliffe, 2013; Glenton et al., 2013b; Glenton et al., 2011b; Gogia et al., 2011; Gogia & Sachdev, 2010; Hall, 2011; Hopkins et al., 2007; Hundley et al., 2013; Joshi et al., 2014; Kamal-Yanni et al., 2012; Kredo et al., 2014; Lassi & Bhutta, 2015; Lassi et al., 2010; Lassi et al., 2012; Lee et al., 2009; Lee et al., 2014; Lefevre et al., 2015; Lewin et al., 2010; Malarcher et al., 2011; Mangham-Jefferies et al., 2014; Mannan et al., 2012; Mdege et al., 2013; Musa et al., 2014; Mutamba et al., 2013; Mwai et al., 2013; Okwundu et al., 2013; Oladapo et al., 2012; Paintain et al., 2014; Pegurri et al., 2005; Rahman et al., 2013; Ruizendaal et al., 2014; Schiffman et al., 2010; Smith et al., 2013; Spaulding et al., 2009; Sunguya et al., 2013; Toczek et al., 2013; van Ginneken et al., 2013; Velema et al., 2008; Volmink & Garner, 2007; Vouking et al., 2014; Vouking et al., 2013; Wiley-Exley, 2007; Wouters et al., 2012).

3. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) adj2 (worker)).tw.
4. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "colaborador voluntario" or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or "drug-kit manager" or anganwadi* or "barefoot doctor" or "distributeurs communautaires" or "relais communautaires" or "agentes polivalentes elementares").tw.
5. ((saude or salud) adj2 (agente or voluntario or tecnicos or promotores)).tw.
6. ((health* or medical*) adj3 (auxiliar* or attendant* or allied or village or volunteer)).tw.
7. 1 or 2 or 3 or 4 or 5 or 6
8. randomised controlled trial.pt.
9. controlled clinical trial.pt.
10. randomised.ab.
11. placebo.ab.
12. drug therapy.fs. (except psychinfo + gh)
13. randomly.ab.
14. trial.ab.
15. groups.ab.
16. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17. exp animals/ not humans.sh.
18. 16 not 17
19. 17 and 18

Terms 1-7 search for the population of interest and terms 8-18 are the Cochrane highly sensitive search strategy for identifying randomised trials. Searches were conducted April 15, 2015 and July 12, 2016. Precise dates of coverage, date last searched, and search strategies used for the other databases listed above can be found in Appendix 4a.

Searching other resources

This section is consistent with what was specified in the criteria for the overview, with the addition of strategy number six:

1. Handsearch of (a) seven journals from 1960 (or first online issue) to May 2014
(*Health Policy and Planning*, *Journal of Community Health*, *Social Science and Medicine*, *Human Resources for Health*, *BMC Public Health*, *Global*

Health Action, and *Global Public Health*) and (b) conference proceedings from the American Society of Tropical Medicine and Hygiene.

2. Review of websites of nineteen leading global public health research institutions, donors, implementers, and technical agencies: International Child Health Review Collaboration; World Health Organization; United Nations Children's Fund; United Nations Development Programme; United Nations Population Fund; Joint United Nations Programme on HIV/AIDS; United States Agency for International Development; UK Department for International Human Development Resource Centre; Partners in Health; the Global Fund to Fight AIDS, Tuberculosis and Malaria; Bill & Melinda Gates Foundation; Clinton Foundation; Centre for Global Development; University Research Co-Center for Human Services (URC-CHS); CARE, Frontline Health Workers Coalition; One Million Community Health Workers Campaign; Global Health Workforce Alliance; and CHW Central.
3. Search of grey literature databases relevant to public health: Open GREY, New York Academy of Medicine Grey Literature Report, BIOSIS, and Proquest Dissertations & Theses.
4. Direct contact with study authors and leading experts in the field of community health to solicit potentially relevant unpublished papers, on-going research, and suggestions for other contacts.
5. Examination of the reference lists of primary studies to identify other studies that meet the inclusion criteria.
6. Given that the unit of analysis in a systematic review is trials rather than systematic reviews, the following trial registries were searched for on-going research: WHO International Clinical Trials Registry Platform (includes:

Australian New Zealand Clinical Trials Registry (ANZCTR); Brazilian Clinical Trials Registry (ReBec); Chinese Clinical Trial Register (ChiCTR); Clinical Research Information Service, Republic of Korea (CRiS); ClinicalTrials.gov; Clinical Trials Registry - India (CTRI); Cuban Public Registry of Clinical Trials (RPCEC); EU Clinical Trials Register (EU-CTR); German Clinical Trials Register (DRKS); Iranian Registry of Clinical Trials (IRCT); ISRCTN Register; Japan Primary Registries Network (JPRN); Pan African Clinical Trial Registry (PACTR); Sri Lanka Clinical Trials Registry (SLCTR); and The Netherlands National Trial Register (NTR)).

4.2.4 Data Collection and Analysis

Selection of studies

Titles and abstracts were screened in order to identify studies that met the eligibility criteria. The DPhil candidate then retrieved the full text of studies selected as meeting or possibly meeting the criteria and verified them again in order to produce a final list of included studies. In order to audit and improve data quality, ten per cent of titles and abstracts were double-screened by a second reviewer; no disagreements occurred.

Data extraction and management

The DPhil candidate extracted data on study design, participants, interventions, outcomes and methodological quality, using a pre-designed data collection form. The extraction sheet was informed by the Cochrane EPOC Group data collection checklist, the Template for Intervention Description and Replication, and the Oxford Implementation Index; it was piloted on the first ten studies (Effective Practice and Organisation of Care Group, 2015; Hoffmann et al., 2014; Montgomery

et al., 2013b). In order to audit data quality, ten per cent of full-text studies were re-extracted by a second reviewer. No discrepancies were found.

Given the diversity of health issues targeted by CHW programmes, interventions were expected to vary widely and so special care was taken to note all relevant features so as to allow for the creation of relevant subgroupings and a careful consideration of the heterogeneity of the data set. Only subgroups in which both the interventions and the outcomes were sufficiently similar were selected for meta-analysis. Where reports were uncertain or included summary measures, reviewers were contacted for clarification. Any consistently underreported data were noted and highlighted along with future research needs in the discussion. Data were entered and managed using Filemaker Pro; a copy of the data collection sheet is included in Appendix 4b.

Assessment of risk of bias in included studies

Risk of bias for all RCTs was assessed using The Cochrane Collaboration's tool for evaluating the risk of bias. This approach involves assessing the risk of bias across six domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential biases (Higgins & Green, 2011). As per guidance in the Cochrane Collaboration Handbook, additional information on (i) recruitment bias, (ii) baseline imbalances, (iii) loss of clusters, (iv) unit of analysis issues and (v) comparability with individually randomised trials was collected for cluster randomised trials. A judgment of low risk of bias, high risk of bias, or unclear risk of bias was made for each domain based on definitions in the Cochrane Handbook (Higgins & Green, 2011). Had there been crossover trials in which both periods were relevant, information on design suitability, the presence of a carry-over effect, whether only first period data are available,

analysis issues, and the comparability of results with those from parallel-group trials would have been collected.

Assessment of the methodological quality of each economic evaluation was guided by the Consensus on Health Economic Criteria (CHEC) checklist (Evers, Goossens, de Vet, van Tulder, & Ament, 2005) as recommended by multiple best practice guidelines for evidence synthesis and economic evaluations (Higgins & Green, 2011; Walker et al., 2012).

The quality of evidence across a body of evidence (i.e., multiple studies with similar interventions and outcomes) was assessed using the GRADE approach (Guyatt et al., 2008b), defining the quality of evidence for each outcome as “the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest” (Higgins & Green, 2011, section 12.2.1).

Data synthesis

In keeping with earlier Cochrane reviews on aspects of CHW efficacy (Lewin et al., 2010; van Ginneken et al., 2013) studies were grouped according to the type of health issue addressed by the CHWs. The ten comparisons that emerged are listed below. In selecting outcomes to report, the Core Outcome Measures in Effectiveness Trials (COMET) Initiative was searched and core outcome sets in each area consulted where available (Williamson et al., 2012). Where possible, objective outcomes have been preferred.

Mental Health Outcomes

1. CHW delivered interventions for treatment of common mental disorders

Maternal and Paediatric Outcomes

2. CHW delivered interventions to reduce mortality and morbidity in children under five
3. CHW delivered interventions to improve maternal and newborn health

4. CHW delivered interventions to promote cognitive development in infants
5. CHW delivered interventions to promote healthy feeding practices

Infectious Disease Outcomes

6. CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support
7. CHW delivered interventions for maintenance of antiretroviral therapy

Public and Preventative Health Outcomes

8. CHW delivered interventions to promote injury prevention in children
9. CHW delivered interventions to promote immunisation uptake and mass drug administration
10. CHW delivered interventions to reduce hypertension

Meta-analysis is the use of statistical methods to summarise the results of independent studies (Glass, 1976). By combining information from all relevant studies, meta-analyses can provide more precise estimates of the effects of an intervention than those derived from the individual studies included within a review (Higgins & Green, 2011). As explained in Higgins & Green (2011), meta-analysis can help:

1. Increase statistical power by combining individual studies that may be too small to detect small effects, thus increasing the chance of detecting a real effect as statistically significant if it exists.
2. Improve precision by deriving an estimate from more information.
3. Allow investigation of the consistency of effect across studies and, if relevant, allow reasons for differences in effect estimates to be explored (see the subsection below on assessing heterogeneity).
4. Settle controversies arising from apparently conflicting studies or to generate new hypotheses.

Meta-analyses result in a point estimate (the summary treatment effect), and a measure of the precision of results (e.g., a 95% confidence interval). These measures

of precision represent the degree of variability or heterogeneity in the results among included studies (Higgins & Green, 2011).

For dichotomous outcomes, the reviewer calculated risk ratios (RR) and standard errors (SE) of the RR for both individual and cluster RCTs. RRs were preferred to odds ratios because event rates were often high and, in these circumstances, odds ratios can be difficult to interpret (Altman, Deeks, & Sackett, 1998). The risk ratio (also, relative risk) is the ratio of the risk of an event in the two groups (Higgins & Green, 2011). For example, a risk ratio of 2 for a treatment indicates that events with treatment are twice as likely than events without treatment. Similarly, a risk ratio of 0.25 is interpreted as the probability of an event with treatment being one-quarter of that without treatment.

For continuous outcomes, the reviewer used the mean difference (MD) or standardised mean difference (SMD). The mean difference (i.e. difference in means) measures the absolute difference between the mean value in two groups (Higgins & Green, 2011). Because the MD is on the same scale of measurement as the clinical outcome measure used in all studies, it is readily interpretable (Faraone, 2008).

The standardised mean difference is used as a summary statistic in meta-analysis when studies assess the same outcome but measure it on different scales (Higgins & Green, 2011). In this circumstance it is necessary to standardise the results of the studies to a uniform scale before they can be combined. SMD scores are interpreted as the number of standard deviation changes in the outcome. Several rules of thumb exist for interpreting SMDs. This study used the following effect sizes: 0.2 represented a small effect, 0.5 a moderate effect, and 0.8 a large effect (Cohen, 1988). Though the reviewer acknowledges the disadvantages of using SMDs (i.e. vulnerability to varying degrees of heterogeneity in the underlying populations and

difficulties in interpretation) the wide variety of instruments used and the fact that minimally important differences have not been established for these outcomes made alternative presentations suggested by Guyatt et al. (2013) unfeasible.

When meta-analysis was conducted, results were represented using forest plots. Forest plots display effect estimates and confidence intervals for both individual studies and meta-analyses (Lewis & Clarke, 2001). In the plots, each study is represented by a block at the point estimate of intervention effect with a horizontal line extending either side of the block (Higgins & Green, 2011). The area of the block indicates the weight assigned to that study in the meta-analysis while the horizontal line depicts the confidence interval (in this review, the 95% level of confidence). The overall meta-analysed measure of effect is represented as a dashed vertical line and plotted as a diamond, the lateral points of which indicate confidence intervals for this estimate.

The reviewer analysed the outcomes for individual RCTs and the adjusted outcomes for cluster RCTs together, using the inverse variance method. The inverse variance method assumes that the variance for each study is inversely proportional to its importance; thus, more weight is given to studies with less variance than studies with greater variance (Kredo et al., 2014). The method allows for the analysis of continuous and dichotomous data and for clustered and non-clustered data to be combined. The choice of whether to use a fixed-effect or random-effects model was based on the extent to which studies had similar PICOS (population, intervention, comparators, outcomes and settings) characteristics. While the fixed-effect estimate and its confidence interval address the question ‘what is the best (single) estimate of the effect?’, the random-effects estimate assumes there to be a distribution of effects, and the estimate and its confidence interval address the question ‘what is the best

estimate of the average effect?’ (Higgins & Green, 2011). No group of studies was sufficiently homogeneous to apply the fixed-effect model. All forest plots were second read for interpretation by an Assistant Professor with epidemiological and statistical training. For the outcomes or studies where heterogeneity made it inappropriate to combine the results quantitatively, a descriptive review is presented.

Where possible, data were re-analysed on an intention-to-treat basis. As in Lewin et al. (2010) beneficial health behaviours were analysed on a worst-case basis (i.e. persons lost to follow up were assumed to be non-adherent). Morbidity and mortality were analysed on a best-case basis (i.e. persons lost to follow-up were assumed to be alive and without morbidity). No assumptions were made about loss to follow-up for continuous data; only results for those who completed the trial were analysed.

For studies that reported continuous data but did not report standard deviations, the reviewer contacted the study authors. If no response was received, theses were calculated from other available data such as standard errors (Higgins & Green, 2011). The reviewer expressed all effect estimates with their 95% confidence intervals (CI).

The reviewer considered only relevant health economics studies conducted alongside effectiveness studies that met eligibility criteria for the effectiveness component of the review (Higgins & Green, 2011). The economics component of this review was conducted according to current guidance on the use of economics methods in the preparation and maintenance of Cochrane reviews (Higgins & Green, 2011). The reviewer classified the included economic evaluations based on an established system (Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005). Briefly, full economic evaluations (e.g. cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis) compare the costs (resource use) associated with one or more alternative interventions (e.g. intervention X versus comparator Y) with their

consequences (outcomes, effects) whereas partial economic evaluations (e.g. cost analyses, cost-description studies and cost-outcome descriptions) do not make explicit comparisons between alternative interventions in terms of both costs (resource use) and consequences (effects) (Higgins & Green, 2011). The characteristics and results of included economic evaluations were summarised using a table supplemented by a narrative summary.

Statistical analysis was carried in R (R Core Team, 2015) using the “meta” package (Schwarzer, 2015).

Unit of analysis issues and measures of treatment effect

Cluster-randomised trials were verified to ensure appropriate analysis had been done (i.e. adjustment for clustering in calculating confidence intervals or p-values). Failure to account for clustering is likely to overestimate the precision of the study. In cases where such an analysis had not been done, the necessary intraclass correlation coefficients (ICCs) were extracted or obtained from the investigators and results re-analysed: the reviewer reduced the size of each C-RCT to its ‘effective sample size’ using the design effect $(1 + (M-1)*ICC)$, where M is the mean cluster size and ICC is the intraclass correlation coefficient (Rao & Scott, 1992). Where no information on the intra-cluster correlation coefficient (ICC) was reported in any of the cluster RCTs included in the analysis group, the reviewer assumed an ICC of 0.02 typical of primary and community care interventions, for this adjustment (Campbell, Grimshaw, & Steen, 2000b). Where an ICC was reported among the studies in a group, this ICC was used for the adjustments to other studies; if multiple studies in a group reported ICCs the most conservative ICC was used. Where results were presented as cluster means the number of events in each group was estimated as $(N*cluster\ mean/100)$.

As recommended in the Cochrane Handbook, two approaches were used for multi-arm studies: (a) In situations where only one pair of interventions was relevant to the meta-analysis, the pair was selected and the other arms excluded, (b) in situations where multiple intervention groups were relevant, all eligible intervention groups were combined to estimate an average treatment effect compared with the control. If intervention groups were sufficiently different that an average treatment effect would be meaningless, the 'shared' group was split into groups with smaller sample size, and included in (reasonably independent) comparisons (Higgins & Green, 2011). For crossover trials, data were extracted and analysed only from the period before the crossover.

Dealing with missing data

Authors of each of the included studies were contacted to obtain missing statistics (i.e. unreported data) and attrition was assessed as a form of bias. Additional outcome data was sought from seven authors; responses were received from two.

Assessment of reporting biases

To reduce possible publication bias, the reviewer employed strategies to search for and include relevant unpublished studies (Dickersin, 2005). These strategies included searching the grey literature and prospective trial registration databases to overcome time-lag bias.

Selective outcome reporting was assessed using the approach described in Chapter 8 of the Cochrane Handbook (Higgins & Green, 2011). Funnel plots (plotting effect size against standard error) were considered to detect potential bias in the five outcomes for which there were greater than 10 studies. Contour lines corresponding to perceived 'milestones' of statistical significance (e.g. $P = 0.01, 0.05, 0.1$) were used to

help differentiate asymmetry due to publication bias from that due to other factors (Peters, Sutton, Jones, Abrams, & Rushton, 2008).

Assessment of heterogeneity and moderator analysis

There exist several possible sources of heterogeneity among studies included in meta-analyses: (1) *methodological* heterogeneity (variability in trial design and quality); (2) *clinical* heterogeneity (variability in the participants, the types or timing of outcome measurements, and intervention characteristics); and (3) *statistical* heterogeneity (variability in summary treatment effects between trials) (Higgins & Green, 2011).

Methodological and clinical sources of heterogeneity contribute to the magnitude and presence of statistical heterogeneity. For example, trials that do not adequately conceal allocation to treatment groups may result in overestimates in the meta-analytic treatment effects (Pildal et al., 2007). Significant statistical heterogeneity arising from methodological heterogeneity suggests that the studies are not all estimating the same effects due to different degrees of bias (Gagnier, Moher, Boon, Beyene, & Bombardier, 2012). Clinical heterogeneity can also cause significant statistical heterogeneity, leading to inaccurate summary effects and associated conclusions.

Three strategies were used to assess these potential sources of heterogeneity. (1) Risk of bias for all included trials was assessed using The Cochrane Collaboration's tool for evaluating the risk of bias (see earlier sub-section on Assessment of risk of bias) and methodological heterogeneity was identified for relevant comparisons. (2) The clinical characteristics of each study were extracted, presented in tables in Appendix 4c, and extensively described in-text for each comparison. To contextualise each outcome, an abbreviated summary of these

characteristics was reiterated for studies reporting that outcome and sources of clinical heterogeneity identified, where relevant. (3) The extent of statistical heterogeneity across comparable studies was assessed using forest plots, the I^2 statistic and the χ^2 test.

In order to explore these potential sources of heterogeneity, moderator analysis—paying due attention to issues of statistical power (Hinshaw, 2002) and ecological fallacy (Thompson & Higgins, 2005)—was planned for the following variables derived from qualitative CHW literature: (1) gender of the CHW, (2) literacy level of the CHW, and (3) alternative/additional employment of the CHW (Blanas & Ndiaye, 2010; Creanga, Bradley, Kidanu, Melkamu, & Tsui, 2007; Crispin et al., 2012; Emukah et al., 2008; Green, Joyce, & Foreit, 2002; Hasegawa, Yasuoka, Ly, Nguon, & Jimba, 2013; Katz et al., 2002; Lopes, Cabral, & Sousa, 2014; MacCorquodale, 1982; Miller, Musominali, Baganizi, & Paccione, 2014; Mumtaz et al., 2013; Rahman et al., 2010; Razee, Whittaker, Jayasuriya, Yap, & Brentnall, 2012; Ronaghy et al., 1976; Ruebush, Weller, & Klein, 1994). These analyses would have been conducted had sufficient trials reported the necessary information (Rothwell, 2005).

A small number of exploratory, post-hoc subgroup and sensitivity analyses were performed to investigate sources of clinical and methodological variability identified over the course of the review. These are identified as post-hoc and exploratory in-text and should be interpreted with caution.

4.3 Results

4.3.1 Description of Studies

Results of the search

Electronic searches retrieved a total of 19 887 references; the six additional sources listed above yielded an additional 107 studies. Excluding duplicates, a total of 15 028 records were screened for inclusion (Fig. 4.1). Of these, 14 865 irrelevant records were excluded and 343 full text papers were considered for inclusion in the review. Of these, 155 papers reporting 86 trials met the inclusion criteria.

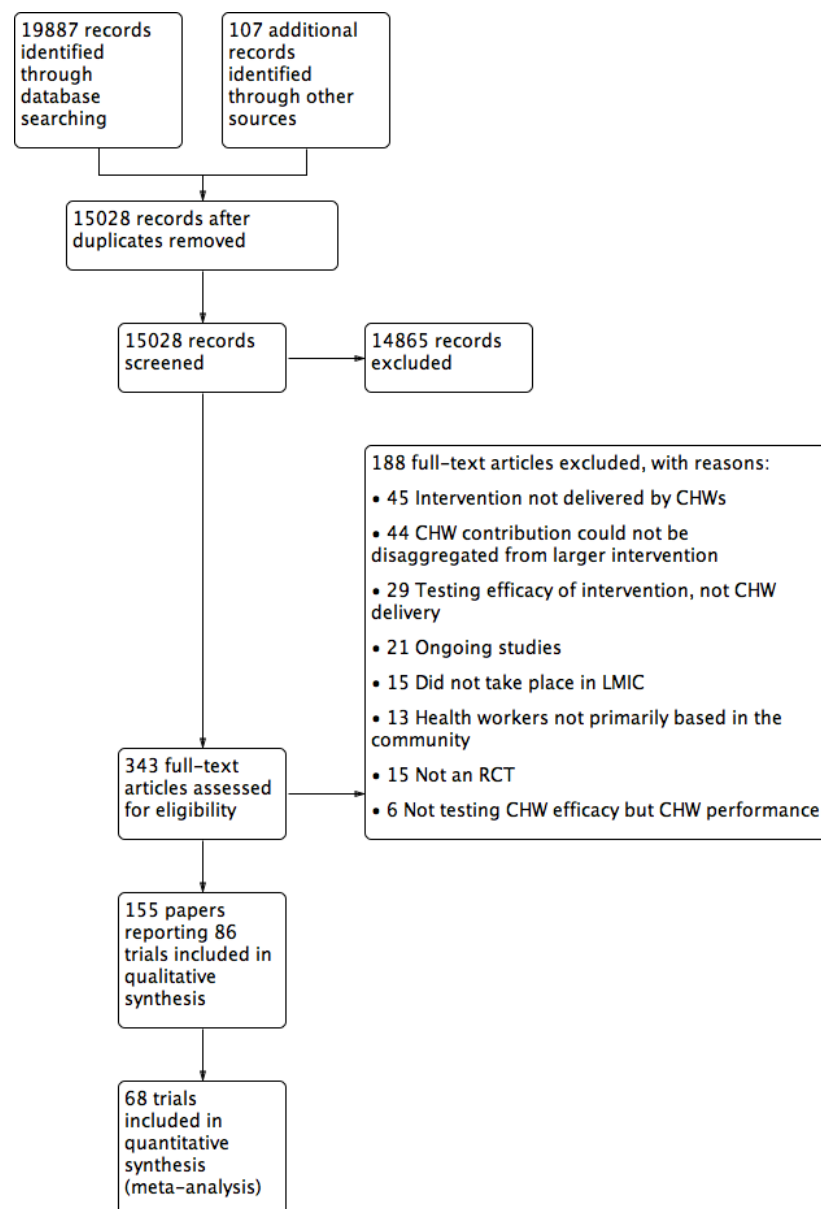


Figure 4.1 PRISMA Flow Diagram

Setting

Of the 86 studies included in this review, twenty-three studies³ (27%) were conducted in nine upper-middle-income countries: Brazil (2), China (2), Iran (1), Iraq (1), Jamaica (4), Mexico (1), South Africa (11), Thailand (1), and Turkey (1). Forty-one studies (48%) were conducted in fourteen lower-middle-income countries: Bangladesh (7), Cameroon (1), Ghana (6), India (10), Kenya (2), Myanmar (2), Nigeria (1), Pakistan (6), Philippines (1), Senegal (1), Swaziland (1), Vietnam (2), and Zambia (1). Twenty-three studies (27%) were conducted in eight low-income countries: Burkina Faso (1), Burundi (1), Ethiopia (2), Malawi (2), Nepal (2), Uganda (10), Tanzania (4), and The Gambia (1).

In sixty-six studies (77%), the intervention was delivered to patients in their home. In fifteen studies (18%) the intervention was delivered to patients in the community. Five studies (6%) involved a combination of home and community-based delivery.

Participants

CHW demographic (age, sex, education), and programme (recruitment, training, supervision, incentives etc.) information was poorly reported. As such, it was not possible to group the studies in terms of these metrics. Different recipients were targeted in the different groups of studies. A full account of CHW and recipient characteristics is outlined in the Effects of Interventions section below and in the Characteristics of Included Studies tables in Appendix 4c.

³ One study was conducted in both an upper-middle income (China) and lower-middle income (India) country and so is counted twice in this list.

Intervention characteristics

The objectives and modes of delivery of the interventions are discussed in more detail for each group of studies in the Effects of Interventions section below and in the Characteristics of Included Studies tables in Appendix 4c.

Outcomes

Most studies reported multiple effect measures and nearly half (45%) did not specify a primary outcome. Relevant outcomes were extracted and were categorised for the analysis according to the type of health issue addressed by the CHWs. This is explained further in the Effects of Interventions section below.

Excluded studies

In total, 167 studies were of interest to this area of study but did not fulfil all inclusion criteria. These studies, together with their reasons for exclusion, are documented in Characteristics of excluded studies table (Appendix 4c). In brief, 45 were excluded because the intervention was not delivered by CHWs, 44 because the CHW contribution could not be disaggregated from the larger intervention, 29 because the trial was testing the efficacy of a task rather than CHW delivery, 15 because the trial did not take place in an LMIC, 13 because the health workers were not primarily based in the community, 15 because the study was not a randomised controlled trial, and 6 because the trial was testing a CHW performance enhancement interventions rather than the CHWs' ability to deliver a particular intervention.

Ongoing studies

Information regarding the 21 ongoing studies identified is provided in the Table of Ongoing Studies in Appendix 4c.

4.3.2 Risk of Bias of Included Studies

Assessments of the risk of bias for each included study are summarised in the Characteristics of included studies tables in Appendix 4c and summarised in Table 4.1 and Figure 4.2. Risk of bias assessments were not used in deciding which studies should be included in the meta-analyses but for (a) interpreting the results and (b) assessing the quality of evidence for specific outcomes of CHW delivered interventions. The first six categories (adequate sequence generation, allocation concealment, blinding, incomplete outcome data, selective reporting, and publication bias) apply to all included trials, while the latter four (recruitment bias, baseline imbalance, loss of clusters, and unit of analysis issues) apply only to cluster-randomised control trials.

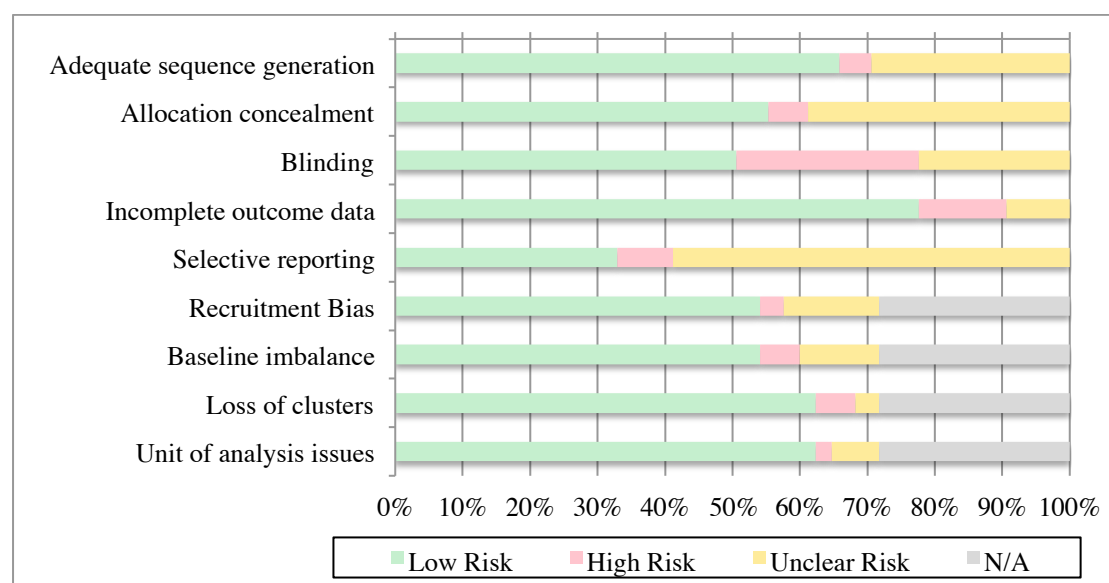


Figure 4.2 Reviewer judgements about each methodological quality item as a percentage across all included studies.

Table 4.1 Review reviewers' judgements about each methodological quality item for all included studies

[illegible]

Continued from previous page

[illegible]

Continued from previous page

	Adequate sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Recruitment Bias	Baseline imbalance	Loss of clusters	Unit of analysis issues
Tripathy 2010	?	+	-	+	+	?	+	+	+
Tripathy 2016	+	+	+	-	+	+	+	-	+
Tylleskar 2011	+	+	+	+	+	+	+	+	+
Vazir 2013	+	+	+	+	?	+	+	+	+
Waiswa 2015	+	+	?	+	?	+	+	+	+
Wang 2007	?	?	?	?	?	+	+	+	+
Wright 2004	?	?	+	+	?				
Yeomans 2010	+	+	+	+	?				
Yousafzai 2014	+	+	+	+	?	+	+	-	+
Zwarenstein 2000	+	+	+	+	?				

Random sequence generation and allocation (selection bias)

The majority (56/86) of included trials met the ‘low risk of bias’ criteria for random sequence generation. In the remaining studies, 4 were at high risk of bias (Gardner et al., 2003; le Roux et al., 2010; le Roux et al., 2011; Neuner et al., 2008) and in 26 the risk of bias was unclear due to poor reporting. A slightly lower proportion (47/86) met the ‘low risk of bias for allocation concealment. Of remaining studies, 34 had unclear risk of bias due to poor reporting and five (le Roux et al., 2010; le Roux et al., 2011; Neuner et al., 2008; Newell et al., 2006; Ohnmar et al., 2012) were at high risk. Two of these studies (le Roux et al., 2010; le Roux et al., 2011) have strong baseline differences that suggested that the assignment of “needier” participants into the intervention arm probably took place. In both cases, the direction of bias caused by this deviation from random assignment is likely to be against the detection of an intervention effect. In one study (Lwilla et al., 2003), though there was likely no deviation from random assignment, some selection bias may still have

occurred because the facility likely to have the sickest patients was randomised to the control group and so may have biased the results in favour of the intervention (community-based CHW care).

Blinding (performance bias and detection bias)

The community-level nature of these interventions precluded the blinding of participants and personnel to the presence or actions of a CHW. The impossibility of blinding intervention staff may have led to performance bias. Half (43/86) of studies blinded their outcome assessors, in 20 studies the risk of bias is unclear due to poor reporting, and 23 studies did not blind their outcome assessors (Aung et al., 2014; Azad et al., 2010; Bari et al., 2011; Bari et al., 2006; Chang et al., 2010a; Chang et al., 2015; Colbourn et al., 2013b; Darmstadt et al., 2010; Datiko & Lindtjorn, 2009; Fylkesnes et al., 2013; Haider et al., 2000; le Roux et al., 2011; Morrow et al., 1999; Mtango & Neuvians, 1986; Ohnmar et al., 2012; Pachón et al., 2002; Selke et al., 2010; Sloan et al., 2008; Stanton et al., 2013; Swart et al., 2008; Tine et al., 2011; Tripathy et al., 2010; UNICEF/UNDP/World Bank/WHO TDR, 2000).

While some studies in which assessors were unblinded measured biological outcomes (e.g. virologic failure in Chang et al., 2010 or sputum conversion in Datiko & Lindtjorn, 2009) and so may not have been at as large a risk of bias, the more subjective outcome collection tools used in several of the studies (e.g. household interviews) and the fact that multiple studies used CHWs to both implement the intervention and collect data (e.g. le Roux et al., 2011; Mtango & Neuvians, 1986; Selke et al., 2010; Tine et al., 2014) make detection bias a notable risk.

In one study (Stanton et al., 2013), there is evidence of performance bias: two CHW behaviours differed between intervention and control groups.

Incomplete outcome data (attrition bias)

Final attrition rates ranged from 1% (Wright et al., 2004) to nearly 42% (Ali et al., 2003). While the majority of trials (66/86) met the 'low risk of bias' criterion for incomplete outcome data, the bias in 9 trials was unclear due to poor reporting and 11 trials should be considered high risk (Ali et al., 2003; Bari et al., 2011; CDI Study Group, 2010; Colbourn et al., 2013a; Eloff et al., 2014; Gökçay et al., 1993; le Roux et al., 2010; Rahman et al., 2009; Siabani et al., 2015; Sloan et al., 2008; Tripathy et al., 2016). Of these, most had attrition between 10-20% but were marked down for excluding dropouts from the analysis. One trial had a final attrition above 30%, which must be considered when judging both internal and external validity (Ali et al., 2003); though this trial reported an attrition analysis indicating that there was no significant difference found between the dropout rates of the intervention and control groups or between the socio-demographic characteristics of the dropouts and non-dropouts within each group.

Selective reporting (reporting bias)

Based on outcomes listed in study protocols, there appeared to be a low risk of selective reporting for only a minority of studies (28/86). For the majority (51/86) no protocol was identified and so the risk of bias remains unclear. In seven studies (Baker-Henningham et al., 2005; Bojang et al., 2011; Chang et al., 2010a; Chatterjee et al., 2014; Chongsuvivatwong et al., 1996; Darmstadt et al., 2010; Kweku et al., 2009) there was a high risk of selective reporting, with Kweku et al. (2009) switching their primary outcome, Chongsuvivatwong et al. (1996) collecting information on the outcomes in the intervention group longer than the control group, and the rest failing to report all outcomes listed in the protocol.

Publication bias

Funnel plots (plotting effect size against standard error) were considered to detect potential bias in the five outcomes for which there were greater than 10 studies. Terrin, Schmid, Lau, and Olkin (2003) have suggested that the funnel plot is inappropriate for heterogeneous meta-analyses, noting that the funnel plot was originally used on the premise that the studies come from a single underlying population (Richard & Pillemer, 1984). Given this, funnel plots were only conducted for the two of five outcomes with low heterogeneity (see Appendix 4c). No meaningful conclusion could be drawn from the first plot; slight asymmetry in the second plot, in which studies appear to be missing in areas of statistical non-significance, suggest the intervention effect may be overestimated (Egger, Smith, Schneider, & Minder, 1997; Villar, Piaggio, Carroli, & Donner, 1997).

C-RCTs: Recruitment bias

The majority of C-RCTs (46/62) had a low risk of recruitment bias, with participants having been recruited before cluster-randomisation or with all participants in a cluster being included in the analysis. In 13 studies the risk of bias is unclear due to poor reporting and in 3 studies participants were recruited after randomisation so there is a high risk of recruitment bias (Aung et al., 2014; Baker-Henningham et al., 2005; Gökçay et al., 1993)

C-RCTs: Baseline imbalance

The majority of C-RCTs (45/61) had a low risk of bias due to baseline imbalance. 11 studies did not provide an explicit statement of group equivalence at baseline and in 5 studies (Azad et al., 2010; Bolton et al., 2003; Sloan et al., 2008; Soofi et al., 2012; Stanton et al., 2013) there were baseline imbalances in demographic and

clinical indicators that could have affected measured outcomes, though the authors adjusted for these in the analysis.

C-RCTs: Loss of clusters

In the majority of C-RCTs (53/62), no clusters were lost. In the 9 remaining studies, 4 had insufficient information was available to make a judgement on risk of bias and 5 had clusters that were lost (Colbourn et al., 2013a; Rotheram-Borus et al., 2014; Stanton et al., 2013; Tripathy et al., 2016; Yousafzai et al., 2014). In Rotheram-Borus et al. (2014) one matched pair was eliminated after 6 months of recruitment due to low numbers of pregnant women. In Stanton et al. (2013) two additional CHW catchment areas were selected to replace non-compliant CHWs, both of whom were in the intervention group. The new catchment areas were adjacent to those of the non-compliant CHWs and had similar characteristics. In Yousafzai et al. (2014) four CHWs left their posts officially, all of whom were in one of the intervention arms of the trial.

C-RCTs: Unit of analysis issues

The majority of C-RCTs properly accounted for clustering (53/62). In 7 studies, risk of bias was not possible to assess due to poor reporting. In 2 studies, no adjustment for clustering was made (Pachón et al., 2002; Pence et al., 2007), meaning the precision of the study is likely to be overestimated. In one of these studies, Pence et al. (2007), there was only one cluster per trial arm. This prevents the application of statistical methods to correct for intra-area correlation of errors and means that it should be read as a plausibility rather than a probability study (Habicht, Victora, & Vaughan, 1999), i.e. interpretation of its results should not rely primarily on statistical significance.

4.3.3 Effects of Interventions

This review covers a wide range of CHWs delivering a wide range of healthcare interventions for a variety of outcomes. Studies were grouped according to the type of health issue addressed by the CHWs as policy decisions regarding task shifting are typically made on a “by disease” basis (Lewin, Zwarenstein, & Dick, 2005). The studies in these broad groups were then subdivided into smaller groups based on clinically similar interventions intended influence a single health care outcome or a group of closely related outcomes. Combining the results of included studies to obtain an overall estimate of effect was feasible for the majority of outcomes in comparisons 1-9. Below are the groupings reported:

MENTAL HEALTH OUTCOMES

1. CHW delivered interventions for treatment of common mental disorders
 - 1.1. Prevalence of depression
 - 1.2. Severity of common mental disorder symptoms (includes anxiety and depression)
 - 1.3. Functional impairment of adults with common mental disorders
 - 1.4. Severity of post-traumatic stress disorder (PTSD)
 - 1.5. Severity of depression symptoms in adolescents

MATERNAL AND PAEDIATRIC OUTCOMES

2. CHW delivered interventions to reduce mortality and morbidity in children under five
 - 2.1. Mortality among children under five years
 - 2.2. Child morbidity, reported illness in children
 - 2.3. Infant mortality
 - 2.4. Infant morbidity
3. CHW delivered interventions to improve maternal and newborn health
 - 3.1. Stillbirths
 - 3.2. Neonatal mortality
 - 3.3. Early neonatal mortality
 - 3.4. Late neonatal mortality
 - 3.5. Perinatal mortality
 - 3.6. Maternal mortality
 - 3.7. Maternal morbidity
 - 3.8. Facility-based deliveries
 - 3.9. Use of clean delivery kits
 - 3.10. Exclusive breastfeeding up to six months post partum.
 - 3.11. Initiation of breastfeeding within 1 h of birth

4. CHW delivered interventions to promote cognitive development in infants
 - 4.1. Maternal knowledge of infant development
 - 4.2. Mental development
5. CHW delivered interventions to promote healthy feeding practices
 - 5.1. Weight-for-age Z score
 - 5.2. Height-for-age Z score
 - 5.3. Weight-for-height Z score

INFECTIOUS DISEASE OUTCOMES

6. CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support
 - 6.1. TB Cure rate
 - 6.2. TB Cure or completion rate (vs. facility-based care)
 - 6.3. TB Completion rate (vs. family DOTS)
7. CHW delivered interventions for maintenance of antiretroviral therapy
 - 7.1. Death
 - 7.2. Loss to follow-up (LTFU)
 - 7.3. Retention in pre-ARV care
 - 7.4. CD4 count

PUBLIC AND PREVENTATIVE HEALTH OUTCOMES

8. CHW delivered interventions to promote injury prevention in children
 - 8.1. Hazard score
9. CHW delivered interventions to promote immunisation uptake and mass drug administration
 - 9.1. Immunisation coverage
 - 9.2. Mass drug administration coverage
10. CHW delivered interventions to reduce hypertension
 - 10.1. Systolic blood pressure

The remaining 7 studies were diverse and could not be usefully grouped. Studies describe CHW delivered interventions for dementia care (Dias et al., 2008), chronic heart failure (Siabani et al., 2015), oral cancer screening (Sankaranarayanan et al., 2005), among other applications. These studies are discussed narratively in the text and their primary outcomes are listed in Appendix 4d (“Outcomes of trials not assigned to the above comparisons”).

In the majority of cases where meta-analysis was performed analyses employed the primary study outcome. Adjusted forest plots and GRADE tables for all meta-analyses conducted are provided in the text; a list of all syntheses conducted for

this review and the forest plots for all unadjusted syntheses are in Appendix 4d. For studies where outcomes were too diverse to be usefully pooled, outcomes are reported narratively in the text and in tables in the relevant sections. Detailed descriptions for each study are provided in the Characteristics of included studies tables in Appendix 4c.

MENTAL HEALTH OUTCOMES

Comparison 1: CHW delivered interventions for treatment of common mental disorders

Eleven trials were examined in this comparison (Ali et al., 2003; Baker-Henningham et al., 2005; Bolton et al., 2007; Bolton et al., 2003; Chatterjee et al., 2014; Eloff et al., 2014; Ertl et al., 2011; Neuner et al., 2008; Pradeep, Isaacs, Shanbag, Selvan, & Srinivasan, 2014; Rahman et al., 2008; Yeomans et al., 2010). Two trials were included in the meta-analysis of outcome 1.1. Prevalence of depression (Bolton et al., 2003; Rahman et al., 2008). Seven trials were included in the meta-analysis of outcome 1.2 Severity of common mental disorder symptoms (Ali et al., 2003; Baker-Henningham et al., 2005; Bolton et al., 2003; Chatterjee et al., 2014; Eloff et al., 2014; Pradeep et al., 2014; Rahman et al., 2008). Three trials were included in the meta-analysis of outcome 1.3 Functional impairment of adults with common mental disorders (Bolton et al., 2003; Chatterjee et al., 2014; Rahman et al., 2008). Two trials were included in the meta-analysis of outcome 1.4 Severity of post-traumatic stress disorder (PTSD) (Neuner et al., 2008; Yeomans et al., 2010). Two trials were included in the meta-analysis of outcome 1.5 Severity of depression in adolescents (Bolton et al., 2007; Ertl et al., 2011).

Setting: Two trials were conducted in upper-middle-income countries, four in lower-middle-income countries, and five in low-income countries. Four trials were conducted in Uganda (Bolton et al., 2003; Bolton et al., 2007; Ertl et al., 2011;

Neuner et al., 2008); two in Pakistan (Ali et al., 2003; Rahman et al., 2008); two in India (Chatterjee et al., 2014; Pradeep et al., 2014); and one each in Jamaica (Baker-Henningham et al., 2005), Burundi (Yeomans et al., 2010) and South Africa (Eloff et al., 2014).

Participants: One study included adults of both sexes with depression (Bolton et al., 2003), while five studies included women with depression (Ali et al., 2003; Pradeep et al., 2014), perinatal depression (Baker-Henningham et al., 2005; Rahman et al., 2008), or at risk for depression (Eloff et al., 2014). One study included adolescents aged 14 to 17 living in displaced persons camps with PTSD symptoms (Bolton et al., 2007) while another included child soldiers aged 12 to 25 years (Ertl 2011 RCT). Two studies included adults of both sexes who were diagnosed with PTSD (Neuner et al., 2008; Yeomans et al., 2010). One study included adults with a primary diagnosis of schizophrenia (Chatterjee et al., 2014).

CHW recruitment was via the research team (Ali et al., 2003; Bolton et al., 2007; Ertl et al., 2011), the health system (Baker-Henningham et al., 2005; Rahman et al., 2008), an NGO (Chatterjee et al., 2014) or by unspecified actors. Scholastic recruitment requirements ranged from 10-12 years of schooling (Chatterjee et al., 2014; Eloff et al. 2014) to basic literacy (Ali et al., 2003; Neuner et al., 2008) or no formal education (Yeomans et al., 2010). In three trials, recruited CHWs were female (Ali et al., 2003; Pradeep et al., 2014; Rahman et al., 2008) while one trial used sex-specific health workers for sex-specific group (Bolton et al., 2003).

Description of interventions: Aside from CHWs in Pradeep et al. (2014) who provided health education, all CHWs in this comparison provided psychological interventions: cognitive behaviour therapy (CBT)-like problem solving (Ali et al., 2003; Rahman et al., 2008), group interpersonal therapy (IPT-G) (Bolton et al., 2003;

Bolton et al., 2007), narrative exposure therapy (NET) (Ertl et al., 2011; Neuner et al., 2008), collaborative community-based schizophrenia care (Chatterjee et al., 2014), trauma counselling (Neuner et al., 2008; Yeomans et al., 2010), or family-centred psychosocial support (Baker-Henningham et al., 2005; Eloff et al. 2014). Yeomans et al. (2010) had two arms: in one the trauma workshop was provided with PTSD psychoeducation and in the other it was provided without. For the four studies with multiple trial arms (Bolton et al., 2007; Ertl et al. 2011; Neuner et al., 2008; Yeomans et al., 2010), eligible intervention groups in each study were, in line with the review protocol, combined to estimate an average treatment effect compared with the control (for additional detail on the trial arms in each study see Characteristics of Included Studies Tables in Appendix 4c). Interventions varied in duration, ranging from 30 (Baker-Hennigham et al., 2005) to 120 minutes (Bolton et al., 2007), in frequency, with most occurring weekly but two increasing the intervals between sessions over time (e.g. Chatterjee et al., 2014 Rahman 2008), and in total time, from three weeks (Neuner et al., 2008) to one year (Baker-Hennigham et al., 2005; Chatterjee et al., 2014). The ratio of CHWs to population ranged from 1:6-8 (Bolton et al., 2007) to 1:30 (Pradeep et al., 2014)

In five studies, interventions were manualised (Bolton et al., 2003; Bolton et al., 2007 - IPT-G; Eloff et al. 2014; Neuner et al., 2008 – NET; Yeomans – both arms; Rahman et al., 2008) or involved manual-based training (Baker-Hennigham et al., 2005; Ertl et al., 2011). Training ranged from 1 day (Yeomans et al., 2010) to 6 weeks (Baker-Henningham et al., 2005; Chatterjee et al., 2014; Neuner et al., 2008). Ongoing education was offered in two studies in the form of continued training (Chatterjee et al., 2014) or a two-day refresher at four months (Baker-Hennigham et al., 2005). Incentives were not described except in Baker-Hennigham et al. (2005)

where CHWs were employed by the government and Bolton et al., (2003) where CHWs received weekly payment of an unspecified amount for the duration of the intervention.

Supervision was described in six studies. It was provided by NGO staff (Bolton et al., 2007), social workers (Chatterjee et al., 2014; Eloff et al., 2014), the research team (Rahman et al., 2008), or by a supervisor for which no characteristics were provided (Baker-Henningham et al., 2005; Ertl et al., 2011; Neuner et al., 2008). The frequency of supervision was weekly (Bolton et al., 2007; Neuner et al., 2008), every two weeks (Baker-Henningham et al., 2005), monthly (Rahman et al., 2008), or quarterly (Chatterjee et al., 2014).

Comparison groups included usual care without the addition of a CHW (Ali et al., 2003; Bolton et al., 2003; Chatterjee et al., 2014; Eloff et al., 2014; Neuner et al., 2008; Pradeep et al., 2014), usual care where the CHW was not trained to deliver the intervention in question (Baker-Henningham et al., 2005; Rahman et al., 2008), or a wait-list control (Bolton et al., 2007; Ertl et al., 2011; Yeomans et al., 2010).

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison one. Prevalence of depression, functional impairment of adults with common mental disorders, and severity of PTSD are also included in Summary of Findings 1. Data reported at longest follow-up were included unless otherwise specified.

1.1 Prevalence of depression

Two trials were included in the meta-analysis (Fig 4.3). The trials were conducted in Uganda (Bolton et al. 2003) and Pakistan (Rahman et al., 2008) on adults with depression. While the CHWs provided different types of manualised short-term psychological therapies across trials (IPT in Bolton et al, 2003; CBT-based

in Rahman et al., 2008), any apparent clinical heterogeneity did not translate into statistical heterogeneity ($I^2 = 0\%$, $P = 0.81$). The meta-analysis showed moderate quality evidence that CHW-led psychological interventions can reduce the prevalence of depression (RR 0.46, 95% CI 0.38 to 0.56, $P < 0.0001$). Evidence was downgraded due to study design and implementation (allocation concealment in Bolton 2003 was unclear).

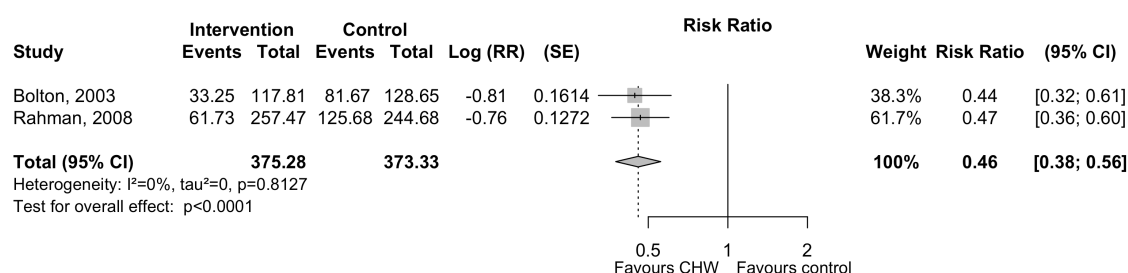


Figure 4.3 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.1 Prevalence of depression – adjusted for clustering

1.2 Severity of common mental disorder symptoms

Seven trials were included in the meta-analysis (Fig 4.4). CHWs delivered a range of psychological interventions to adults with common mental disorders (e.g. schizophrenia, depression) in the Caribbean, South Asia, and Africa. While some clinical heterogeneity exists across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. country, disorder, CHW programme features etc.). The meta-analysis showed low quality evidence that CHW-led interventions lead to moderate clinical benefit with regards to the severity of common mental disorders (SMD -0.64, 95% CI -0.95 to -0.32, $P < 0.0001$). Evidence was downgraded due to study design and implementation (allocation was unclear in four studies) and substantial unexplained statistical heterogeneity ($I^2 = 87.4\%$, $P < 0.0001$).

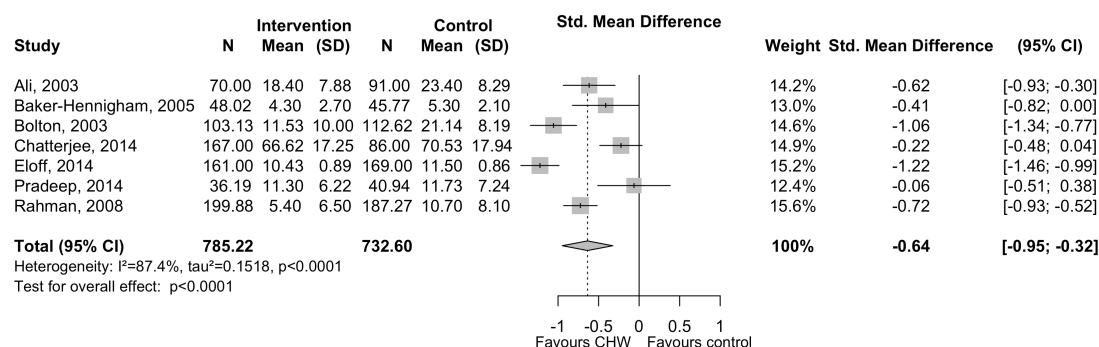


Figure 4.4 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.2 Severity of common mental disorder symptoms (includes anxiety and depression) – adjusted for clustering

1.3 Functional impairment of adults with common mental disorders

Three trials were included in the meta-analysis (Fig 4.5). CHWs delivered a range of psychological interventions to adults with common mental disorders (e.g. schizophrenia, depression) in South Asia and East Africa. While some clinical heterogeneity existed across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. region, disorder, CHW programme features etc.). The meta-analysis showed moderate quality evidence that CHW delivered interventions led to a moderate clinical benefit with regards to functional impairment within six months of interventions (SMD -0.44, 95% CI -0.80 to -0.09, $P = 0.041$). Evidence was downgraded due to substantial unexplained statistical heterogeneity ($I^2 = 85\%$, $P = 0.001$).

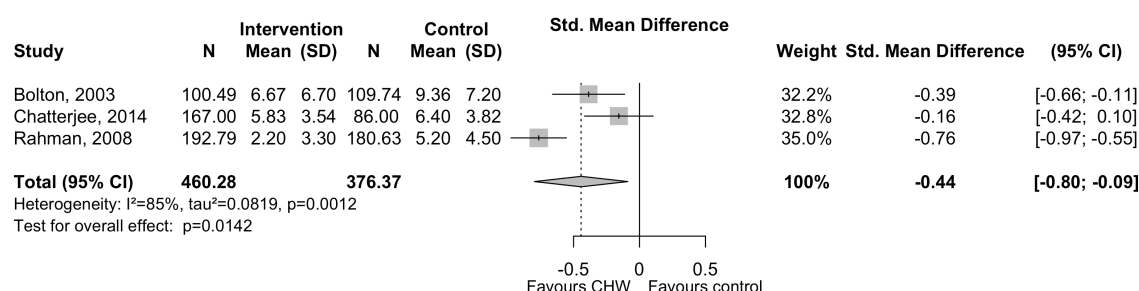


Figure 4.5 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.3 Functional impairment of adults with common mental disorders – adjusted for clustering

1.4 Severity of post-traumatic stress disorder (PTSD)

Two trials were included in the meta-analysis (Fig 4.6). CHWs delivered trauma counselling to adults with PTSD in East (Neuner et al. 2008) and West

(Yeomans et al., 2008) Africa. The meta-analysis showed moderate quality evidence that CHW delivered interventions can lead to a moderate clinical benefit with regards to improvement in PTSD symptoms (SMD -0.54, 95% CI -0.81 to -0.28, $P < 0.0001$). Evidence was downgraded due to study design and implementation (sequence generation and allocation concealment in Neuner 2008 were high risk). There was little evidence of statistical heterogeneity ($I^2 = 16.4\%$, $P = 0.27$).

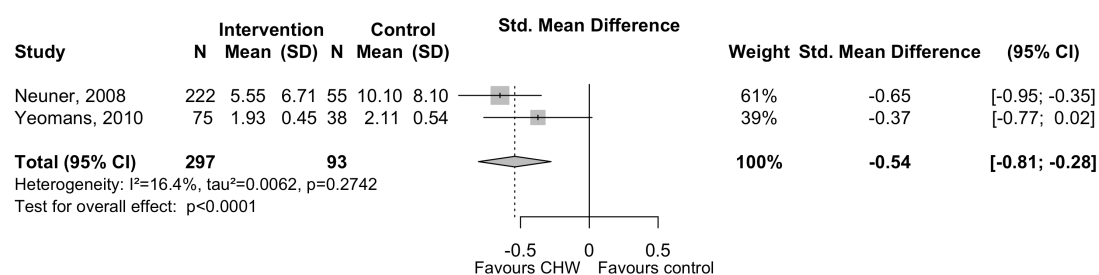


Figure 4.6 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.4 Severity of post-traumatic stress disorder (PTSD)

1.5 Severity of depression symptoms in adolescents

Two trials, both set in Uganda, were included in the meta-analysis (Fig 4.7). CHWs delivered short-term therapy to adolescents who had experienced trauma. The meta-analysis showed very low quality evidence that CHW delivered interventions may have a small to moderate effect on reducing depressive symptoms (SMD -0.41, 95% CI -1.01 to 0.20, $P = 0.185$). Evidence was downgraded due to imprecision, study design and implementation (adequate sequence generation was unclear in Bolton et al., 2007 and both adequate sequence generation and allocation concealment were unclear in Ertl et al., 2011) and substantial unexplained statistical heterogeneity ($I^2 = 80.7\%$, $P = 0.02$). As only two studies report on this outcome it was not possible to conduct sensitivity analysis on clinical differences between studies (e.g. individual-based vs. group-based treatment) that may relate to the difference in effect sizes.

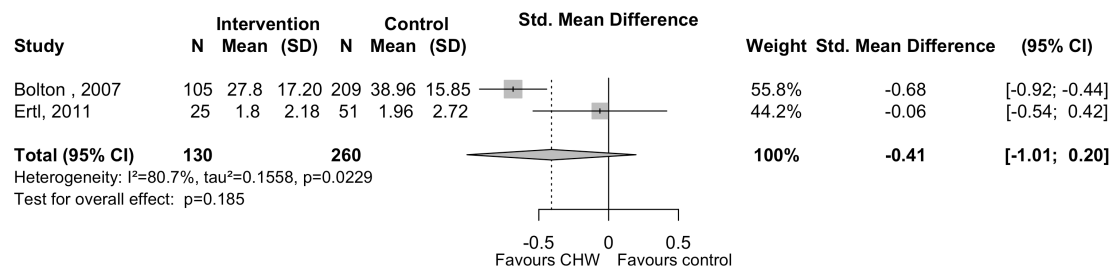


Figure 4.7 Comparison 1: CHW for treatment of common mental disorders compared to usual care, Outcome: 1.5 Severity of depression in adolescents

MATERNAL AND PAEDIATRIC OUTCOMES

Comparison 2: CHW to reduce mortality and morbidity in children under five

Fifteen trials were examined in this comparison (Agrasada et al., 2005; Azad et al., 2010; Bari et al., 2011; Bari et al., 2006; Chinbuah et al., 2012; Chongsuvivatwong et al., 1996; Kidane & Morrow, 2000; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; Mtango & Neuvians, 1986; Pence et al., 2007; Sloan et al., 2008; Soofi et al., 2012; Sripaipan et al., 2002). Four trials reported on outcome 2.1 Mortality among children under five years (0-59 months) (Chinbuah et al., 2012; Kidane & Morrow, 2000; Mtango & Neuvians, 1986; Pence et al., 2007). Four trials were included in the meta-analysis of outcome 2.2 Child morbidity (Bari et al., 2011; Chongsuvivatwong et al., 1996; Soofi et al., 2012; Sripaipan et al., 2002). Two trials were included in outcome 2.3 Infant mortality (Agrasada et al., 2005; Lewycka et al., 2013; Sloan et al., 2008). Seven trials were included in outcome 2.4 Infant morbidity (Agrasada et al., 2005; Azad et al., 2010; Bari et al., 2006; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; Sloan et al., 2008).

Setting: One trial was conducted in an upper-middle-income country, ten in lower-middle-income countries, and four in low-income-countries. Three trials were conducted in Bangladesh (Azad et al., 2010; Bari et al., 2006; Sloan et al., 2008); two

trials were conducted in Pakistan (Bari et al., 2011; Soofi et al., 2012) and Ghana (Chinbuah et al., 2012; Pence et al., 2007); and one trial was conducted in Ethiopia (Kidane & Morrow, 2000), India (Kumar et al., 2008), Malawi (Lewycka et al., 2013), Nepal (Manandhar et al., 2004), the Philippines (Agrasada et al., 2005), Tanzania (Mtango & Neuvians, 1986), Thailand (Chongsuvivatwong et al., 1996), and Vietnam (Sripaipan et al., 2002).

Participants: All studies were conducted on children less than five years of age and their mothers. CHW recruitment was via the study organisers (Agrasada et al. 2005; Azad et al., 2010; Bari et al., 2006; Mtango & Neuvians, 1986), community leaders or communities (Chinbuah et al., 2012; Pence et al., 2007), the health system (Bari et al., 2011; Sloan et al., 2008; Soofi et al., 2012), the health system and the host villages together (Lewycka et al., 2013; Manandhar et al., 2004) or was not specified. Recruitment criteria ranged from completion of secondary school (Kumar et al., 2008) or middle school (Soofi et al., 2012) to basic literacy (Lewycka et al., 2013; Manandhar et al., 2004). In one study CHWs were illiterate (Kidane & Morrow, 2000). Other criteria included performing well in a training course (Agrasada et al. 2005) or being female (Azad et al., 2010; Lewycka et al., 2013; Soofi et al., 2012).

Description of interventions: In six trials, CHWs treated childhood illness during home visits (Bari et al., 2011; Chinbuah et al., 2012; Mtango & Neuvians, 1986; Pence et al., 2007; Soofi et al., 2012) or taught local mothers to provide treatment (Kidane & Morrow, 2000). In seven trials, CHWs promoted birth preparedness and/or childcare behaviours during structured home visits (Agrasada et al., 2005; Bari et al., 2006; Chongsuvivatwong et al., 1996; Kumar et al., 2008; Lewycka et al., 2013; Sloan et al., 2008; Sripaipan et al., 2002). In three studies, CHWs led women's groups through an action-learning cycle in which the women

identified local health problems and formulated strategies to address them (Azad et al., 2010; Lewycka et al., 2013; Manandhar et al., 2004). Four studies had multiple trial arms (Agrasada et al., 2005; Chinbuah et al., 2012; Kumar et al., 2008; Lewycka et al., 2013; Pence et al., 2007); in line with the review protocol, eligible intervention groups in each study were combined to estimate an average treatment effect compared with the control (for additional detail on the trial arms in each study, see Characteristics of Included Studies Tables in Appendix 4c). The ratio of CHWs to population ranged from 1:8 (Agrasada et al., 2005) to 1:4000 (Bari et al., 2006).

CHW training was described in sixteen studies and ranged from two days (Chongsuvivatwong et al., 1996) to three months (Bari et al., 2011). Initial instruction was followed by up to one year of in-service training (Bari et al., 2011; Mtango & Neuvians, 1986) or refresher training on a monthly (Soofi et al., 2012), quarterly (Chinbuah et al., 2012; Pence et al., 2007), or annual (Lewycka et al., 2013) basis.

Incentives were described in nine studies; in three studies CHWs were provided a salary or incentivized payments (Kumar et al., 2008; Sloan et al., 2008; Soofi et al., 2012), in one study they were reimbursed for expenses (Agrasada et al., 2005), in one study they received monetary and non-monetary (e.g. bicycle, calendars) incentives (Lewycka et al., 2013), and in two studies (Chinbuah et al., 2012; Sripaipan et al., 2002) no incentives were provided.

Supervision was described in five studies. It was provided variously by village committees (Pence et al., 2007), certified health workers (Agrasada et al. 2005), local field supervisors (Azad et al., 2010; Bari et al., 2006; Bari et al., 2011; Chinbuah et al., 2012; Kidane & Morrow, 2000), government health surveillance assistants (Lewycka et al., 2013), or by a supervisor for which no characteristics were provided (Kumar et al., 2008; Manandhar et al., 2004). The frequency with which supervision

was undertaken ranged from fortnightly (Chinbuah et al., 2012) to quarterly (Lewycka et al., 2013).

Comparison groups included usual care without the addition of a CHW (Agrasada et al., 2005; Bari et al., 2006; Chinbuah et al., 2012; Chongsuvivatwong et al., 1996; Kumar et al., 2008; Lewycka et al., 2013; Mtango & Neuvians, 1986; Pence et al., 2007; Sloan et al., 2008; Sripaipan et al., 2002), usual care where the CHW was not trained to deliver the intervention in question (Bari et al., 2011; Kidane & Morrow, 2000; Soofi et al., 2012), and health-facility strengthening activities undertaken in both intervention and control areas (Azad et al., 2010; Manandhar et al., 2004).

Results: Findings are discussed for the main outcomes measured by included studies: mortality among children under five years, child morbidity, infant mortality, infant morbidity are included in Summary of Findings 2. Data reported at longest follow-up were included.

2.1 Mortality among children under five years

Four trials reported on outcome 2.1; data from three studies were included in the meta-analysis (Fig 4.8). CHWs treated or trained mothers to treat common paediatric illnesses in children under 5 years in East (Kidane & Morrow, 2000; Mtango & Neuvians, 1986) and West (Pence et al., 2005) Africa. The meta-analysis showed low quality evidence that CHW delivered interventions may reduce child mortality (RR 0.65, 95% CI 0.38 to 1.10, $P = 0.11$). Evidence was downgraded for imprecision and limitations in study design (Mtango & Neuvians, 1986 and Pence et al., 2005 only had 2 clusters each, causing issues with baseline differences). There was no evidence of statistical heterogeneity ($I^2 = 0\%$, $P = 0.83$).

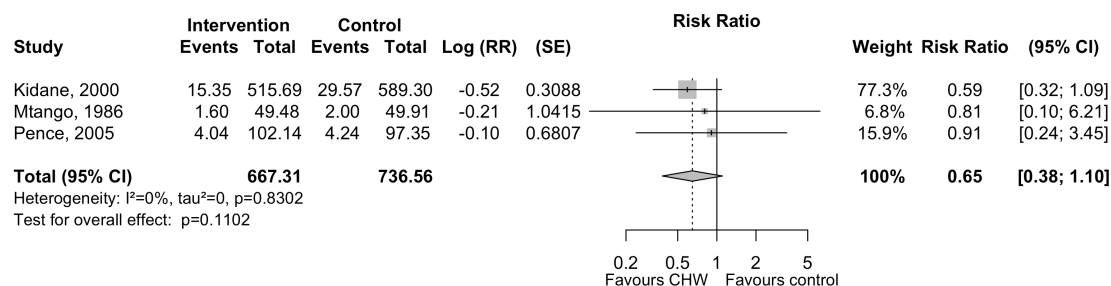


Figure 4.8 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.1 Mortality among children under five years (0-59 months) – adjusted for clustering

Though the control group in Chinbuah et al. (2012) was only wait-listed for six months, the paper reports mortality by year. Efforts to track down the six-month data by trial arm have been unsuccessful.

2.2 Child morbidity

Data from four studies were included in the meta-analysis (Fig 4.9). CHWs treated childhood illness or promoted childcare behaviours during structured home visits to children under five years in South and Southeast Asia. The meta-analysis showed low quality evidence CHW delivered intervention may reduce child morbidity (RR 0.70, 95% CI 0.51 to 0.96, $P = 0.026$). Evidence was downgraded for limitations in study design (it was unclear whether assessors were blinded in the majority of studies) and substantial statistical heterogeneity ($I^2 = 93.8\%$, $P = 0.1$).

In order to investigate this statistical heterogeneity a subgroup analysis was performed based on CHW delivered intervention type. While the magnitude of the effect was statistically significantly different between interventions in which CHWs provided treatment to children and those in which CHWs only conducted health promotion ($P < 0.0001$), the interaction is quantitative, rather than qualitative (i.e. the size of the effect varies but not the direction, see: Yusuf, Wittes, Probstfield, & Tyroler, 1991). More critically, this is a post-hoc, observational analysis on a very small number of trials; the results could be confounded by other study-level

characteristics and do not indicate that the subgroup factor explains heterogeneity. As a result, the results of this subgrouping should be interpreted with caution.

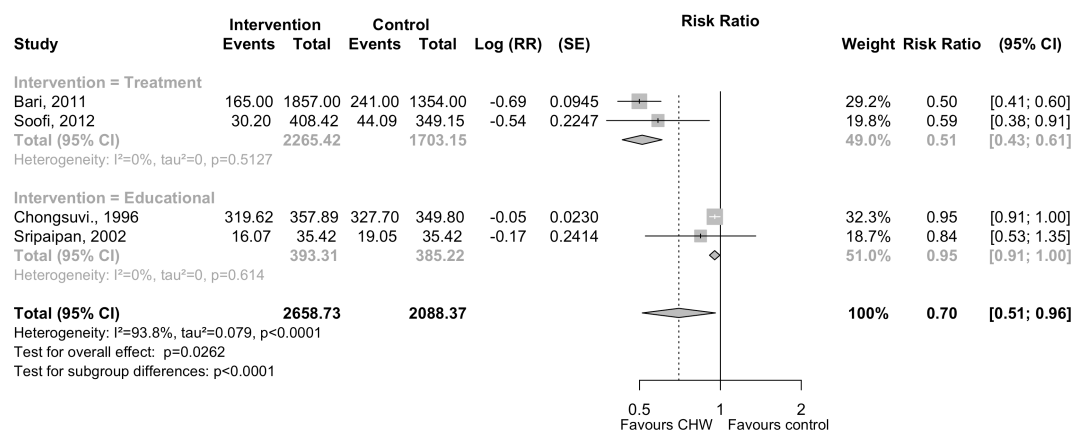


Figure 4.9 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.2 child morbidity – adjusted for clustering

2.3 Infant mortality

Data from two studies were included in a meta-analysis (Fig 4.10). CHWs promoted birth preparedness and childcare behaviours to pregnant mothers in Malawi (Lewycka et al., 2013) and Bangladesh (Sloan et al., 2008). The meta-analysis showed low quality evidence of no difference between CHW delivered interventions and usual care in reducing infant mortality (RR 0.99, 95% CI 0.80 to 1.23, $P = 0.96$). Evidence was downgraded for imprecision and limitations in study design (Sloan et al. 2008 was at high risk of performance and attrition bias). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.67$)

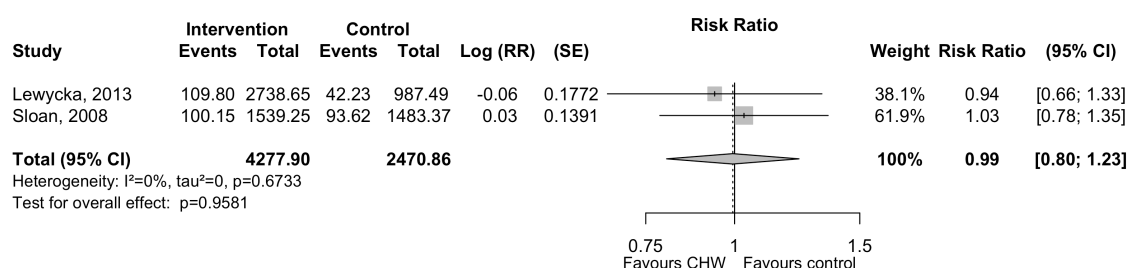


Figure 4.10 Comparison 2: CHW delivered interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.3 Infant mortality – adjusted for clustering

2.4 Infant morbidity

Seven studies reported on this outcome and six were included in the meta-analysis (Fig 4.11). CHWs promoted birth preparedness and childcare behaviours in women's groups or structured home visits to mothers in Southeast Asia, South Asia and Southeast Africa. While some clinical heterogeneity exists across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. region, CHW programme characteristics etc.). The meta-analysis showed low quality evidence CHW delivered intervention may reduce child morbidity (RR 0.73, 95% CI 0.56 to 0.96, $P = 0.023$). Evidence was downgraded due to study design and implementation (it was unclear whether assessors were blinded in the majority of studies), and considerable unexplained statistical heterogeneity in the results ($I^2 = 87.5\%$, $P < 0.0001$). Data from Sloan et al. (2008) was not available in a form that allowed statistical pooling; their trial found no difference in the infants' reported morbidity at the 30- to 45-day follow-up.

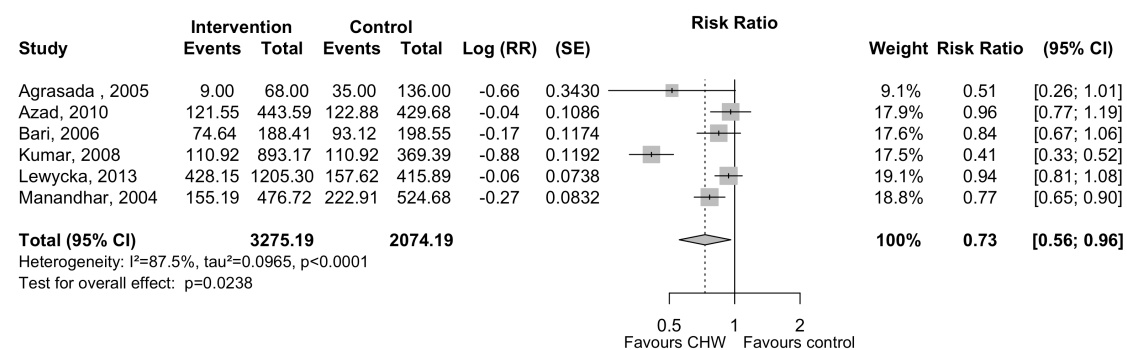


Figure 4.11 Forest plot of comparison 2: Usual care versus community health worker to reduce mortality and morbidity in children under five, outcome: 2.4 Infant morbidity – adjusted for clustering

Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care

Twenty-five reports describing twenty-six trials were examined in this comparison (Agrasada et al., 2005; Azad et al., 2010; Baqui et al., 2008; Colbourn et al., 2013a; Coutinho et al., 2005; Darmstadt et al., 2010; Fottrell et al., 2013; Haider

et al., 2000; Hanson et al., 2015; Kirkwood et al., 2013; Kumar et al., 2008; Leite et al., 2005; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012; Morrow et al., 1999; Penfold et al., 2014; Rotheram-Borus et al., 2014; Sloan et al., 2008; Stanton et al., 2013; Tomlinson et al., 2014; Tripathy et al., 2010; Tripathy et al., 2016; Tylleskar et al., 2011; Waiswa et al., 2015). Seven trials were included in the meta-analysis of outcome 3.1 Stillbirths (Azad et al., 2010; Colbourn et al., 2013a; Fottrell et al., 2013; Kumar et al., 2008; Manandhar et al., 2004; More et al., 2012; Tripathy et al., 2016). Fifteen trials were included in the meta-analysis of outcome 3.2 Neonatal morbidity (Azad et al., 2010; Baqui et al., 2008; Colbourn et al., 2013a; Darmstadt et al., 2010; Fottrell et al., 2013; Hanson et al., 2015; Kirkwood et al., 2013; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012; Sloan et al., 2008; Tomlinson et al., 2014; Tripathy et al., 2010; Tripathy et al., 2016) a subset of these trials reported on outcome 3.3 Early neonatal mortality and 3.4 Late neonatal mortality and were included in these meta-analyses (Azad et al., 2010; Colbourn et al., 2013a; Fottrell et al., 2013; Kumar et al., 2008; Manandhar et al., 2004; Tripathy et al., 2010; Tripathy et al., 2016). Seven trials were included in the meta-analysis for outcome 3.5 Perinatal mortality (Azad et al., 2010; Colbourn et al., 2013a; Fottrell et al., 2013; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; Tripathy et al., 2016). Eleven trials were included in the meta-analysis for outcome 3.6 Maternal mortality (Azad et al., 2010; Colbourn et al., 2013a; Fottrell et al., 2013; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012; Sloan et al., 2008; Tomlinson et al., 2014; Tripathy et al., 2010; Tripathy et al., 2016). Three trials reported on outcome 3.7 Maternal morbidity (Lewycka et al., 2013; Manandhar et al., 2004; Stanton et al., 2013). Thirteen trials were included in the meta-analysis for outcome 3.8 Facility based delivery (Azad et al., 2010;

Colbourn et al., 2013a; Darmstadt et al., 2010; Fottrell et al., 2013; Hanson et al., 2015; Kirkwood et al., 2013; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012; Tripathy et al., 2010; Tripathy et al., 2016; Waiswa et al., 2015). Four trials were included in the meta-analysis for outcome 3.9 Clean delivery kits (Azad et al., 2010; Manandhar et al., 2004; Tripathy et al., 2010; Tripathy et al., 2016). Ten studies were included in the meta-analysis for outcome 3.10 Exclusive breastfeeding (Agrasada et al., 2005; Coutinho et al., 2005; Haider et al., 2000; Leite et al., 2005; Lewycka et al., 2013; Morrow et al., 1999; Rotheram-Borus et al., 2014; Tylleskar et al., 2011). Eleven trials were included in the meta-analysis for outcome 3.11 Breastfeeding initiation within one hour of birth (Baqui et al., 2008; Darmstadt et al., 2010; Hanson et al., 2015; Kirkwood et al., 2013; Kumar et al., 2008; Lewycka et al., 2013; Manandhar et al., 2004; Sloan et al., 2008; Tomlinson et al., 2014; Tripathy et al., 2016; Waiswa et al., 2015).

Setting: Six trials were conducted in upper-middle-income countries, thirteen in lower-middle-income countries, and seven in low-income countries. Six trials were conducted in Bangladesh (Azad et al., 2010; Baqui et al., 2008; Darmstadt et al., 2010; Fottrell et al., 2013; Haider et al., 2000; Sloan et al., 2008); four in India (Kumar et al., 2008; More et al., 2012; Tripathy et al., 2010; Tripathy et al., 2016); three in South Africa (Rotheram-Borus et al., 2014; Tomlinson et al., 2014; Tylleskar et al., 2011-SA); two in Malawi (Colbourn et al., 2013; Lewycka et al., 2013), Uganda (Tylleskar et al., 2011-UG; Waiswa et al., 2015), Ghana (Kirkwood et al., 2013; Stanton et al., 2013), and Brazil (Coutinho et al., 2005; Leite et al., 2005); and one in the Philippines (Agrasada et al., 2005), Nepal (Manandhar et al., 2004), Mexico (Morrow et al., 1999), Tanzania (Hanson et al., 2015), and Burkina Faso (Tylleskar et al., 2011-BF).

Participants: All studies were conducted on women of reproductive age and their newborn children. CHW recruitment was via community input (Penfold et al., 2014; Tripathy et al., 2010), the health system (Sloan et al., 2008; Stanton et al., 2013), the health system and the host villages together (Lewycka et al., 2013; Manandhar et al., 2004), an NGO (Baqui et al., 2008; Colbourn et al., 2013), local advertisements (Darmstadt et al., 2010), the researcher team (Agrasada et al., 2005; More et al., 2012; Tripathy et al., 2016) or by unspecified actors. In eleven trials, scholastic attainment was a recruitment criterion: requirements ranged from basic literacy (Lewycka et al., 2013; Manandhar et al., 2004) or a minimum of 4 years of education (Haider et al., 2000) to completion (Kumar et al., 2008; More et al., 2012) or near completion (Baqui et al. 2008; Colbourn et al., 2013; Darmstadt et al., 2010; Rotheram-Borus et al., 2014; Tomlinson et al. 2014; Tylleskar et al., 2011-SA) of secondary school. Previous breastfeeding experience was required for CHWs in three trials (Agrasada et al., 2005; Haider et al., 2000; Leite et al., 2005) while being female was a prerequisite in seven others (Azad et al., 2010; Colbourn et al., 2013; Fottrell et al., 2013; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012, Tripathy et al., 2010). Other criteria included performing well in a training course (Agrasada et al. 2005).

Description of interventions: In nineteen trials, CHWs promoted birth and newborn-care preparedness and/or breastfeeding counselling during structured home visits (Agrasada et al., 2005; Baqui et al., 2008; Coutinho et al., 2005; Darmstadt et al., 2010; Haider et al., 2000; Hanson et al., 2015; Kirkwood et al., 2013; Kumar et al., 2008; Leite et al., 2005; Lewycka et al., 2013; Morrow et al., 1999; Rotheram-Borus et al., 2014; Sloan et al., 2008; Tomlinson et al., 2014; Tylleskar et al., 2011; Waiswa et al., 2015). In seven trials, CHWs led women's groups through an action-

learning cycle in which the women identified local health problems and formulated strategies to address them (Azad et al., 2010; Colbourn et al., 2013; Fottrell et al., 2013; Lewycka et al., 2013; Manandhar et al., 2004; More et al., 2012; Tripathy et al., 2010; Tripathy et al., 2016). In one trial (Stanton et al., 2013) CHWs provided one injection of oxytocin one minute after birth. Six studies had multiple trial arms (Agrasada et al., 2005; Baqui et al., 2008; Colbourn et al., 2013a; Kumar et al., 2008; Lewycka et al., 2013; Morrow et al., 1999); in line with the review protocol, eligible intervention groups in each study were combined to estimate an average treatment effect compared with the control (for additional detail on the trial arms in each study see Characteristics of Included Studies Tables in Appendix 4c). The ratio of CHWs to population ranged from 1:8 (Agrasada et al., 2005) to 1:4000 (Baqui et al., 2008; Darmstadt et al., 2010). Reported CHW training ranged from 5 days (Agrasada et al., 2005; Haider et al., 2000; Lewycka et al., 2013; Stanton et al., 2013; Waiswa et al., 2015) to 3 months (Morrow et al., 1999). Initial instruction was followed by “ongoing” support (Colbourn et al., 2013; More et al., 2012) or annual refresher trainings (Lewycka et al., 2013).

Incentives were described in twelve studies; in seven studies CHWs were provided a salary or incentivised payments (Haider et al., 2000; Kumar et al., 2008; Leite et al., 2005; Rotheram-Borus et al., 2014; Sloan et al., 2008; Tomlinson et al., 2014; Tripathy et al., 2016), in two studies they were reimbursed for expenses (Agrasada et al., 2005; Colbourn et al., 2013a), and in three studies they received monetary and non-monetary (e.g. bicycle, calendars) incentives (Kirkwood et al., 2013; Lewycka et al., 2013; Waiswa et al., 2015).

Supervision was described in eight studies. It was provided by certified health workers (Agrasada et al. 2005), local field supervisors (Azad et al., 2010; Colbourn et

al., 2013; More et al., 2012; Tripathy et al., 2010), NGO staff (Haider et al., 2000; Tripathy et al., 2016), government health surveillance assistants (Lewycka et al., 2013) or by a supervisor for which no characteristics were provided (Kumar et al., 2008; Manandhar et al., 2004; Tomlinson et al. 2014). The frequency of supervision was biweekly (Rotheram-Borus et al., 2014), weekly (More et al., 2012; Tomlinson et al. 2014), fortnightly (Darmstadt et al., 2010; Tripathy et al., 2010; Tylleskar et al., 2011-BF-UG-SA), monthly (Kirkwood et al., 2013; Penfold et al., 2014; Tripathy et al., 2016), or quarterly (Haider et al., 2000; Lewycka et al., 2013; Waiswa et al., 2015).

Comparison groups included usual care without the addition of a CHW (Agrasada et al., 2005; Baqui et al., 2008; Coutinho et al., 2005; Darmstadt et al., 2010; Haider et al., 2000; Kumar et al., 2008; Leite et al., 2005; Lewycka et al., 2013; More et al., 2012; Morrow et al., 1999; Penfold et al., 2014; Rotheram-Borus et al., 2014; Sloan et al., 2008; Tripathy et al., 2010; Tripathy et al., 2016; Tylleskar et al., 2011-BF-UG), usual care where the CHW was not trained to deliver the intervention in question (Stanton et al., 2013; Tomlinson et al., 2014; Tylleskar et al., 2011-SA), and health-facility strengthening activities undertaken in both intervention and control areas (Azad et al., 2010; Fottrell et al., 2013; Kirkwood et al., 2013; Manandhar et al., 2004; Waiswa et al., 2015). In Colbourn et al. (2013) there was two control arms: no intervention and health facility strengthening.

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison three. Neonatal mortality, maternal mortality, facility-based delivery, and exclusive breastfeeding are also featured in Summary of Findings 3. Data reported at longest follow-up were included unless otherwise specified.

3.1 Stillbirths

Data from seven trials were included in the meta-analysis (Fig 4.12). CHWs promoted newborn-care preparedness in women's groups or structured home visits to mothers in South Asia and Southeast Africa. The meta-analysis showed moderate quality evidence of no difference between CHW delivered interventions and usual care in reducing stillbirths (RR 0.89, 95% CI 0.69 to 1.13, $P = 0.34$). Evidence was downgraded due to imprecision. There was no evidence of statistical heterogeneity ($I^2 = 0\%$, $P = 0.93$).

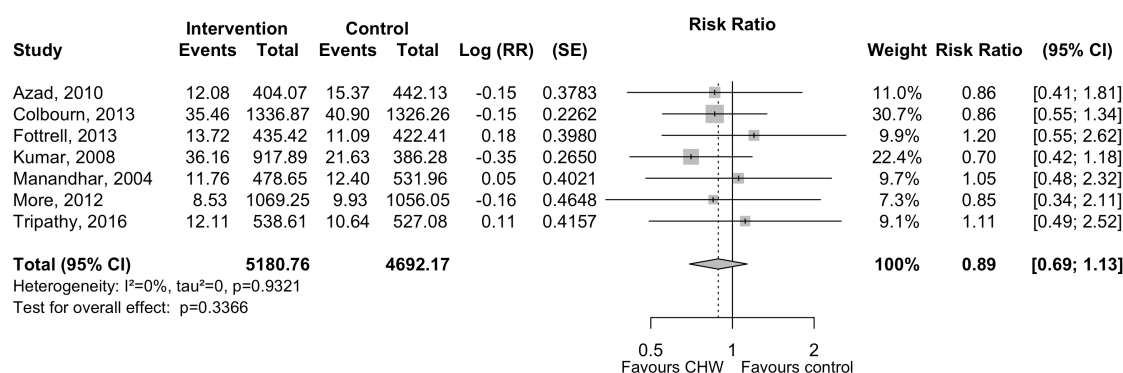


Figure 4.12 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.1 Stillbirths – adjusted for clustering

3.2 Neonatal mortality (3.3 Early NMR and 3.4 Late NMR)

Fifteen trials were included in the meta-analysis (Fig 4.13). CHWs promoted birth preparedness and childcare behaviours in women's groups or structured home visits to mothers in South Asia and across Africa. The meta-analysis showed moderate quality evidence that CHW delivered interventions reduce neonatal mortality (RR 0.76, 95% CI 0.61 to 0.96, $P = 0.02$). Evidence was downgraded for substantial unexplained heterogeneity ($I^2 = 75.8\%$, $P < 0.0001$). One study at low risk of bias, Lewycka et al., (2013) appears to be an outlier, though it is not clear from the extracted characteristics why this would be the case. When this study was excluded in a sensitivity analysis, statistical heterogeneity reduced and the magnitude of effect dropped slightly (RR 0.86, 95% CI 0.76 to 0.96, $P = 0.009$, $I^2 = 36\%$, $P = 0.087$).

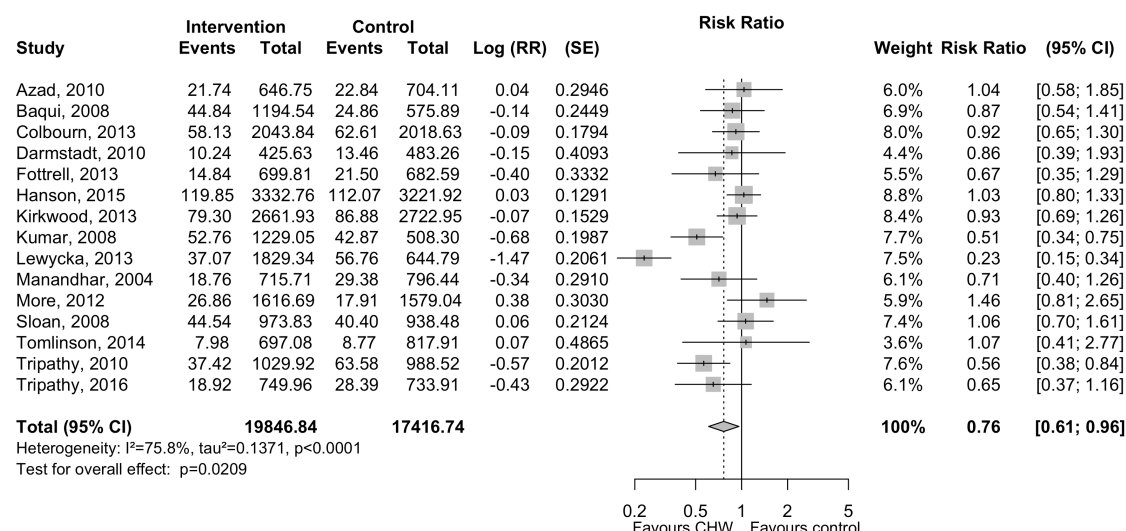


Figure 4.13 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.2.Neonatal mortality – adjusted for clustering

Seven of these trials reported on outcome 3.3 Early neonatal mortality and outcome 3.4 Late neonatal mortality and were included in these meta-analyses (Appendix 4d). The meta-analysis showed moderate quality evidence CHW delivered interventions are likely to both reduce early neonatal mortality (RR 0.69, 95% CI 0.57 to 0.85, $P = 0.0004$) and late neonatal mortality (RR 0.70, 95% CI 0.49 to 1.01, $P = 0.053$). There was no evidence of heterogeneity in either the early neonatal mortality ($I^2 = 0.5\%$, $P = 0.42$) or late neonatal mortality ($I^2 = 0\%$, $P = 0.48$) trials.

3.5 Perinatal mortality

Seven trials were included in the meta-analysis (Fig 4.14). CHWs promoted birth preparedness and childcare behaviours in women's groups or structured home visits to mothers in South Asia and Southeast Africa. The meta-analysis showed high quality evidence that CHW delivered interventions reduce perinatal mortality (RR 0.83, 95% CI 0.72 to 0.95, $P = 0.009$). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.78$).

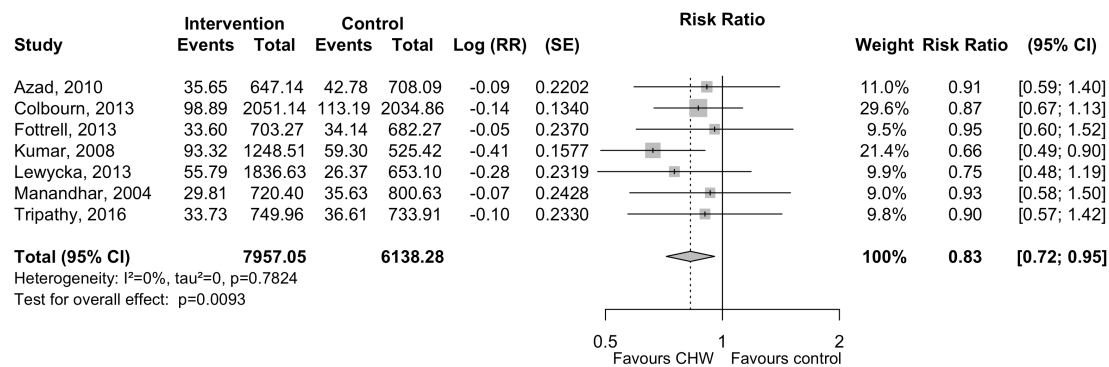


Figure 4.14 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.5.Perinatal mortality – adjusted for clustering

3.6 Maternal mortality

Eleven trials were included in the meta-analysis (Fig 4.15). CHWs promoted birth preparedness in women's groups or structured home visits to mothers in South Asia and South Africa. The meta-analysis showed moderate quality evidence of no difference between CHW delivered interventions and usual care for reducing maternal mortality (RR 0.66, 95% CI 0.35 to 1.25, $P = 0.20$). Evidence was downgraded due to imprecision. There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.999$).

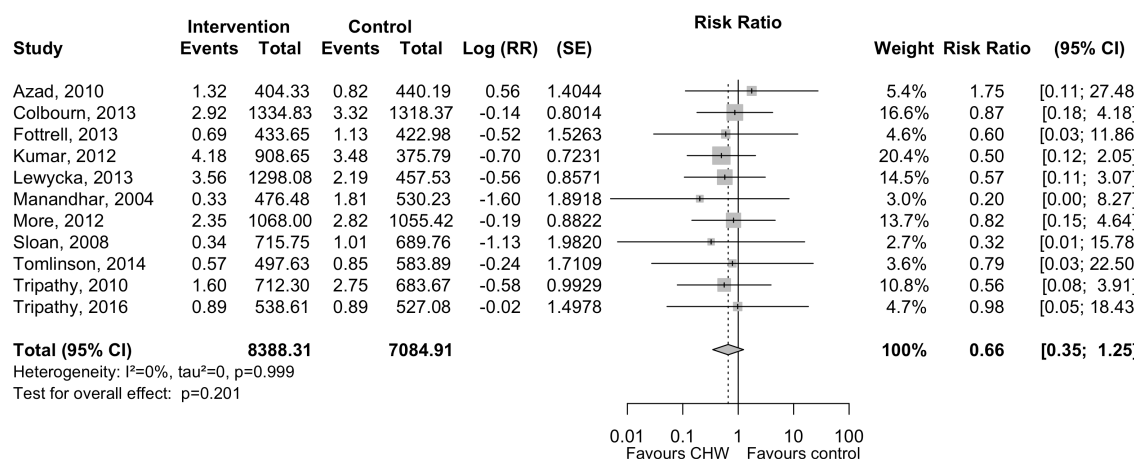


Figure 4.15 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.6. Maternal mortality – adjusted for clustering

3.7. Maternal morbidity

Three trials reported on maternal morbidity; two measured any perceived antenatal, delivery, or postnatal maternal problems and were included in the meta-analysis (Fig 4.16). The trials were conducted in South Asia and across Africa. In the

two meta-analysed trials, CHWs promoted birth preparedness in women's groups or structured home visits to mothers. In Stanton et al., (2013) CHWs provided one injection of oxytocin one minute after birth. The meta-analysis showed low quality evidence that CHW delivered interventions have no effect on maternal morbidity (RR 1.04, 95% CI 0.63 to 1.73, $P = 0.88$). Evidence was downgraded due to imprecision and a considerable degree of statistical heterogeneity ($I^2 = 93.4\%$, $P < 0.0001$). The fact that Lewycka et al. (2013) had baseline imbalances after randomisation favouring the control group may explain this discrepancy.

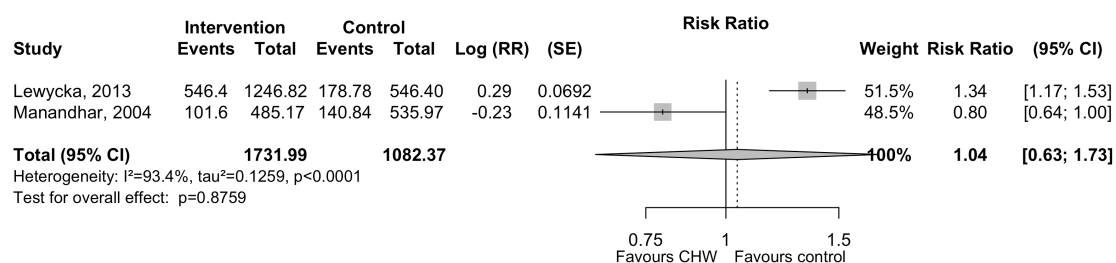


Figure 4.16 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.7. Maternal morbidity – adjusted for clustering

The third trial, Stanton et al. 2013, reported only on postpartum haemorrhage: PPH-1 was defined as blood loss <500 mL; PPH-2 as PPH-1 plus women who received early treatment for PPH; and PPH-3 as PPH-2 plus any other women referred to hospital for postpartum bleeding. Adjusted for clustering, PPH-1 rates were 2.6% in the intervention group versus 5.5% in the control group (RR: 0.49; 95% CI: 0.27–0.88); PPH-2 rates were 3.8% versus 10.8% (RR: 0.35; 95% CI: 0.18–0.63), and PPH-3 rates were similar to those of PPH-2 (4.1 % vs. 11.1%, RR 0.361; 95% CI: 0.196–0.664). This suggests that maternal health care planners can consider extending the use of oxytocin into peripheral settings with the use of CHWs.

3.8 Facility based delivery

Thirteen trials were included in the meta-analysis (Fig 4.17). CHWs promoted birth preparedness in women's groups or structured home visits to mothers in South

Asia and across Africa. While some clinical heterogeneity exists across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. region, CHW programme characteristics etc.). The meta-analysis showed low quality evidence that CHW delivered interventions may increase facility-based delivery (RR 0.94, 95% CI 0.86 to 1.02, $P = 0.14$). Evidence was downgraded due to imprecision and a considerable degree of unexplained statistical heterogeneity ($I^2 = 92.4\%$, $P < 0.0001$). One study at low risk of bias, Hanson et al. (2015), appears to be an outlier, having reported a dramatic improvement of facility-based delivery. It is not clear from the extracted clinical characteristics why this would be the case. When this study was excluded in a sensitivity analysis, statistical heterogeneity dropped to a small degree and the magnitude of effect was reduced slightly (RR 0.97, 95% CI 0.92 to 1.03, $P = 0.3$, $I^2 = 85\%$, $P < 0.0001$).

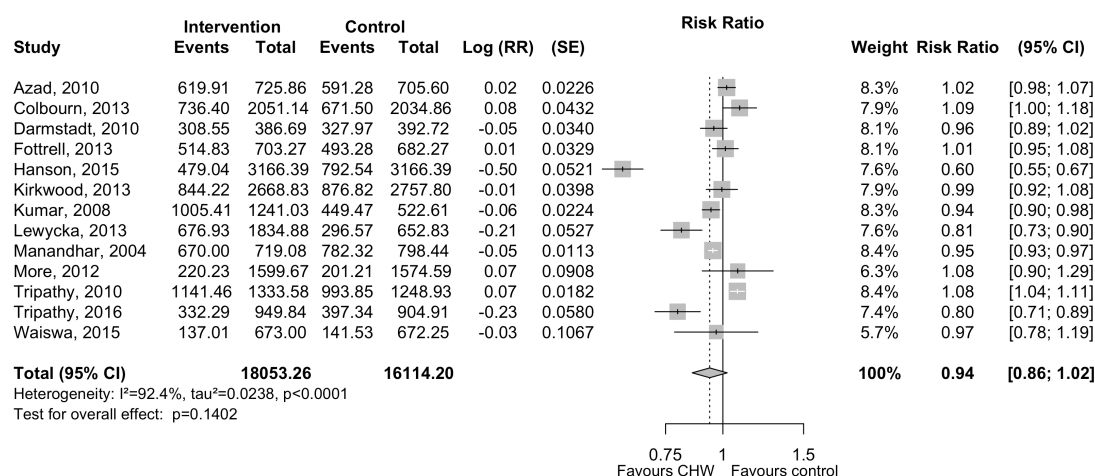


Figure 4.17 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.8. Facility-based delivery – adjusted for clustering

3.9 Clean delivery kits

Four trials were included in the meta-analysis for outcome (Fig 4.18). CHWs led women's groups in South Asia through an action-learning cycle in which the women identified local health problems and formulated strategies to address them. This showed low quality evidence that CHW delivered interventions increase the

uptake of clean delivery kits (RR 0.89, 95% CI 0.83 to 0.96, $P < 0.0001$). Evidence was downgraded due to study design and implementation (lack of blinding and baseline imbalances were present or unclear in the majority of studies) and considerable unexplained statistical heterogeneity ($I^2 = 96\%$, $P < 0.0001$). One study, Tripathy et al. (2016), appears to be an outlier. It is not clear from the extracted clinical or methodological characteristics why this would be the case. When this study was excluded in a sensitivity analysis, statistical heterogeneity dropped and the magnitude of effect was increased slightly (RR 0.86, 95% CI 0.82 to 0.89, $P < 0.001$, $I^2 = 46.6\%$, $P < 0.15$).

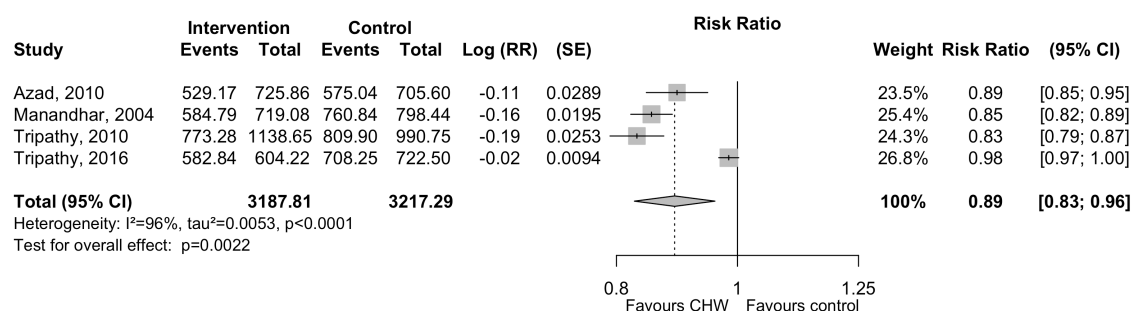


Figure 4.18 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – adjusted for clustering

3.10 Exclusive breastfeeding up to six months post partum

Data from ten trials on the outcome were included the meta-analysis (Fig 4.19). CHWs promoted birth and newborn-care preparedness and/or breastfeeding counselling during structured home visits to women across Asia, Africa, and Latin America. This showed moderate quality evidence that CHW delivered interventions had a substantial impact on exclusive breastfeeding up to six months postpartum (RR 0.65, 95% CI 0.52 to 0.83, $P = 0.0005$). Evidence was downgraded due to considerable statistical heterogeneity in the results ($I^2 = 97.3\%$, $P < 0.0001$). Post-hoc subgroup analysis suggests, however, that CHW delivered interventions devoted

exclusively to breastfeeding counselling tended to produce larger effect sizes (RR 0.60, 95% CI 0.46 to 0.78, $P = 0.0002$, eight studies) than those devoted to breastfeeding counselling in addition to more general newborn preparedness (RR 0.93, 95% CI 0.90 to 0.96, $P < 0.0001$, two studies)—though the interaction is quantitative, rather than qualitative (i.e. the size of the effect varies but not the direction, see: Yusuf et al., 1991). As with all subgroup analyses, these results should be interpreted with caution; they are observational in nature and so the results could be confounded by other study-level characteristics and do not indicate that the subgroup factor explains heterogeneity.

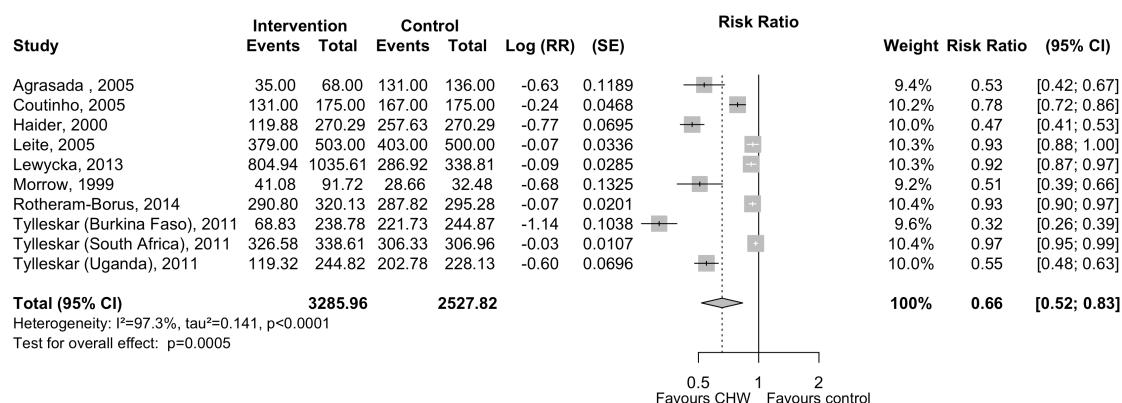


Figure 4.19 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – adjusted for clustering

3.11 Breastfeeding initiation within one hour of birth

Eleven trials were included in the meta-analysis (Fig 4.20). CHWs promoted newborn-care preparedness during structured home visits or women's group meetings in South Asia and across Africa. While some clinical heterogeneity exists across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. region, group-based vs. one-on-one intervention etc.). This showed moderate quality evidence that CHW delivered interventions increase the uptake of the initiation of breastfeeding within one hour of birth (RR 0.61, 95% CI

0.45 to 0.82, $P = 0.001$). Evidence was downgraded due to considerable unexplained statistical heterogeneity in the results ($I^2 = 98\%$, $P < 0.0001$).

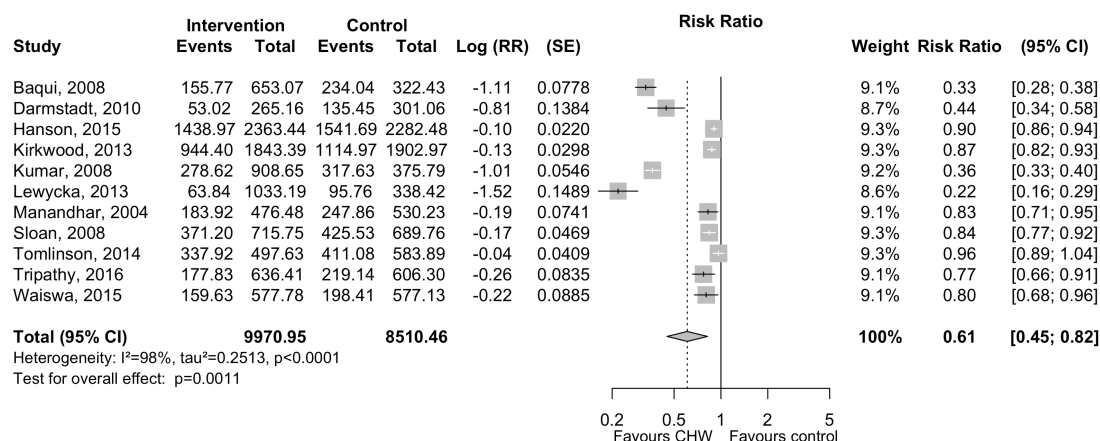


Figure 4.20 Comparison 3: CHW delivered interventions to improve maternal and newborn health compared to usual care, Outcome: 3.11. Initiation of breastfeeding within 1 h of birth – adjusted for clustering

Comparison 4: CHW delivered interventions to promote cognitive development in infants

Eight studies were examined in this comparison (Grantham-McGregor et al., 1991; Murray et al., 2015; Powell et al., 2004; Powell & Grantham-McGregor, 1989; Rahman et al., 2009; Vazir et al., 2013; Walker et al., 2004; Yousafzai et al., 2014). Two trials were included in the meta-analysis of Outcome 4.1 Maternal knowledge of infant development (Rahman et al., 2009; Yousafzai et al., 2014). Seven trials were included in the meta-analysis of outcome 4.2 Cognitive Development (Grantham-McGregor, Powell, Walker, & Himes, 1991; Murray, Cooper, Arteche, Stein, & Tomlinson, 2015; Powell, Baker-Henningham, Walker, Gernay, & Grantham-McGregor, 2004; Powell & Grantham-McGregor, 1989; Vazir et al., 2013; Walker, Chang, Powell, & Grantham-McGregor, 2004; Yousafzai et al., 2014).

Setting: Five trials were conducted in upper middle-income countries and three in lower-middle-income countries. Four trials were conducted in Jamaica (Grantham-McGregor et al., 1991; Powell & Grantham-McGregor, 1989; Powell et

al., 2004; Walker et al., 2004); two trials were conducted in Pakistan (Rahman et al., 2009; Yousafzai et al., 2014); and one trial was conducted in both India (Vazir et al., 2013) and South Africa (Murray et al., 2015).

Participants: In four studies, infants were enrolled at birth (Murray et al., 2015; Rahman et al., 2009; Walker et al., 2004; Yousafzai et al., 2014); in five studies, infants were enrolled after birth: at 3 months in Vazir et al. (2013), 9-24 months in Grantham-McGregor et al. (1991); 9-30 months in Baker-Henningham et al. (2005), and 16-30 months in Powell & Grantham-McGregor (1989). In three of these studies, included infants were malnourished: low birth weight (< 2500 g) term in Walker et al. (2004); stunted (lengths below -2 standard deviations of the National Center for Health Statistics reference data for age and sex) in Grantham-McGregor et al. (1991); and undernourished (WAZ <-1.5) in Baker-Henningham et al. (2005).

Two studies used existing government CHWs (Rahman et al., 2009; Powell et al., 2004); the recruiter in the other studies was not specified. Recruitment criteria ranged from 8 years of education (Yousafzai et al., 2014) to 10 years of education (Walker et al., 2004).

Description of interventions: In all eight studies, CHWs provided psychosocial stimulation. Three of these studies had multiple intervention groups. In Yousafzai et al. (2014) and Grantham-McGregor et al. (1991) CHWs delivered nutritional supplementation in one arm, responsive stimulation in the second arm, or a combination of both interventions in the third intervention arm. In Vazir et al. (2013) CHWs delivered an education intervention on complementary feeding in arm A or the same intervention in arm B plus psychosocial stimulation. For the three studies with multiple trial arms (Grantham-McGregor et al., 1991; Vazir et al., 2013; Yousafzai et al., 2014), eligible intervention groups in each study were, in line with the review

protocol, combined to estimate an average treatment effect compared with the control (for additional detail on the trial arms in each study see Characteristics of Included Studies Tables in Appendix 4c). The ratio of CHWs to population ranged from 1:8 (Rahman et al., 2009) to 1:1000 (Yousafzai et al., 2014).

CHW training was described in four studies and ranged from one day with a one-hour refresher session (Rahman et al., 2009) to three months classroom-based training and 12 months field-based training with 15 days of refresher training per year (Yousafzai et al., 2014).

Incentives were described in three studies; in all three studies CHWs were provided a salary or other monetary payment (Powell & Grantham-McGregor, 1989; Powell et al., 2004; Yousafzai et al., 2014).

Supervision was described in six studies. It was provided by a nurse-supervisor (Powell & Grantham-McGregor, 1989; Powell et al., 2004), a lady health supervisor (Yousafzai et al., 2014), a community clinical psychologist (Murray et al., 2015), graduates in nutrition (Vazir et al., 2013) or by a supervisor for which no characteristics were provided (Grantham-McGregor et al., 1991;). The frequency with which it was undertaken ranged from weekly (Murray et al., 2015; Powell & Grantham-McGregor, 1989) to monthly (Powell et al., 2004).

Comparison groups included usual care without the addition of a CHW (Powell & Grantham-McGregor, 1989; Powell et al., 2004; Walker et al., 2004; Yousafzai et al., 2014) and usual care where the CHW was not trained to deliver the intervention in question (Grantham-McGregor et al., 1991; Murray et al., 2015; Rahman et al., 2009; Vazir et al., 2013).

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison four. Infant cognitive development is included in

Summary of Findings 4. Data reported at longest follow-up were included unless otherwise specified.

4.1 Maternal knowledge of infant development

Two trials, conducted in Pakistan, were included in the meta-analysis (Fig 4.21). CHWs provided psychosocial stimulation to infants starting at birth. This showed low quality evidence that CHW delivered interventions have a large effect on maternal knowledge of infant development (SMD -1.37, 95% CI -3.03 to 0.29, $P = 0.11$). Evidence was downgraded for imprecision and considerable statistical heterogeneity ($I^2 = 97.8\%$, $P < 0.0001$). As only two studies report on this outcome it was not possible to conduct sensitivity analysis on clinical differences between studies (e.g. related to CHW programme characteristics) that may relate to the difference in effect sizes.

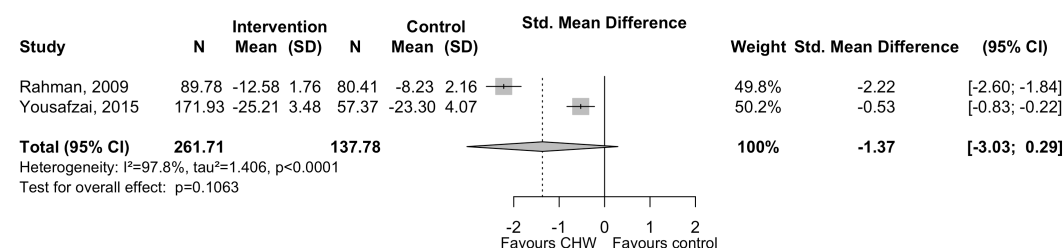


Figure 4.21 Comparison 4: CHW delivered interventions to promote cognitive development in infants, Outcome: 4.1 Maternal knowledge of infant development – adjusted for clustering

4.2 Cognitive development

Seven trials were included in the meta-analysis (Fig 4.22). CHWs provided psychosocial stimulation to infants in South Asia, Africa, and the Caribbean. While some clinical heterogeneity existed across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. whether the infants were malnourished, supervisor qualifications etc.) The meta-analysis showed low quality evidence that CHW delivered interventions have a moderate effect on infant cognitive development (SMD -0.56, 95% CI -0.86 to -0.27, $P = 0.0002$). Evidence

was downgraded for design and implementation (adequate sequence generation and allocation concealment were unclear or high risk in five of the trials) and substantial unexplained heterogeneity ($I^2 = 80.8\%$, $P < 0.0001$).

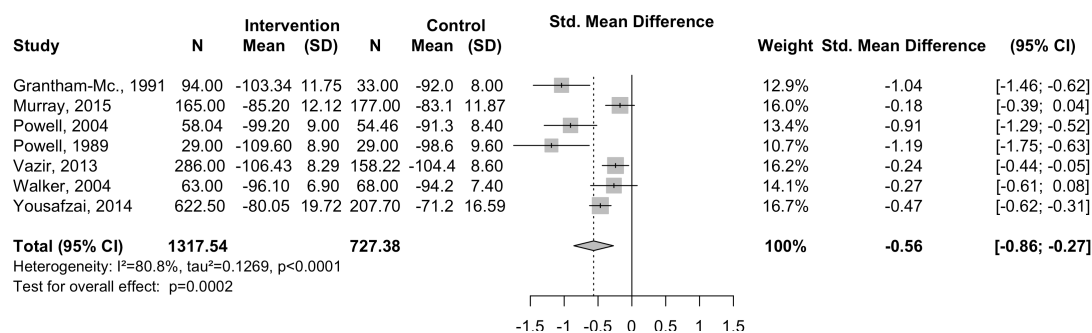


Figure 4.22 Comparison 4: CHW delivered interventions to promote cognitive development in infants, Outcome: 4.2 Cognitive development – adjusted for clustering

Comparison 5: CHW delivered interventions to promote healthy feeding practices

Seven trials were examined in this comparison (le Roux et al., 2010; le Roux et al., 2011; Pachón et al., 2002; Rahman et al., 2008; Saleem et al., 2014; Vazir et al., 2013; Yousafzai et al., 2014). Seven trials reported on outcome 5.1 Weight-for-Age Z Score (WAZ) (le Roux et al., 2010; le Roux et al., 2011; Pachón et al., 2002; Rahman et al., 2008; Saleem et al., 2014; Vazir et al., 2013; Yousafzai et al., 2014). Three trials were included in the meta-analysis of outcome 5.2 Height-for-Age Z Score (HAZ) (Pachón et al., 2002; Rahman et al., 2008; Yousafzai et al., 2014). Two trials were included in the meta-analysis of outcome 5.2 Weight-for-Height Z Score (WHZ) (Pachón et al., 2002; Yousafzai et al., 2014).

Setting: Two trials were conducted in upper-middle-income countries and five in lower-middle-income countries. Three trials were conducted in Pakistan (Rahman et al., 2008; Saleem et al., 2014; Yousafzai et al., 2014), two South Africa (le Roux et al., 2010; le Roux et al., 2011), and one in both Vietnam (Pachón et al., 2002) and India (Vazir et al., 2013).

Participants: Six studies were conducted on children under 25 months and their caregivers (le Roux et al., 2010; Pachón et al., 2002; Rahman et al., 2008; Saleem et al., 2014; Vazir et al., 2013; Yousafzai et al., 2014); the remaining study (le Roux et al. 2011) included children aged 0-6 years. In three studies (le Roux et al., 2010; le Roux et al. 2011; Pachón et al., 2002) children were malnourished at the start of the intervention, while in one (Rahman et al., 2008) the mothers had perinatal depression.

CHW recruitment was via the research team (le Roux et al. 2011), the health system (Rahman et al., 2008), an NGO (le Roux et al., 2010) or by unspecified actors. In three trials, recruited CHWs were female (le Roux et al., 2010; le Roux et al., 2011; Rahman et al., 2008; Saleem et al., 2014; Vazir et al., 2013; Yousafzai et al., 2014). Education requirements ranged from 8 years (Yousafzai et al., 2014), to 10 years (Saleem et al., 2014), to 12 years (Vazir et al., 2013).

Description of interventions: Interventions included home visits (le Roux et al., 2010; le Roux et al., 2011) or intensive nutrition rehabilitation sessions with growth monitoring and nutrition counselling (Pachón et al., 2002), cognitive behaviour therapy-like problem solving (Rahman et al., 2008), or breastfeeding and complementary feeding education (Saleem et al., 2014). Yousafzai et al., 2014 was a 2x2 factorial trial with a nutrition education and multiple micronutrient powders arm, a responsive stimulation arm, and a combination arm. Vazir et al. (2013) had two intervention arms: in Arm A, CHWs delivered an education intervention on complementary feeding; in Arm B CHWs delivered the same intervention as in arm A plus psychosocial stimulation. In Yousafzai et al., (2014) eligible intervention groups in each study were combined to estimate an average treatment effect compared with the control; Vazir et al. (2013) was not meta-analysed (for additional detail on the trial

arms in each study, see Characteristics of Included Studies Tables in Appendix 4c). The ratio of CHWs to population ranged from 1:12 (Rahman et al., 2008) to 1:1000 (Yousafzai et al., 2014).

CHW training was described in all seven trials. Duration ranged from 2 days (Rahman et al., 2008) to three months (Yousafzai et al., 2014) and initial instruction was followed by up to one year of in-service training (Yousafzai et al., 2014) or a refresher at four months (Rahman et al., 2008).

Supervision was described in six studies. It was provided by local field supervisors (Yousafzai et al., 2014), a commune management steering committee (Pachón et al., 2002), the research team (Rahman et al., 2008), trained graduates in nutrition (Vazir et al., 2013), or by a supervisor for which no characteristics were provided (le Roux et al., 2010; le Roux et al., 2011) on a monthly (le Roux et al., 2010; le Roux et al., 2011; Pachón et al., 2002; Rahman et al., 2008) or unspecified basis (Yousafzai et al., 2014).

Incentives were described in five studies. In four studies CHWs were provided a salary or incentivised payments (le Roux et al., 2010; le Roux et al., 2011; Rahman et al., 2008; Yousafzai et al., 2014) and in one study no incentives were provided (Pachón et al., 2002).

Comparison groups included usual care without the addition of a CHW (le Roux et al., 2010; le Roux et al., 2011; Pachón et al., 2002; Saleem et al., 2014; Vazir et al., 2013) or usual care where the CHW was not trained to deliver the intervention in question (Rahman et al., 2008; Yousafzai et al., 2014).

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison five. WAZ, HAZ, and WHZ are also featured in

Summary of Findings 5. Data reported at longest follow-up were included unless otherwise specified.

5.1 Weight-for-age Z score

Seven trials reported infant weight measures; three of these trials were included in the meta-analysis of weight-for-age z scores (Fig 4.23). CHWs provided home-based nutritional education and/or rehabilitation to children in South Asia, Southeast Asia, and Africa. The meta-analysis showed high quality evidence that CHW delivered interventions have a small effect on improving WAZ (MD -0.12, 95% CI -0.21 to -0.02, $P = 0.0185$). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.91$).

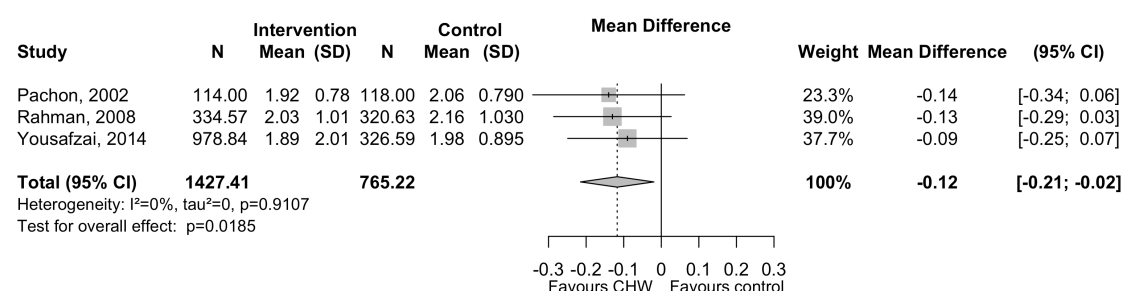


Figure 4.23 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.1 Weight-for-age Z score – adjusted for clustering

Data in the four remaining trials (le Roux et al., 2010; le Roux et al., 2011; Saleem et al., 2014; Vazir et al., 2013) were not available in a form that allowed statistical pooling. In le Roux et al. (2010) low-weight infants in the intervention condition were more likely to rehabilitate (reach a healthy weight for their ages) than children in the control condition (43% vs. 31%, $\chi^2 = 11.24$, $P < 0.01$) over the course of one year. In le Roux et al. (2011) data on WAZ were provided, but without a measure of variability; authors were contacted but were unable to supply this data. The study found that children in the intervention condition gained significantly more weight (WAZ -2.28) than children in the control condition (WAZ -2.45) over 12 months. Saleem et al. (2014) dichotomised WAZ scores into underweight (-2 standard

deviations) or not. There was a 6% reduction in the proportion of under-weight infants in the intervention group and a 9% increase in underweight infants in the control clusters 7.5 months after baseline. In Vazir et al. (2013) no significant differences in change in weight-for-age were observed between the control and intervention groups at one year of follow-up in either unadjusted or adjusted analyses. These studies are largely consistent with the meta-analysed results in showing that CHW-led interventions can have a small but statistically significant impact on WAZ.

5.2 Height-for-age Z score

Three trials were included in the meta-analysis (Fig 4.24). CHWs provided home-based nutritional education and/or rehabilitation to children in South and Southeast Asia. This showed low quality evidence that CHW delivered interventions have no effect on improving HAZ (MD -0.10, 95% CI -0.29 to 0.09, $P = 0.31$). Evidence was downgraded for imprecision and due to substantial unexplained heterogeneity ($I^2 = 46.7\%$, $P = 0.15$). As only three studies report on this outcome it was not possible to conduct sensitivity analysis on clinical differences between studies (e.g. type of CHW delivered intervention) that may relate to the difference in effect sizes.

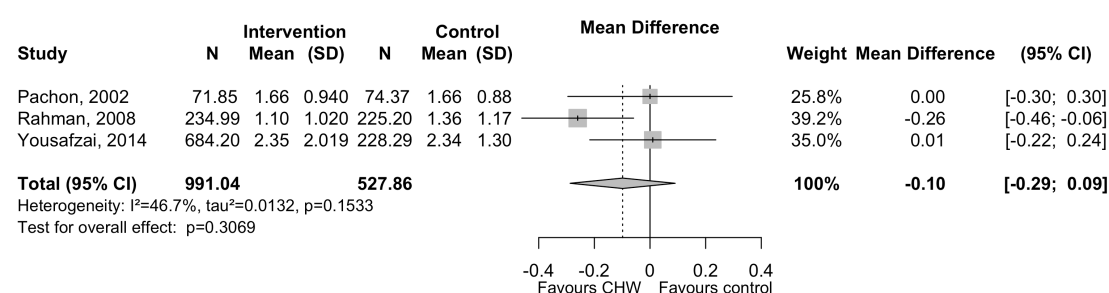


Figure 4.24 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.2 Height-for-age Z score – adjusted for clustering

5.3 Weight-for-height Z score

Two trials were included in the meta-analysis (Fig 4.25). CHWs provided home-based nutritional education and/or rehabilitation to children in Pakistan

(Yousafzai et al., 2014) and Vietnam (Pachón et al., 2002). This showed moderate quality evidence that CHW delivered interventions have a slight effect on improving WHZ (MD -0.14, 95% CI -0.27 to -0.01, $P = 0.04$). Evidence was downgraded due to study design and implementation (assessors were not blinded in Pachón 2002). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 1$).

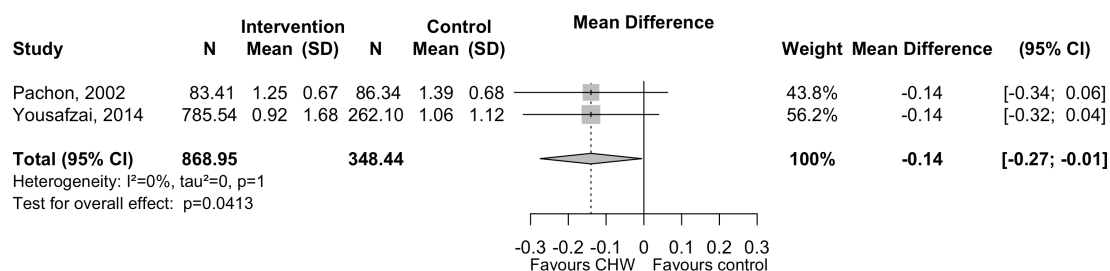


Figure 4.25 Forest plot of comparison 5: CHW delivered interventions to promote healthy feeding practices, outcome: 5.3 Weight-for-height Z score – adjusted for clustering

INFECTIOUS DISEASE OUTCOMES

Comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support

Eight studies were examined in this comparison (CDI Study Group, 2010; Clarke, 2005; Datiko & Lindtjorn, 2009; Lwilla et al., 2003; Mohan et al., 2003; Newell et al., 2006; Wright et al., 2004; Zwarenstein et al., 2000). Five trials were included in the meta-analysis of outcome 6.1 Cure for smear positive TB patients (new and retreatment) (Clarke, 2005; Datiko & Lindtjorn, 2009; Lwilla et al., 2003; Mohan et al., 2003; Zwarenstein et al., 2000); four trials were included in the meta-analysis of outcome 6.2 Treatment completion rate for all pulmonary TB patients (vs. facility-based care) (CDI Study Group, 2010; Clarke, 2005; Datiko & Lindtjorn, 2009; and Zwarenstein et al., 2000); and two trials were included in the meta-analysis of outcome 6.3 Treatment completion rate for all pulmonary TB patients (vs. family DOTS) (Newell et al., 2006; Wright et al., 2004).

Setting: Four trials were conducted in upper-middle-income countries, one in a lower-middle-income country, and three in low-income countries. One trial each was conducted in Ethiopia (Datiko & Lindtjorn, 2009), Tanzania (Lwilla et al., 2003), Iraq (Mohan et al., 2003), Nepal (Newell et al., 2006), and Swaziland (Wright et al., 2004), and two were conducted in South Africa (Clarke, 2005; Zwarenstein et al., 2000). One trial (CDI Study Group, 2010) was conducted across two middle-income countries (Cameroon and Nigeria) and one low-income country (Uganda).

Participants: Five studies were conducted on new smear-positive TB adult patients (Clarke et al., 2005; Datiko & Lindtjorn, 2009; Lwilla et al., 2003; Mohan et al., 2003; Newell et al., 2006) and three were conducted on adults with pulmonary TB (CDI Study Group, 2010; Wright et al., 2004; Zwarenstein et al., 2000). CHW recruitment was via community input (CDI Study Group, 2010; Clarke, 2005; Datiko & Lindtjorn, 2009; Lwilla et al., 2003), the research team (Mohan et al., 2003), or was not specified. No other recruitment criteria beyond basic availability were specified.

Description of interventions: CHWs provided screening, referral and adherence motivation (Clarke, 2005; Mohan et al., 2003) or directly observed therapy short course (DOTS) (CDI Study Group, 2010; Datiko & Lindtjorn, 2009; Lwilla et al., 2003; Newell et al., 2006; Wright et al., 2004; Zwarenstein et al., 2000). One study (Datiko & Lindtjorn, 2009) reported the ratio of CHWs to population (1:5000).

CHW training was described in seven studies and ranged from one day (Newell et al., 2006) to thirty-five days (Clarke, 2005). No refresher training was reported.

Incentives were described in five studies. In two studies CHWs were provided a salary or incentivized payments (Clarke, 2005; Datiko & Lindtjorn, 2009), in one study they were reimbursed for expenses (Mohan et al., 2003), in one study incentives

were determined on a community-by-community basis (CDI Study Group, 2010), and in one study no incentives were provided (Newell et al., 2006).

Supervision was described in seven studies. It was provided variously by TB coordinators (Mohan et al., 2003), facility-based health workers and community members (CDI Study Group, 2010) or by a supervisor for which no characteristics were provided (Clarke, 2005; Datiko & Lindtjorn, 2009; Lwilla et al., 2003; Newell et al., 2006). The frequency with which it was undertaken ranged from fortnightly (Lwilla et al., 2003) to monthly (Clarke, 2005).

Comparison groups included usual care without the addition of a CHW (CDI Study Group, 2010; Clarke, 2005; Mohan et al., 2003; Zwarenstein et al., 2000), usual care where the CHW was not trained to deliver the intervention in question (Datiko & Lindtjorn, 2009), institution-based DOTS (Lwilla et al., 2003; Zwarenstein et al., 2000) or family member DOTS (Newell et al., 2006; Wright et al., 2004).

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison six. Cure for smear positive TB patients (new and retreatment) and treatment completion rate for all pulmonary TB patients (vs. facility-based care) are also included in Summary of Findings 6. Data reported at longest follow-up were included unless otherwise specified.

6.1 Cure for smear positive TB patients (new and retreatment)

Five trials were included in the meta-analysis (Fig 4.26). CHWs provided adherence motivation or DOTS to adult TB patients in Africa and the Middle East. The meta-analysis showed low quality evidence that CHW delivered interventions improve TB cure rates (RR 0.66, 95% CI 0.42 to 1.04, $P = 0.075$). Evidence was downgraded for imprecision and considerable statistical heterogeneity ($I^2 = 78.2\%$, $P = 0.001$). One study at unclear risk of bias, Mohan et al., (2003) appears to be an

outlier (possibly because it was conducted in a different region than the other four studies). When this study was excluded in a sensitivity analysis, the magnitude of effect dropped, but reduced variability in the observed effects led to a more precise, statistically significant pooled estimate (RR 0.85, 95% CI 0.72 to 0.99, $P = 0.043$, $I^2 = 0\%$, $P = 0.69$).

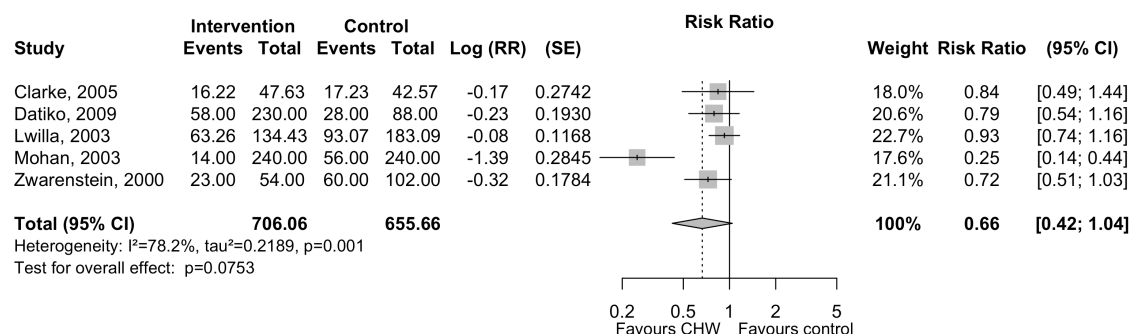


Figure 4.26 Comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, Outcome: 6.1 Cure for smear positive TB patients (new and retreatment) – adjusted for clustering

6.2 Treatment completion rate for all pulmonary TB patients (vs. facility-based care)

Four trials were included in the meta-analysis (Fig 4.27). CHWs provided adherence motivation or DOTS to adult TB patients across Africa. The meta-analysis showed high quality evidence that CHW delivered interventions can improve TB completion rates versus facility-based care (RR 0.63, 95% CI 0.45 to 0.88, $P = 0.006$). There was no evidence of heterogeneity ($I^2 = 0\%$, $P = 0.91$).

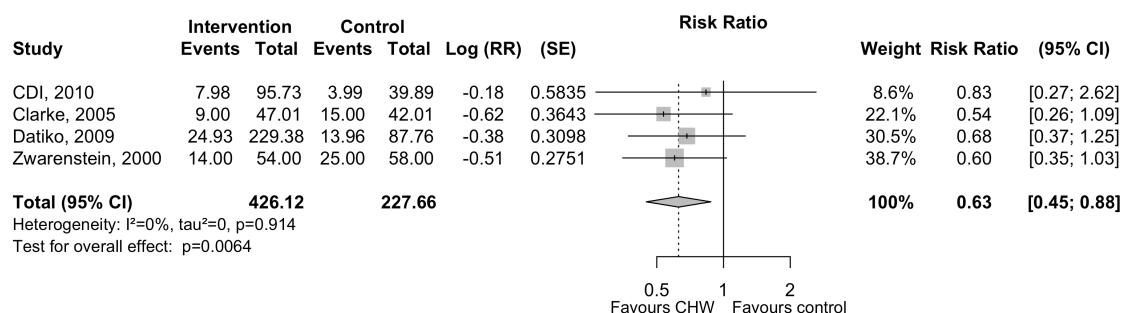


Figure 4.27 Forest plot of comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, outcome: 6.2 Treatment completion rate for all pulmonary TB patients (vs. facility-based care) – adjusted for clustering

6.3 Treatment completion rate for all pulmonary TB patients (vs. family DOTS)

Two trials were included in the meta-analysis (Fig 4.28). CHWs provided DOTS to adult TB patients in Nepal (Newell et al., 2006) and Swaziland (Wright et al., 2004). This showed moderate quality evidence that CHW DOTS is not superior to family member DOTS (RR 0.97, 95% CI 0.84 to 1.12, $P = 0.68$). Evidence was downgraded due to study design and implementation (allocation concealment in both trials was unclear or high risk). There was some evidence of heterogeneity ($I^2 = 37.2\%$, $P = 0.20$).

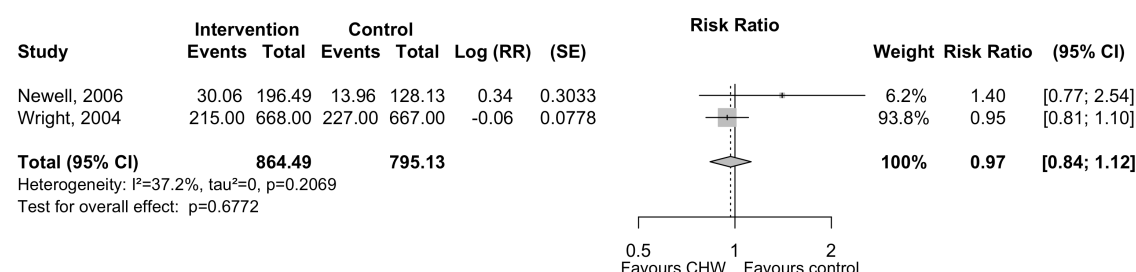


Figure 4.28 Forest plot of comparison 6: CHW delivered interventions to improve TB treatment outcomes compared with other forms of adherence support, outcome: 6.2 Treatment completion rate for all pulmonary TB patients (vs. family DOTS) – adjusted for clustering

Comparison 7: CHW delivered interventions for maintenance of antiretroviral therapy

Nine papers reporting five trials were examined in this comparison (Chang et al., 2013; Chang et al., 2010a; Chang et al., 2015; Jaffar et al., 2009; Lubega et al., 2015; Lugada et al., 2010; Nakigozi et al., 2015; Selke et al., 2010). Two trials were included in the meta-analysis of outcome 7.1. Death (Chang et al., 2010a; Jaffar et al., 2009); three trials were included in the meta-analysis of outcome 7.2 Loss to follow-up (Chang et al., 2010a; Jaffar et al., 2009; Selke et al., 2010); two trials reported on outcome 7.3 Retention in pre-ARV care (Chang et al., 2015; Lubega et al., 2015); and three trials reported on outcome 7.4 CD4 counts (Chang et al., 2010a; Jaffar et al., 2009; Selke et al., 2010),

Setting: Four trials were conducted in a low-income country, Uganda (Chang et al., 2010; Chang et al., 2015; Jaffar et al., 2009; Lubega et al., 2015), and one in a lower-middle-income country, Kenya (Selke et al., 2010).

Participants: All studies were conducted on HIV-infected adults. In Jaffar et al. (2009) adults were starting antiretroviral therapy for the first time; in Chang et al. (2010) they were either already on ART at the start of the trial or were started on ART at any time during the trial; and in Selke et al. (2010) they were clinically stable on ART for a minimum of three months with no adherence issues. In Chang et al. (2015) and Lubega et al. (2015) patients were newly screened and in pre-antiretroviral (pre-ARV) care.

CHW recruitment was via fellow patient nomination with final approval by clinic staff (Chang et al., 2015) or was not specified. In four trials CHW were chosen from the HIV clinic population (Chang et al., 2010; Chang et al., 2015; Lubega et al., 2015; Selke et al., 2010). Other recruitment criteria included good ART adherence for at least six months (Chang et al. 2010), literacy (Chang et al., 2010), and clinical stability (self-reported 100% ART adherence over the 6 months prior to recruitment; Selke et al., 2010). CHWs had degree qualifications or college diplomas (Jaffar et al., 2009), secondary education (Selke et al., 2010) or a minimum of primary education (Lubega et al., 2015).

Description of interventions: In two studies, CHWs provided monthly home counselling visits to provide psychosocial support and encourage patients to attend routine pre-ARV care (Chang et al., 2015; Lubega et al., 2015). In three studies, CHWs made home visits to monitor participants for symptoms of drug toxicity or disease progression, perform a pill count, and counsel on adherence (Chang et al., 2010; Jaffar et al., 2009; Selke et al., 2010). In Selke et al. (2010) they used Personal

Digital Assistants for decision support. The ratio of CHWs to population ranged from 1:10 (Lubega et al., 2015) to 1:28 (Chang et al., 2010).

CHW training was described in all five studies and ranged from two days (Chang et al., 2010; Chang et al., 2015) to four weeks (Jaffar et al., 2009). Initial instruction was followed by yearly refresher courses (Jaffar et al., 2009); two months of clinical mentoring (Selke et al., 2010); or regularly conducted field-based skills reinforcement trainings (Chang et al., 2015).

Supervision was described in four studies. It was provided by clinic staff (Chang et al., 2010; Jaffar et al., 2009; Lubega et al., 2015) or part-time coordinators (Chang et al., 2015) on a daily basis (Jaffar et al., 2009) or with a frequency that was not specified.

Incentives were described in four studies; in two studies CHWs were provided a regular salary and expenses (Chang et al., 2010; Chang et al., 2015), in two studies they received non-monetary (e.g. bicycle, t-shirts) incentives (Chang et al., 2010; Lubega et al., 2015), and in one study no incentives were provided (Jaffar et al., 2009).

In all five trials, comparison groups were clinic-based care.

Results: Findings are discussed for the main outcomes measured by included studies: death (12-24 months), loss to follow-up (12-24 months), CD4 count, and retention in pre-ARV care. The first two outcomes are featured in Summary of Findings 7. Data reported at longest follow-up were included.

7.1. Death (12-24 months)

Data from two studies, both conducted in Uganda, on the outcome were included in the meta-analysis (Fig 4.29). CHWs conducted home visits to provide ARV adherence support to HIV positive adults. The meta-analysis showed moderate

quality evidence of no difference in death at 12-24 months between home-based care provided by CHWs and facility-based care (RR 1.03, 95% CI 0.76 to 1.39; $P = 0.85$). Evidence was downgraded due to imprecision. Data were not heterogeneous ($I^2 = 0\%$, $P = 0.83$).

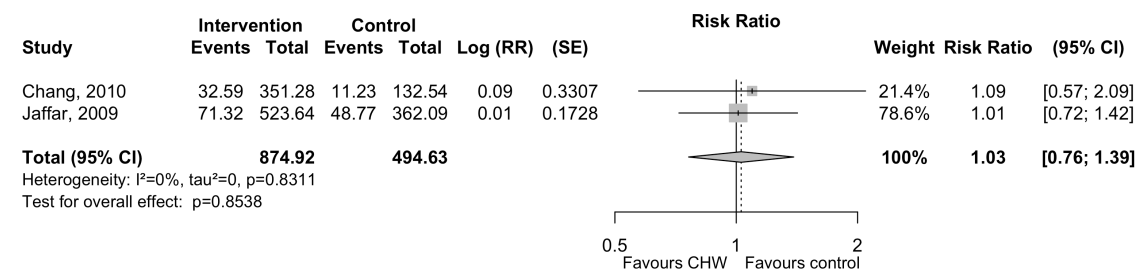


Figure 4.29 Forest plot of comparison 7: Usual care versus community health worker for maintenance of ART, outcome 7.1 Death (12-24 months) – adjusted for clustering

7.2 Loss to follow-up (12-24 months)

Data from three studies were included in the meta-analysis (Fig 4.30). CHWs conducted home visits to provide ARV adherence support to HIV positive adults in Uganda (Chang et al., 2010; Jaffar et al., 2009) and Kenya (Selke et al., 2010). The meta-analysis showed moderate quality evidence that loss to follow-up is similar in home-based care provided by CHWs and facility-based care (RR 0.82, 95% CI 0.46 to 1.54; $P = 0.58$). Evidence was downgraded due to imprecision. Data were not heterogeneous ($I^2 = 0\%$, $P = 0.58$).

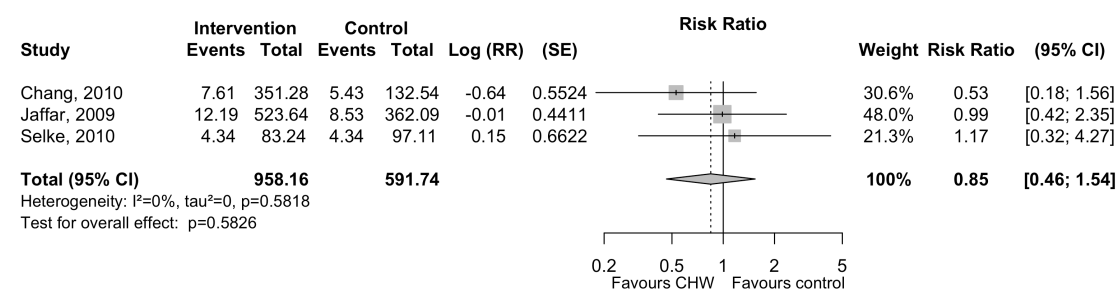


Figure 4.30 Forest plot of comparison 7: Usual care versus community health worker for maintenance of ART, outcome 7.2 Loss to follow-up (12-24 months) – adjusted for clustering

7.3 Retention in pre-ARV care

Two studies reported on retention in pre-ARV care; data were not available in a form that allowed statistical pooling. In Chang et al. (2015) intervention participants

were more likely to report being in care (92 vs. 84 %; RR 1.09, 95% CI 1.00 to 1.19) at one year. In Lubega et al. (2015) participants receiving monthly home counselling visits were 2.5 times more likely to be retained in pre-ARV compared with those in standard care over a period of 24 months (RR, 2.5, 95% CI 2.0 to 3.0). Overall, these studies suggest that CHW-led interventions can improve retention in pre-ARV care.

7.4 CD4 count

Three studies reported on CD4 counts; data were not available in a form that allowed statistical pooling. In Jaffar et al., (2009) 81% of participants in the home care arm and 80% of participants in facility care arm had CD4-cell counts of greater than 200 cells per μL at two years and counts at this visit were lower than baseline in 32 (4%) of those in home care compared with 29 (6%) in facility care. In Selke et al. (2010) intervention–control differences in median CD4 count at 12 months were not significantly different ($P = 0.5$). In Chang et al. (2010) CD4 change scores were similar at 2 year ($P = 0.49$). Overall, these studies suggest that CD4 count may be similar in home-based care provided by CHWs and facility-based care.

PUBLIC AND PREVENTATIVE HEALTH OUTCOMES

Comparison 8: CHW delivered interventions to promote injury prevention in children

Two trials were examined in this comparison (Odendaal et al., 2009; Swart et al., 2008); both are included in outcome 8.1 Hazard score—a standardised instrument composed of 88 items related to house-hold hazards associated with burn, poison and fall injuries (see: Jordaan, Atkins, Van Niekerk, & Seedat, 2005).

Setting: Both trials (Odendaal et al., 2009; Swart et al., 2008) were conducted in South Africa, an upper-middle income country.

Participants: Both studies were conducted on households with children 10 years or younger who lived in the house, visited there more than once a month, or stayed there during school holidays. Both studies employed paraprofessional home visitors. In Odendaal et al. (2009) they were recruited from a cohort of community health workers from the community employed by the collaborating NGO. No recruitment process was described in Swart et al. (2008).

Intervention and Comparator: CHWs in both trials conducted four home visits on child development, and the prevention of burn, poison, and fall injuries. CHWs conducted information dissemination about specific injury prevention practices; home inspection accompanied by information about home hazards; and the supply of safety devices. In Odendaal et al. (2009) visits were conducted over three months; in Swart et al. (2008) they were conducted over four months and lasted 40 minutes. The CHW to population ratio was 1:10 in Odendaal et al. (2009) and not specified in Swart et al. (2008).

Neither study specified the length of CHW training or the indicated whether refresher training was used. CHWs in both studies received training in unintentional injury prevention methods. CHWs in Swart et al. (2008) were supervised daily by two site coordinators; CHWs in Odendaal et al. (2009) were supervised weekly by unspecified actors. CHWs in Swart et al. (2008) were paid ZAR30 (approximately US\$4) per visit; in Odendaal et al. (2009) incentives were not described. In both trials, comparison groups did not receive visits until after the post-study assessment, at which point they were provided safety devices.

Results: Findings are discussed for the main outcome measured by included studies: hazard score. This outcome is featured in Summary of Findings 8. Data reported at longest follow-up were included.

8.1 Hazard score

Two trials were included in the meta-analysis (Fig 4.31). CHWs provided hazard prevention education to households with children 10 years or younger in South Africa. The meta-analysis showed very low quality evidence that CHW delivered interventions may reduce the presence of home hazards (SMD -1.8, 95% CI -5.02 to 1.42, $P = 0.27$). Evidence was downgraded for design and implementation (adequate sequence generation, allocation concealment, and blinding were unclear or high risk in both trials), imprecision, and substantial unexplained statistical heterogeneity ($I^2 = 79\%$, $P = 0.029$). The two trials were very similar in terms of population, intervention, and implementation; it is unclear why Odendaal et al. (2009) achieved a much larger effect.

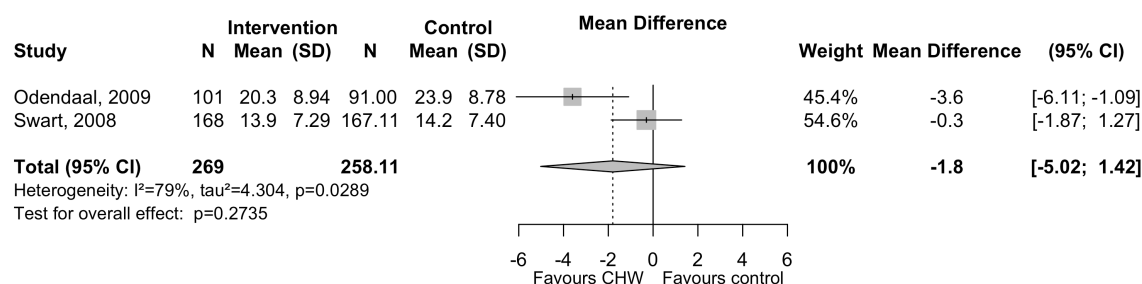


Figure 4.31 Forest plot of comparison 7: CHW delivered interventions to promote injury prevention in children, outcome: 7.1 hazard score – adjusted for clustering

Comparison 9: CHW delivered interventions to promote immunisation uptake and mass drug administration

Eight studies were examined in this comparison (Andersson et al., 2009; Bojang et al., 2011; Gökçay et al., 1993; Kweku et al., 2009; Ramaiah et al., 2001; Tine et al., 2011; UNICEF/UNDP/World Bank/WHO TDR, 2000; Wang et al., 2007). Three trials reported on outcome 9.1 Immunisation coverage (Andersson et al., 2009; Gökçay et al., 1993; Wang et al., 2007). Five trials reported on outcome 9.2 Mass drug administration coverage (Bojang et al., 2011; Kweku et al., 2009; Ramaiah et al., 2001; Tine et al., 2011; UNICEF/UNDP/World Bank/WHO TDR, 2000).

Setting: Four trials were conducted in upper middle-income countries, two trials in a lower middle-income country, and two trials in low-income countries. One trial each was conducted in Pakistan (Andersson et al., 2009), Turkey (Gökçay et al., 1993), China (Wang et al., 2007), The Gambia (Bojang et al., 2011), India (Ramaiah et al., 2001), Ghana and Kenya (UNICEF/UNDP/World Bank/WHO TDR, 2000), and Senegal (Tine et al., 2011)

Participants: Four studies were conducted on children under 5 years (Andersson et al., 2009; Gökçay et al., 1993; Kweku et al., 2009; Wang et al., 2007), two were conducted on children under 10 years (Bojang et al., 2011; Tine et al., 2011), and two were conducted on participants of any age save young children (Ramaiah et al., 2001; UNICEF/UNDP/World Bank/WHO TDR, 2000). CHW recruitment was via the research team (Andersson et al., 2009; UNICEF/UNDP/World Bank/WHO TDR, 2000), village leaders (Ramaiah et al., 2001) or by unspecified actors. Recruitment criteria ranged from the ability to read and write (UNICEF/UNDP/World Bank/WHO TDR, 2000) to primary school graduation (Gökçay et al., 1993). No other recruitment criteria were specified.

Description of interventions: In three studies, CHWs aimed to improve vaccination uptake. In Andersson et al. (2009) and Gökçay et al. (1993) CHWs offered information and assisted community members to access vaccinations; in Wang et al. (2007) CHWs administered vaccinations at home. In five studies CHWs administered mass treatment or preventative therapy. In Bojang et al. (2011), Kweku et al. (2009), and Tine et al. (2011) CHWs delivered Intermittent Preventive Treatment in children (IPTc); in Ramaiah et al. (2001) and UNICEF/UNDP/World Bank/WHO TDR (2000) CHWs delivered diethylcarbamazine. The ratio of CHWs to population ranged from 1:291 (Gökçay et al., 1993) to 1:>400 (Bojang et al., 2011).

CHW training was described in four studies and ranged from 2 days (UNICEF/UNDP/World Bank/WHO TDR, 2000) to 21 days (Gökçay et al., 1993). No refresher training was reported.

Incentives were described in three studies. In one study CHWs were provided a monthly salary (Gökçay et al., 1993) and in two study CHWs received one-off payments during drug administration (Bojang et al., 2011; Kweku et al., 2009). Supervision was not described in any study.

Comparison groups included usual, typically facility-based, care without the addition of a CHW (Andersson et al., 2009; Tine et al., 2011; UNICEF/UNDP/World Bank/WHO TDR, 2000; Wang et al., 2007), home visits by midwives or other health service personnel (Gökçay et al., 1993; Ramaiah et al., 2001), drug delivery by reproductive and child health (RCH) trekking teams (Bojang et al., 2011), or a mix of facility-based care and outreach clinics (Kweku et al., 2009).

Results: Findings are discussed for the main outcomes measured by included studies and listed under comparison nine. Mass treatment coverage is included in Summary of Findings 9. Data reported at longest follow-up were included unless otherwise specified.

9.1 Immunisation coverage

Three trials (Andersson et al., 2009; Gökçay et al., 1993; Wang et al., 2007) reported on outcome 9.1 Immunisation coverage. Because of the diversity of services provided by CHWs (e.g. education vs. vaccine provision) and comparators (midwife visits vs. no home visits) meta-analysis would not have been instructive; results are summarised narratively. In Andersson et al. (2009) measles and DPT vaccination uptake among children aged 12-23 months was significantly higher in intervention than in control clusters. Adjusting for baseline differences, the intervention doubled

the odds of measles vaccination in the intervention communities (OR 2.20, 95% CI 1.24 to 3.88) and trebled the odds of full DPT vaccination (OR 3.36, 95% CI 2.03 to 5.56). In Gökçay et al., (1993) no significant differences were found in immunisation status in children between areas visited by CHWs and areas visited by midwives. In Wang et al. (2007) timely (within 24 hours of birth) coverage of HepB vaccine was 72.5 % in the two intervention groups where CHWs administered the out of cold chain vaccine to infants at home (with or without a prefilled injection device) and 57.9% in the group where the vaccine was administered in township hospitals; a significant difference adjusting for clustering (RR 0.65, 95% CI 0.55 to 0.77, $P < 0.0001$).

Overall, these studies suggest that home visitation, whether by CHWs or other health professionals, to educate patients or deliver vaccines is likely to be useful in improving the coverage of a range of vaccines for children.

9.2 Mass drug administration coverage

Five trials reported on mass drug administration coverage, of which four were included in the meta-analysis (Fig 4.32). CHWs administered mass treatment or preventative therapy to patients in South Asia and across Africa. While clinical heterogeneity exists across trials, there was no discernible pattern in the varied effect sizes related to extracted characteristics (e.g. drug delivered, comparator type etc.). The meta-analysis showed low quality evidence that CHW delivered interventions can increase the coverage of mass treatment or preventative therapy (RR 0.56, 95% CI 0.31 to 1.00, $P = 0.048$). Evidence was downgraded due to study design and implementation (assessors were not blinded or it was unclear if they were blinded in all three studies) and the presence of considerable statistical heterogeneity ($I^2 = 96.2\%$, $P < 0.0001$).

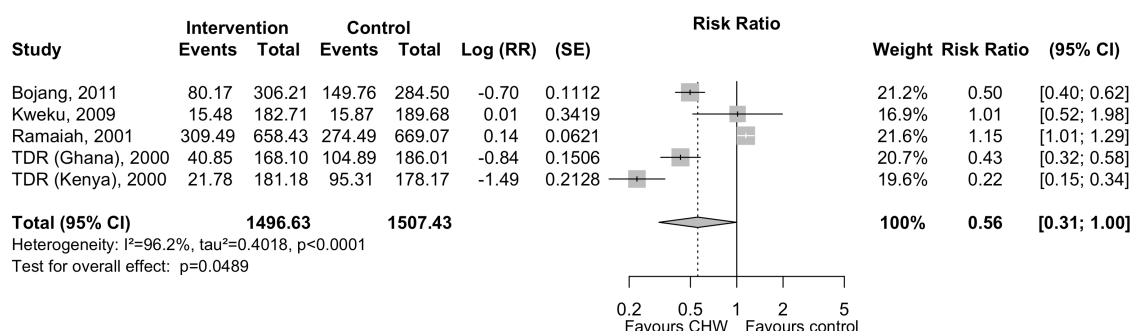


Figure 4.32 Forest plot of comparison 9: CHWs to promote immunisation uptake and mass drug administration, outcome: 9.2 Mass therapy coverage – adjusted for clustering

In Tine et al., (2014) seasonal malaria chemoprevention coverage was not reported for the control arm. In the intervention arm it was 97.3% (95% CI: 91.3 to 100%) in year one, and 88.8% (95% CI: 84.2 to 93.6%) in year two.

Comparison 10: CHW delivered interventions to reduce hypertension

Two trials were examined in this comparison (Jafar et al., 2009; Tian et al., 2015); both reported on outcome 10.1 Systolic blood pressure.

Setting: One trial was conducted in an upper-middle-income country (China) and lower-middle-income country (India) (Tian et al., 2015) while the second was conducted in Pakistan, a lower-middle-income country (Jafar et al., 2009).

Participants: Both trials were conducted on patients 40 years or older. In Jafar et al. (2009) patients had hypertension (systolic blood pressure ≥ 140 mm Hg, diastolic blood pressure ≥ 90 mm Hg, or already receiving treatment) whereas in Tian et al. (2015) patients were at high cardiovascular risk (self-reported history of coronary heart disease, stroke, diabetes mellitus, and/or measured systolic blood pressure ≥ 160 mmHg).

In Jafar et al. (2009) CHWs were recruited in line with the requirements of the government-sponsored Lady Health Workers Programme of Pakistan (i.e. 8 or preferably 10 years of schooling). In Tian et al. (2015) CHWs were pre-existing

village-based paraprofessionals in China and volunteer members of the community in India. No further details on the recruitment process was provided in either trial.

Description of comparator: Jafar et al. (2009) was a 2x2 factorial trial in which (i) CHWs provided health education every 3 months, (ii) general practitioners (GPs) received general training in hypertension management, (iii) both, or (iv) neither. CHWs in Tian et al. (2015) managed patients using an Android-powered app focused on the appropriate prescription of two medications (blood pressure-lowering agents and aspirin) and two lifestyle modifications (smoking cessation and salt reduction). The ratio of CHWs to patients in Jafar et al. (2009) was 1:110; in Tian et al. (2015) it was not specified.

Neither study specified the length of CHW training or indicated whether refresher training was used.

CHWs in Jafar et al. (2009) received training over six weeks in behaviour-changing communication strategies to convey standardised health education messages. In Tian et al. (2015) CHWs received one day of training and a refresher every 3 to 4 months during the implementation of the intervention.

CHWs in Jafar et al. (2009) were supervised quarterly by unspecified actors and provided a salary consistent with the requirements of the government-sponsored Lady Health Workers Programme of Pakistan. CHWs in Tian et al. (2015) were supported by a partnership with licensed physicians in healthcare centres.

Performance feedback and performance-based incentives to the CHWs were also part of the intervention package in China, whereas Indian CHWs received fixed monthly honorariums. On average, the payment for CHWs in the intervention group for the entire duration of the intervention was US\$400 in China and US\$500 in India. In both trials, comparison groups were usual care without the addition of a CHW.

Results: Findings are discussed for systolic blood pressure, the main outcome measured by included studies. Data reported at longest follow-up were included.

10.1 Systolic blood pressure

Two studies reported on systolic blood pressure; because of the diversity of populations (i.e. hypertensive adults vs. adults at risk for hypertension) and services provided by the CHWs (e.g. provision of drugs vs. health education) meta-analysis would not have been instructive, and instead results are summarised narratively. In Jafar et al. (2009) the age, sex, and baseline blood pressure–adjusted decrease in systolic blood pressure was significantly greater ($P < 0.001$) in the CHW and GP group (10.8 mm Hg, 95% CI 8.9 to 12.8 mm Hg) than in the GP-only (5.6 mm Hg, 95% CI 3.7 to 7.5 mm Hg), CHW-only (5.6 mm Hg, 95% CI 3.7 to 7.4 mm Hg), or no intervention groups (5.8 mm Hg, 95% CI, 3.9 to 7.7 mm Hg) after 2 years. In Tian et al. (2015) the intervention group had statistically significantly lower systolic blood pressure (−2.7 mm Hg; $P = 0.04$) than the intervention group after one year. Overall, these studies suggest that CHW-led interventions with appropriate clinical support can lower systolic blood pressure.

Summary of Findings 1

CHW delivered interventions versus usual care for treatment of common mental disorders

Patient or population: Patients with common mental disorders (includes anxiety disorders and depression)

Settings: Low-, lower-middle, and upper-middle-income countries – Uganda (4), Pakistan (2), India (1), Jamaica (1), Burundi (1), and South Africa (1)

Intervention: CHW led-psychological interventions

Comparison: Usual care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			

Depression prevalence Follow-up: 0.5-12 months	445 per 1000 ⁴	205 per 1000 (169 to 249)	RR 0.46 (0.38 to 0.56)	1244 (2 studies)	⊕⊕⊕□ Moderate ¹
Functional impairment Follow-up: 6-12 months		The amount of functional impairment was 0.44 standard deviations lower (0.09 to 0.80 lower) ⁵	SMD -0.44 (-0.80 to -0.09)	1392 (3 studies)	⊕⊕⊕□ Moderate ²
Severity of PTSD Follow-up: 0.5-6 months		The amount of functional impairment was 0.54 standard deviations lower (0.81 to 0.28 lower) ⁵	SMD -0.54 (-0.81 to -0.28)	390 (2 studies)	⊕⊕⊕□ Moderate ³

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. Downgraded for design and implementation as allocation concealment in Bolton 2003 was unclear.
2. There are considerable levels of heterogeneity across these studies ($I^2 = 85.2\%$, $P = 0.001$) and the confidence intervals do not overlap for all of the studies. The reasons for this heterogeneity are not clear.
3. Downgraded for design and implementation as sequence generation and allocation concealment in Neuner 2008 were high risk.
4. Control group risk is based on the mean baseline control group risk among included studies (Schünemann, Brożek, & Oxman, 2015).
5. A standard deviation of 0.4 represents a moderate effect.

Summary of Findings 2

CHW delivered interventions versus usual care to reduce mortality and morbidity among children under five years

Patient or population: Children under 5 years

Settings: Low-, lower-middle-, and upper-middle-income countries – Bangladesh (3), Pakistan (2), Ghana (1), Ethiopia (1), India (1), Malawi (1), Nepal (1), the Philippines (1), Tanzania (1), Thailand (1), Vietnam (1)

Intervention: CHW care

Comparison: Usual care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Child mortality Follow-up: 1-2 years	44 per 1000 ¹⁰	29 per 1000 (17 to 48)	RR 0.65 (0.38 to 1.10)	106378 (3 studies)	⊕⊕□□ Low ^{1,2}
Child morbidity Follow-up: 0-3 years	Low risk population ¹¹		RR 0.70 (0.51 to 0.96)	9453 (4 studies)	⊕⊕□□ Low ^{3,4}
	178 per 1000	125 per 1000 (91 to 171)			
	High risk population ¹¹				
	538 per 1000	377 per 1000 (274 to 517)			
Infant mortality Follow-up: 1-3 years	55 per 1000 ¹⁰	54 per 1000 (44 to 68)	RR 0.99 (0.80 to 1.23)	9459 (2 studies)	⊕⊕□□ Low ^{5,6,7}
Infant morbidity Follow-up: 1-3 years	Low risk population ¹¹		RR 0.73 (0.56 to 0.96)	48768 (7 studies)	⊕⊕□□ Low ^{8,9}
	286 per 1000	212 per 1000 (160 to 275)			
	High risk population ¹¹				
	425 per 1000	310 per 1000 (238 to 408)			

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. Downgraded for imprecision, as the pooled estimate of effect included both no effect and appreciable benefit.
2. Downgraded for design and implementation as Mtango 1986 and Pence 2005 only had 2 clusters each, causing issues with baseline differences.
3. Downgraded for design and implementation as it was unclear whether assessors were blinded in the majority of studies.
4. Downgraded, as there were moderate levels of heterogeneity across these studies ($I^2 = 44.8\%$, $P = 0.1$) and the reasons for this heterogeneity are not clear.
5. Downgraded for imprecision, as the pooled estimate of effect included both no effect and appreciable benefit.
6. Downgraded for design and implementation as Sloan 2007 was at high risk of performance and attrition bias

7. Refer to narrative summary of additional trials not contained in the forest plot.
8. There are considerable levels of heterogeneity across these studies ($I^2 = 90.9\%$, $P < 0.0001$) and the confidence intervals do not overlap for all of the studies. The reasons for this heterogeneity are not clear.
9. Downgraded for design and implementation as it was unclear whether assessors were blinded in the majority of studies.
10. Control group risk is based on the mean baseline control group risk among included studies (Schünemann et al., 2015).
11. Control group risks based on baseline risks found in the included studies, specifically the next to lowest and the next to highest (Schünemann et al., 2015).

Summary of Findings 3

CHW delivered interventions versus usual care for maternal and newborn health					
Patient or population: Women of reproductive age and their newborn children					
Settings: Low-, lower-middle, and upper-middle-income countries – Bangladesh (6), India (4), South Africa (3), Malawi (2), Uganda (2), Brazil (2), the Philippines (1), Ghana (1), Nepal (1), Mexico (1), Tanzania (1), and Burkina Faso (1)					
Intervention: CHW care					
Comparison: Usual care					
Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Neonatal mortality Follow-up: 1-3 years	Low risk population⁵		RR 0.76 (0.61 to 0.96)	157161 (15 studies)	⊕⊕⊕□ Moderate ¹
	11 per 1000	8 per 1000 (7 to 11)			
	High risk population⁵				
	84 per 1000	64 per 1000 (51 to 81)			
Perinatal mortality Follow-up: 1-3 years	55 per 1000 ⁶	46 per 1000 (40 to 52)	RR 0.83 (0.72 to 0.95)	72646 (7 studies)	⊕⊕⊕⊕ High
Maternal mortality Follow-up: 1-3 years	Low risk population⁵		RR 0.66 (0.35 to 1.25)	102775 (11 studies)	⊕⊕⊕□ Moderate ²
	1 per 1000	1 per 1000 (0 to 1)			
	High risk population⁵				
	5 per 1000	3 per 1000 (2 to 6)			
Exclusive breastfeed- ing Follow-up: 3-6 months	Low risk population⁵		RR 0.66 (0.52 to 0.83)	9382 (10 studies)	⊕⊕⊕□ Moderate ³
	847 per 1000	560 per 1000 (440 to 703)			
	High risk population⁵				
	975 per 1000	649 per 1000 (507 to 809)			
Initiation of	Low risk population⁵		RR 0.61 (0.45 to	75740 (11 studies)	⊕⊕⊕□ Moderate ⁴
	344 per 1000	210 per 1000			

breastfeed	(155 to 282)	0.82)
-ing	High risk population⁵	
Follow-up: 0.5-3 years	726 per 1000 443 per 1000 (327 to 595)	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. There are considerable levels of heterogeneity across these studies ($I^2 = 75.5\%$, $P < 0.0001$) and the confidence intervals do not overlap for all of the studies. The reasons for this heterogeneity are not clear.

2. Downgraded for imprecision, as the pooled estimate of effect included both no effect and appreciable benefit.

3. There are considerable levels of heterogeneity across these studies ($I^2 = 92.8\%$, $P < 0.0001$) and the confidence intervals do not overlap for all of the studies. The reasons for this heterogeneity are not clear.

4. There are considerable levels of heterogeneity across these studies ($I^2 = 95.8\%$, $P < 0.0001$) and the confidence intervals do not overlap for all of the studies. The reasons for this heterogeneity are not clear.

5. Control group risks based on baseline risks found in the included studies, specifically the next to lowest and the next to highest (Schünemann et al., 2015).

6. Control group risk is based on the mean baseline control group risk among included studies (Schünemann et al., 2015).

Summary of Findings 4

CHW delivered interventions versus usual care to promote cognitive development in infants

Patient or population: Infants 0-30 months

Settings: Lower-middle and upper-middle-income countries – Jamaica (4), Pakistan (1), India (1), and South Africa (1)

Intervention: CHW led-psychosocial stimulation

Comparison: Usual care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Infant cognitive development Follow-up: 12-24 months		The level of infant cognitive development was 0.56 standard deviations	SMD -0.56, (-0.86 to -0.27)	2676 (7 studies)	⊕⊕□□ Low ^{2,3}

higher
(0.86 to 0.27
higher)¹

* The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. A standard deviation of 0.6 represents a moderate effect.

2. Downgraded for design and implementation as adequate sequence generation and allocation concealment were unclear or high risk in five of the trials.

3. There are considerable levels of heterogeneity across these studies ($I^2 = 80.4\%$, $P < 0.0001$). The reasons for this heterogeneity are not clear.

Summary of Findings 5

CHW delivered interventions versus usual care to promote healthy feeding practices

Patient or population: Children 0-6 years

Settings: Lower-middle and upper-middle-income countries – Pakistan (2), Vietnam (1), and South Africa (1)

Intervention: CHW-led interventions

Comparison: Usual care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Weight for age: z-scores (WAZ) Follow-up: 6-12 months	Change in WAZ in the control group ranged from 1.98 to 2.16	The mean change in WAZ in the intervention group was 0.12 higher (0.21 to 0.02 higher)		2328 (3 studies)	⊕⊕⊕⊕ High ¹
Height-for-age: z-scores (HAZ) Follow-up: 6-12 months	Change in HAZ in the control group ranged from 1.36 to 2.34	The mean change in HAZ in the intervention group was 0.10 higher (0.29 higher to 0.09 lower)		2328 (3 studies)	⊕⊕□□ Low ^{2,3}

Weight-for-height: z-scores (WHZ) Follow-up: 6 months	Change in WHZ in the control group ranged from 1.06 to 1.39	The mean change in WHZ in the intervention group was 0.14 higher (0.27 to 0.01 higher)	1623 (2 studies)	⊕⊕⊕□ Moderate ⁴
---	---	---	---------------------	-------------------------------

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is the risk in the control group (Schünemann et al., 2015). The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. Refer to narrative summary of additional trials not contained in the forest plot.

2. There are considerable levels of heterogeneity across these studies ($I^2 = 68.8\%$, $P = 0.04$).

3. Downgraded for imprecision as the pooled estimate of effect included both no effect and appreciable benefit.

4. Downgraded for design and implementation as assessors were not blinded in Pachon 2002.

Summary of Findings 6

CHW delivered interventions versus other forms of adherence support to improve TB treatment outcomes

Patient or population: Adults with TB

Settings: Low-, lower-middle, and upper-middle-income countries – Ethiopia (1), Tanzania (1), Iraq (1), Nepal (1), Swaziland (1), South Africa (2), and Cameroon/Nigeria/Uganda (1)

Intervention: CHW-led adherence support

Comparison: Without CHW-led support

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Cure ² Follow-up: 5-7 months	Low risk population ⁴		RR 0.66 (0.42 to 1.04)	1565 (5 studies)	⊕⊕□□ Low ¹
	318 per 1000	210 per 1000 (134 to 331)			
	High risk population ⁴				
	508 per 1000	335 per 1000 (213 to 528)			
Completion	Low risk population ⁴		RR 0.63	655	⊕⊕⊕⊕

rate (vs. facility- based care) Follow-up: 5 to 36 months	159 per 1000	100 per 1000 (72 to 140)	(0.45 to 0.88)	(4 studies)	High
	High risk population⁴				
	357 per 1000	225 per 1000 (161 to 314)			
Completion rate (vs. family) Follow-up: 5 to 36 months	303 per 1000 ⁵	294 per 1000 (255 to 339)	RR 0.97 (0.84 to 1.12)	2242 (2 studies)	⊕⊕⊕□ Moderate ³

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. Downgraded for imprecision as the pooled estimate of effect included both no effect and appreciable benefit.

2. For smear positive, new & retreatment patients.

3. Downgraded for design and implementation as allocation concealment in both trials was unclear or high risk.

4. Control group risks based on baseline risks found in the included studies, specifically the next to lowest and the next to highest (Schünemann et al., 2015).

5. Control group risk is based on the mean baseline control group risk among included studies (Schünemann et al., 2015).

Summary of Findings 7

CHW delivered interventions versus clinic-based care for maintenance of antiretroviral therapy for HIV-infected patients

Patient or population: HIV-infected patients

Settings: Low- and lower-middle-income countries – Uganda (2), Kenya (1)

Intervention: CHW home visits

Comparison: Clinic-based care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Death Follow-up: 1-2 years	120 per 1000	124 per 1000 (91 to 167)	RR 1.03 (0.76 to 1.39)	2789 (2 studies)	⊕⊕⊕□ Moderate ¹

Loss to follow-up	31 per 1000	25 per 1000 (12 to 39)	RR 0.85 (0.46 to 1.54)	2997 (3 studies)	⊕⊕⊕⊖ Moderate ²
Follow-up: 1-2 years					

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is the mean baseline control group risk among included studies (Schünemann et al., 2015). The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. Downgraded for imprecision as the pooled estimate of effect included both no effect and appreciable benefit.

2. Downgraded for imprecision as the pooled estimate of effect included both no effect and appreciable benefit.

Summary of Findings 8

CHW delivered interventions versus usual care to promote injury prevention in children

Patient or population: Children 0-10 years

Settings: Upper-middle-income country – South Africa (2)

Intervention: CHW-led injury prevention intervention

Comparison: Usual care

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Hazard Score Follow-up: 1-2 weeks	The mean hazard score ranged across control groups from 14 to 24 points	The mean hazard score in the intervention groups was 1.8 points lower (5.02 points lower to 1.42 points higher)		569 (2 studies)	⊕□□□ Very low ^{1,2,3}

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is the risk in the control group (Schünemann et al., 2015). The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. There are considerable levels of heterogeneity across these studies ($I^2 = 79\%$, $P = 0.03$).

2. Downgraded for imprecision as the pooled estimate of effect included both no effect and appreciable benefit.

3. Downgraded for design and implementation as adequate sequence generation, allocation concealment, and blinding were unclear or high risk in both trials.

Summary of Findings 9

CHW delivered interventions versus usual care to promote immunisation uptake and mass drug administration

Patient or population: Entire community save young children

Settings: Low- and lower-middle-income countries – The Gambia (1), India (1), Ghana (1), and Kenya (1)

Intervention: CHW drug distribution

Comparison: Drug distribution by higher-level health workers

Outcomes	Illustrative comparative risk* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of evidence (GRADE)
	Assumed risk Usual care	Corresponding risk CHW			
Mass drug admin coverage Follow-up: Immediate	467 per 1000	262 per 1000 (145 to 467)	RR 0.56 (0.31 to 1.00)	11420 (4 studies)	⊕⊕□□ Low ^{1,2}

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is the mean baseline control group risk among included studies (Schünemann et al., 2015). The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1. There are considerable levels of heterogeneity across these studies ($I^2 = 97.1\%$, $P < 0.0001$).

2. Downgraded for study design and implementation as assessors were not blinded or it was unclear if they were blinded in all three studies.

Outcomes of trials not assigned to the above comparisons

Due to the range of interventions described and outcomes measured, six studies could not be assigned to a comparison (Aung et al., 2014; Dias et al., 2008; Fylkesnes et al., 2013; Ohnmar et al., 2012; Sankaranarayanan et al., 2005; Siabani et al., 2015). In these studies CHWs took part in a social franchising network to provide communities with ORS and zinc (Aung et al., 2014), provided interventions to alleviate caregiver burdens (Dias et al., 2008), provided home-based voluntary HIV counselling and testing (Fylkesnes et al., 2011), treated adults with malaria (Ohnmar et al., 2012), provided oral cancer screening (Sankaranarayanan et al., 2005) and provided education to patients with chronic heart failure (Siabani et al., 2015). These studies are fully described in the Characteristics of included studies tables and their primary outcomes are summarised in Appendix 4d.

Economic studies

Very limited data were available on unit cost and resource requirements: less than a quarter of included trials reported economic data. Of these papers, less than two thirds were full economic evaluations. Because the data on resource use and costs were very heterogeneous, meta-analysis was not appropriate and available data are presented in Table 4.2. These studies generally highlight the feasibility and potential cost effectiveness of CHWs in providing various tasks, particularly in comparison to facility-based care. The quality of the economic evaluations was assessed using the Cochrane-recommended CHEC list (Evers et al., 2005), a consensus criteria list that incorporates items from Drummond and Jefferson (1996) and other systematically-identified guidelines. These assessments are reported in a table in Appendix 4c. Because CHEC was designed to assess full economic evaluations, items not applicable to the seven included partial economic evaluations are shaded grey. The

majority of studies had significant risk of bias. Not all relevant alternatives and costs (e.g. training, supervision) were reported, one third of studies did not perform sensitivity analysis, nearly half did not properly discount, and no studies adequately addressed ethical and distributional implications.

Table 4.2 Economic analysis of included studies

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study population	CHW delivered intervention vs. Comparator	Economic results (Currency & price year)
Aung 2014 (Bishai 2015)	CEA/CUA (Societal, 1 year)	Children <5 years	Treat diarrhoea with ORS and zinc at home vs. government clinics	The incremental cost effectiveness is estimated at \$5,955 (IQR: \$3437-\$7589) per death averted and \$214 (IQR: \$127-\$287) per discounted DALY averted. (2010 USD)
Baqui 2008 (LeFevre 2013)	CEA/CUA (Societal, 3 years)	Ever married women 15–49 years	Home visits to promote birth and newborn-care preparedness & refer or treat sick neonates vs. facility care	The incremental costs of implementing the home-care package were \$2971 (95% CI: 1844 to 7628) per neonatal death averted and \$104.62 (95% CI: 65.15 to 266.60), per DALY averted. (2010 USD)
Bari 2011 (Sadrudin 2012)	Cost description (Household)	Children 2–59 months of age w/severe pneumonia	Home treatment of pneumonia vs. facility-based care	The average household cost/case for a CHW managed case was \$1.46 compared with \$7.60 for referred cases. (Year unspecified USD)
Bojang 2011	CEA (Provider, 4 months)	Children <6 years	Delivery of IPT vs. trekking teams	CHW strategy is dominant. ICER of saving per malaria episode averted: \$19.76. (2008 USD)
CDI Study Group 2010	Cost description (Provider)	Specific to intervention	Deliver vitamin A, bed nets, home mgmt of malaria and DOTS vs. usual care	The district-level costs of delivering all five interventions were lower in the CDI districts, but no cost difference was found at the first-line health facility level. (2005 USD)
Chang 2010 (Chang 2013)	CA (Provider)	HIV positive adults	Home based HIV adherence support vs. usual care	Excluding supervisory staffing costs, total yearly cost for the intervention was \$8,475, resulting in a yearly cost/patient of \$8.74. (2006-08 USD)

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study population	CHW delivered intervention vs. Comparator	Economic results (Currency & price year)
<i>Continued from previous page</i>				
Clarke 2005 (Clarke 2006)	CEA (Provider, 6 months/1 year)	New smear positive adult TB patients	Community based case finding and adherence support vs. usual care	The observed cost reduction to the Boland Health District was 74% per case detected and cured for the CHW int. relative to the control. (2004 USD)
Colbourn 2013 (Colbourn 2015)	CEA (Provider, 10 years)	Pregnant women	CHWs facilitate a learning action cycle with (FICI) or without (CI) facility improvements vs. facility improvements alone (FI)	The incremental cost-effectiveness of CI, FI, and combined FICI was \$79, \$281, and \$146 per DALY averted respectively, compared to current practice. FI is dominated by CI and FICI. (2013 international dollars)
Datiko 2009 (Datiko 2010)	CEA (Societal, N.S.)	TB suspects, who had cough for two weeks or more	Identify suspects, collect sputum, and provide DOTS vs. usual care	The cost per successfully treated patient was US\$161.9 and US\$60.7 by facility and CHW, respectively. (2007 USD)
Fottrell 2013	CEA (Provider, N.S.)	Ever-married women of reproductive age	CHWs facilitate a learning action cycle vs. health system strengthening	The cost-effectiveness \$10 053 per neonatal death averted and \$330 per year of life lost averted. The estimated cost-effectiveness of going straight to scale is \$6695 per neonatal death averted and \$220 per year of life lost averted. (2011 USD)
Jafar 2009 (Jafar 2011)	CEA (Societal, 2 years)	Patients 40 years or older with hypertension	Home health education vs. usual care	HE plus trained GP was the most cost-effective intervention, with an incremental cost-effectiveness ratio of \$23 (95% CI, 6 to 99) per mm Hg reduction in systolic BP compared with usual care. (2007 USD)
Jaffar 2009	Cost description (Societal)	HIV patients	Deliver drugs and provide adherence support vs. facility-based care	Cost per patient per year (mean) \$838 facility-based care and \$793 CHW care. (2008 USD)

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study population	CHW delivered intervention vs. Comparator	Economic results (Currency & price year)
<i>Continued from previous page</i>				
Kirkwood 2013 (Pitt 2016)	CEA (Public provider, 1 year)	All pregnancies that included still or live birth	Home visits to promote newborn-care practices & refer sick neonates vs. facility care	Newhints intervention cost a mean of \$10343 (95% CI 2963 to 7674) per newborn life saved, or \$352 (95% CI 104 to 268) per discounted life-year saved (2009 USD)
Kweku 2009 (Patouillard 2011)	CA (Provider)	Children aged 3–59 months	Delivery of IPT vs. facility-based delivery	The unit economic cost of receiving all 3 doses of all 4 courses of iPTc was \$7.56 and \$8.51 when IPTc was delivered by CHWs or facility-based nurses respectively. (2008 USD)
Lewycka 2013	CEA (Provider, 4.5 Years)	Women aged 10–49 years	Facilitate a learning action cycle vs. no intervention	The cost of the women's groups was \$114 per YLL averted; the average cost of volunteer peer counselling was \$33 per YLL averted. (2010 USD)
Manandhar 2004 (Borghi 2005)	CEA (Provider, N.S.)	Married women of reproductive age	Facilitate a learning action cycle vs. health system strengthening alone	The incremental cost per life-year saved (LYS) was \$211. Expansion could reduce the cost per LYS to \$138 (2003 USD)
Sankaranarayanan 2005 (Subramanian 2009)	CEA (Societal, 9 years)	Healthy participants aged 35 years and older	Vs. standard cars	The incremental cost per life-year saved was \$835 for all individuals eligible for screening and \$156 for high-risk individuals. (2004 USD)
Tylleskar 2011 (Chola 2011, Nkonki 2014)	Uganda: CA (Provider) South Africa: CA (Provider)	Mother-infant pairs	Home based exclusive breastfeeding counselling vs. standard care (Uganda) and unrelated CHW visits (South Africa)	Uganda: The cost per mother counselled and visit were \$139 & \$26, respectively (2005-'08 USD) South Africa: The average costs per woman and visit in were \$137 and \$32 per woman and visit, respectively. (2007 USD)

CA: Cost analysis; **CEA:** Cost effectiveness analysis; **DALY:** disability-adjusted life year; **ICER:** incremental cost effectiveness ratio; **IPT:** intermittent preventive treatment for malaria; **N.S.:** not stated; **USD:** US dollars; **YLL:** year of life lost

4.4 Discussion

4.4.1 Summary of Main Results

The review included 155 records reporting 86 randomised controlled trials evaluating the effects of CHW delivered interventions for primary health care outcomes. Studies were published between 1986 and 2016, with nearly 40% of studies published in the past five years. Twenty-three studies (27%) were conducted in nine upper-middle-income countries, forty-one studies (48%) were conducted in fourteen lower-middle-income countries, and twenty-three studies (27%) were conducted in seven low-income countries. In the majority of studies (77%), the intervention was delivered to patients in their home; otherwise it was delivered to patients in the community (18%) or involved a combination of home and community-based delivery (6%).

Relevant outcomes were extracted and grouped in ten comparisons according to the type of health issue addressed by the CHWs. Comparisons addressed mental health outcomes, maternal and paediatric health outcomes, infectious disease outcomes, and public and preventative health outcomes. The diversity of included studies limited meta-analysis to outcomes for nine of these ten comparison groups.

Mental health outcomes

For comparisons addressing mental health outcomes, included analyses showed moderate quality evidence that CHW delivered interventions can reduce the prevalence of depression, improve PTSD symptoms, and reduce functional impairment related to CMD. There was also low quality evidence that CHW-led interventions can lessen the severity of common mental disorders and very low quality evidence that CHW delivered interventions may have a small to moderate effect on reducing adolescent depressive symptoms. The latter three outcomes were

downgraded due to unexplained statistical heterogeneity. As noted in the results, while some clinical and methodological heterogeneity exists across trials, there were no discernible patterns in the varied effect sizes related to extracted characteristics (e.g. country, disorder etc.).

Maternal and paediatric health outcomes

For comparisons addressing maternal and paediatric health outcomes, the included analyses showed high quality evidence that CHW delivered interventions reduce perinatal mortality and improve WAZ. They showed moderate quality evidence that CHW delivered interventions reduce child morbidity, reduce neonatal mortality, improve WHZ, and increase exclusive breastfeeding and the speed of its initiation. They also showed moderate quality evidence of no difference in infant mortality, stillbirths, maternal mortality, and maternal morbidity between care provided by CHWs and usual, usually facility-based, care. There is low-quality evidence that CHW delivered interventions may reduce child mortality, increase the uptake of clean delivery kits, improve infant cognitive development, and produce no difference in HAZ. Finally, there is very low quality evidence that CHW delivered interventions have a large effect on maternal knowledge of infant development, may reduce infant morbidity, and may increase facility-based delivery.

The child morbidity, infant morbidity, NMR, maternal morbidity, facility based delivery, clean delivery kits, maternal knowledge of infant cognitive development, infant cognitive development, HAZ, and breastfeeding outcomes were downgraded for heterogeneity. While some clinical and methodological heterogeneity exists across trials in each comparison, there were largely no discernible patterns in the varied effect sizes related to extracted characteristics. Exploratory, post-hoc analyses did, however, indicate that (a) interventions in which CHWs provided

treatment had a larger effect on child morbidity than those in which CHWs only conducted health promotion and (b) CHW delivered interventions devoted solely to breastfeeding counselling tended to encourage more women to exclusively breastfeed than those devoted to breastfeeding counselling in addition to more general newborn preparedness. These results are observational and so could be confounded by other study-level characteristics and do not indicate that the subgroup factor explains heterogeneity. These factors were not identified *a priori* and a number of potential explanatory factors were considered, making these analyses a hypothesis-generating exercise.

Three outcomes, neonatal mortality, facility-based delivery, and clean delivery kits, were affected by an outlier. All outliers were of generally high quality and in no case was it clear from extracted characteristics what was driving the larger effect sizes. When the outlier for each was excluded the magnitude of the effect increased slightly for the latter dropped slightly for the former two outcomes.

Infectious disease outcomes

For comparisons assessing infectious disease outcomes, included analyses showed high quality evidence that CHW delivered interventions can improve TB completion rates versus facility-based care. They also showed moderate quality evidence that CHW DOTS is not superior to family member DOTS, and that facility-led maintenance of ARVs is no better than CHW-led maintenance in terms of loss to follow-up or death. There is low-quality evidence that CHW delivered interventions may improve TB cure rates versus facility-based care. Studies too diverse for meta-analysis also suggest that CHW-led interventions can improve retention in pre-ARV care, and that CD4 count may be similar in home-based care provided by CHWs and facility-based care.

In the analysis of the cure rate for smear positive TB patients, one study at unclear risk of bias appeared to be an outlier (possibly because it was conducted in a different region than the other studies). When this study was excluded in a sensitivity analysis, the magnitude of effect dropped, but reduced variability in the observed effects led to a more precise, statistically significant pooled estimate.

Public and preventative health outcomes

For comparisons assessing public and preventative health outcomes, the included analyses showed low quality evidence that CHW delivered interventions may reduce the presence of home hazards and increase the coverage of mass treatment or preventative therapy. Studies too diverse for meta-analysis suggest that home visitation, whether by CHWs or other health professionals, to educate or deliver vaccines is likely to be useful in improving the coverage of a range of vaccines for children and that CHW-led interventions with appropriate clinical support can lower systolic blood pressure.

The home hazard and mass treatment outcomes were downgraded for heterogeneity. While clinical heterogeneity exists across the trials reporting the latter outcome, there was no discernible pattern in the varied effect sizes related to extracted characteristics. Trials reporting the hazard outcome were very similar in terms of population, intervention, and implementation—the source of the statistical heterogeneity is likewise unclear.

Assessment of cost

There is mixed quality but consistent evidence that by making care more accessible and given reduced training and remuneration costs, CHWs could potentially reduce the direct and indirect costs of health care if substituted for professionals.

Interpreting findings across outcomes

It is worth re-emphasising that the results presented here are random-effects estimates, meaning they indicate the *average* intervention effect and so may not reflect the actual effect in any particular population under study. Poor reporting of CHW characteristics, CHW programme features, and regional or political context unfortunately made it difficult to explore heterogeneity stemming from such differences and, consequently, to make inferences regarding how to implement and optimise CHW programmes in different settings. That said, careful grouping of studies by PICOS characteristics has nonetheless allowed for insight into specific areas of CHW efficacy. Importantly, while the magnitude of the effect of CHW led-interventions varied, the results indicate that for no outcomes did CHWs cause harm or perform worse than usual care.

The internal and external validity of the results will be discussed further in the next two sections.

Questions that can and cannot be answered by this systematic review

The systematic review is a method of critically appraising, summarising, and attempting to reconcile the evidence about a particular issue (Higgins & Green, 2011; Petticrew & Roberts, 2006). There are, however, considerable technical and conceptual challenges in conducting public health systematic reviews (Petticrew & Roberts, 2008). As explained in Chapter 1, the evidence base for public health interventions has been described as under-populated, dispersed, and different in nature than that for clinical interventions (Millward, Kelly, & Nutbeam, 2003).

Under-populated because there are few outcome evaluations of public health interventions, *dispersed* because existing evidence is much more widely scattered than evidence of the effectiveness of clinical interventions, and *different* because

although randomised trials of complex social interventions exist, there is also a great deal of non-experimental evidence (Petticrew & Roberts, 2008).

There is an increasing range of positivist and interpretative synthesis methods that can help in reviewing complex public health interventions; in 2013, the *Journal of Clinical Epidemiology* presented an entire series on considering complexity in systematic reviews of interventions (Anderson et al., 2013a; Tugwell, Knottnerus, & Idzerda, 2013). Qualitative and mixed-method approaches include, for example, thematic synthesis (Thomas & Harden, 2008), framework synthesis (Dixon-Woods, 2011), and realist synthesis (Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013). Statistical approaches include the use of meta-analysis (Higgins & Green, 2011), meta-regression (Thompson & Higgins, 2002), and Bayesian methods (Pitchforth & Mengersen, 2012). One important dimension of difference between methods is whether a given approach aims predominantly to aggregate information using a pre-determined conceptual framework, or to interpret, configure and arrange theories and concepts (Gough, Oliver, & Thomas, 2012).

Aggregative, hypothesis-testing reviews of complex interventions, like the one in this chapter, have been critiqued for being potentially “simplistic and misleading” (Petticrew, 2015). Rather than evaluate interventions in terms of whether they ‘work’ or ‘do not work’, Petticrew (2015) suggests that reviews ought to answer the broader question of ‘What has happened previously when this intervention has been implemented across a range of contexts, populations and subpopulations, and how have those effects come about?’ This, he claims, could be used to better inform a decision about how an intervention is likely to behave in other settings (Petticrew, 2015).

In a similar vein, Ogilvie, Egan, Hamilton, and Petticrew (2005) argue that producing generalisable estimates of effect sizes is only one possible objective of evidence synthesis and that mapping the available evidence is also important. Using a systematic review of interventions promoting a population shift from cars toward walking and cycling as a case example, they demonstrated that population level interventions are sometimes less likely than individual level interventions to have been studied using randomised trials (Ogilvie et al., 2005). This supports concerns raised elsewhere that “unthinking adherence” to a hierarchy of study design as a means of selecting studies may reduce the value of evidence synthesis for public health and reinforce an “inverse evidence law” whereby the least is known about the effects of interventions most likely to influence whole populations (Ogilvie et al., 2005; Petticrew & Roberts, 2003; Rychetnik, Frommer, Hawe, & Shiell, 2002).

Although these and other aspects of complexity have been described in the literature, as Petticrew et al. (2013) have noted, they do not negate the usefulness of answering the question, “What is the average effect (or range of effects) of the complex intervention on the outcome(s) of interest?” While this approach leaves many potentially important questions unanswered (e.g. regarding context and implementation), it provides a clear answer to a valid question and provides valuable information to policymakers and other research users (Petticrew et al., 2013). The superior potential of RCTs to minimise bias when addressing questions of effectiveness is discussed in the next section.

By applying meta-analytical synthesis methods to trials, this review has addressed questions about: (1) whether an overall effect exists across a larger body of evidence than an individual study; (2) whether effects are consistent across studies; (3) what is the actual magnitude and variation of effects across studies; and (4)

whether particular study-level factors are associated with the magnitude of effect (Petticrew et al., 2013b). These questions, though clear and concise, are not “simplistic.” While they should not be the only questions asked by systematic reviews, answers to them make an important contribution to a broader inquiry into the chain of events that flow from the introduction of an intervention in a complex system (Anderson et al., 2013b).

As explained in Chapter 1, CHWs operate at the intersection of health promotion and individual patient care—some of their tasks fall under the umbrella of medicine while others are better classified as public health. Because this paper reviews randomised trials, evidence about interventions led by CHWs to influence the wider determinants of health—for which CHWs were originally thought to be superior (Berman et al., 1987)—may be under-represented. This is why it is important to understand that this review, though substantial, is but a starting point. In order to broaden the scope of the enquiry, it should be paired with complementary syntheses focused on the effects of implementation and process, context and setting, synergies or interactions between components, and individual intervention components (Anderson et al., 2013a). While addressing all of these questions is beyond the scope of this dissertation, the latter is the focus of Chapter 6.

4.4.2 Internal Validity

It is possible that a Hawthorne effect (improved outcomes due simply to the presence of a home visitor) was produced in studies where the comparator did not include home visits. A number of studies documented that “standard care” was poor, or non-existent; in such cases, the presence of any intervention is likely to be a motivational factor. Future studies would do well to adjust for the attention effect by, for instance, ensuring some type of regular contact between patients and data

collectors. That said, the Hawthorne effect cannot be overestimated. In Morrow et al. (1999), which found that CHWs can increase exclusive breastfeeding, the authors note that prior to their trial they undertook longitudinal studies that involved even more frequent home visits, yet exclusive breastfeeding remained uncommon. Moreover, it may be argued that the ability to give attention is a primary feature of CHW-led care and that it is thus appropriate to capture this effect.

While it has been argued that the contribution of observational studies tends to be unfairly downplayed in public health (Petticrew & Roberts, 2008) and objections to the use of randomised trials in evaluating public health and social interventions have been well chronicled (Oakley et al., 2003; Threlfall et al., 2015; Zachariah et al., 2009), the pragmatically-oriented trials in this review demonstrate the potential weakness of ecological evaluation and the value of randomisation in assessing intervention effects. As Bolton et al. (2007) and Bolton et al. (2003) point out in their trials of CHW-led group interpersonal psychotherapy, the severity of mental health symptoms can fluctuate over time. Due to regression to the mean, the CHW delivered interventions in both trials would have likely shown an effect regardless of true efficacy—without a control group, apparent improvement cannot be differentiated from true intervention efficacy. In their trial of CHW-led women's groups for perinatal outcomes, More et al. (2012) point out that this occurrence is not limited to trials of mental health. In their study, both the control and intervention groups improved over the course of the trial. If their assessment had been based on the favourable trends in the intervention arm, the intervention would have appeared an unqualified success, whereas the study actually found no statistically significant differences between trial arms in extended perinatal mortality or perinatal care practices.

That said, while randomised trials have a unique value in assessing intervention effects, the exclusion of all non-randomised studies without further consideration places a zero weighting on evidence from such studies. When methodologically rigorous observational studies show a large magnitude of effect, a dose-response gradient, and when all plausible confounders or biases would decrease an apparent treatment effect, they can yield high quality evidence (Guyatt et al., 2011). Given this, updates of this systematic review less constrained by the financial and temporal limitations of a DPhil should consider including such evidence.

4.4.3 Overall Completeness and Applicability (External Validity) of Evidence

The objective of this review was to determine the efficacy of CHW delivered interventions for community-based primary healthcare outcomes in low- and middle-income countries. The evidence gathered by this search is somewhat able to address this objective, but factors related to reporting and trial design limit the applicability of these results (Schünemann, Fretheim, & Oxman, 2006).

Reporting

Better reporting of CHW characteristics and programme features, control conditions, adverse events and health context is warranted. Several papers noted that certain CHW programme characteristics were important for programme success, yet did not report such features in sufficient detail for their effects to be explored across studies. For example, studies identified CHW selection (Agrasada et al., 2005; Haider et al., 2000), training (Agrasada et al., 2005; Haider et al., 2000; Rahman et al., 2008; Rotheram-Borus et al., 2014), supervision (Agrasada et al., 2005; Azad et al., 2010; Bari et al., 2011; le Roux et al., 2013; Rahman et al., 2008; Rotheram-Borus et al., 2014), and compensation (Bari et al., 2011) as “critical” (Agrasada et al., 2005) to the

“success” (Haider et al., 2000) and “sustainab[ility]” (Bari et al., 2001) of the intervention. Yet, none of these studies describe how these conclusions were reached. While a few included studies reported implementation information like dosage (Azad et al., 2010; Baker-Henningham et al., 2005; Baqui et al., 2008; Darmstadt et al., 2010; Kirkwood et al., 2013; Lewycka et al., 2013; Penfold et al., 2014) or frequency (Haider et al., 2000), the vast majority neglected to document the CHW characteristics (e.g. level of education) or programme features (e.g. type of supervision) in enough detail to permit the sub-group analyses necessary for this heterogeneity to be explored. Perhaps as a response to this uncertainty, several other recent studies called for more investigation on the design, delivery, and intensity of precisely these features (Bojang et al., 2011; Chang et al., 2010a; Engebretsen et al., 2014).

Similarly, studies included in this review compared CHW delivered interventions with multiple control groups, including similar services delivered by professionals (substitution) and “usual care,” which was often described poorly and in several cases appeared to be no care or care that was inaccessible due to geographic, financial, or other constraints. There were too few studies available for each comparison to explore these differences, but given that the vast majority of studies were not focused on substitution but rather extending care to places where it wasn’t available, caution should be employed when extrapolating these studies as evidence for CHWs substituting for professional providers.

Moreover, because only nine trials reported on the potential for adverse events (Agrasada et al., 2005; Bari et al., 2011; Bojang et al., 2011; Ertl et al., 2011; Fylkesnes et al., 2013; Jaffar et al., 2009; Kweku et al., 2009; Soofi et al., 2012; Stanton et al., 2013) and only three trials reported on the equity effects of their

interventions (Fylkesnes et al., 2013; Houweling et al., 2013; Kumar et al., 2012), it is difficult to draw conclusions regarding the trade-off between benefits and harms (Lavis, Oxman, Fretheim, & Lewin, 2009a; Welch et al., 2013). That said, many of the outcomes were adverse events (e.g. death) and CHW-led interventions, by nature of being community-based, tend to be targeted at the most vulnerable.

Intervention, comparator and outcomes notwithstanding, trials often also failed to report relevant information regarding context. This is true both of the cultural context in which CHWs operate and the way in which they fit into the larger health system. Information about, for example, the degree to which patients seek care from traditional healers as an alternative to or in addition to CHWs, whether male CHWs faced cultural barriers working with women in the home etc. are examples of norms and attitudes that might have shaped intervention effects and so would be relevant to interpreting trial outcomes or exploring differential effects across settings. Similarly, as with most complex interventions, it is likely that components of the CHW delivered intervention interact with contextual factors and each other (Craig et al., 2008; Montgomery et al., 2013b). For example, several trials conducted health facility strengthening in both arms alongside the CHW delivered intervention in one arm (Waiswa et al., 2015). The degree to which this continuity of care (Kerber et al., 2007; Liu, Sullivan, Khan, Sachs, & Singh, 2011) enhanced or enabled the concurrent CHW delivered intervention should be considered more carefully in future studies. Without better reporting of culture and context it is difficult to assess the applicability of these strategies elsewhere.

In order to aid in interpretation of context, both Lewin et al. (2010) and Lavis et al. (2009b) have delineated factors about a potential implementation setting that ought to be considered by decision-makers when assessing whether CHW

programmes are likely to be successful. These are adapted here for use when considering the applicability and transferability of the reviewed CHW-led interventions:

1. Whether there are important differences in health system arrangements that may alter the feasibility and acceptability of the given CHW delivered intervention. For example, it would be difficult to implement the oxytocin-delivery for postpartum haemorrhage prevention intervention described in Stanton et al. (2013) in a country where CHWs were prohibited from performing injections.
2. Whether there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting. For example, as Bojang et al. (2011) point out in their trial of CHW-led IPTc distribution, observed results were contingent on a routine supply of medicine and may not be replicable where health facility stock-out are common.
3. Whether there are important differences in the baseline health conditions between where the studies were done and the implementation setting. For example, Kumar et al. (2008) note that the effect of their CHW-led behaviour change intervention in regions with low neonatal mortality rates is not known, but effectiveness might be expected to diminish.
4. Whether studies from which evidence was drawn were conducted in settings with similar social norms. For example, male CHWs in Kumar et al. (2008) faced substantial barriers in accessing new mothers in the room of confinement on day 0; while these were successfully breached through

community engagement and acceptance, such change of practices may not be as easily achieved in other settings.

5. Whether sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time. For example, as noted above, several trials highlighted strong supervision as being important for programme success.

More complete information about CHW characteristics and programme features, control conditions, health context, and adverse events and equity effects would be of great relevance to policy and practice. Lack of information is, however, both a reporting and trial design issue—the latter is considered in the next section.

Trial Design

As explained in Chapter 1, thorough development and testing of interventions makes the final product easier to evaluate, implement, and understand. Medical Research Council guidance for developing and evaluating complex interventions indicates that unless time is taken in the early stages to iteratively incorporate evidence, theory, and pilot test data into an intervention, larger effectiveness trials are often fail to yield useful information about the processes and mechanisms might make their results generalisable (Campbell et al., 2000a; Campbell et al., 2007; Craig et al., 2008). Incorporating process evaluations would help explain the success, failure, or unexpected consequences of a trial; illuminate causal mechanisms of a successful intervention as well as avenues for optimisation; evaluate fidelity and quality of implementation; and reveal the degree to which the treatment can be standardised (Collins et al., 2011; Linnan & Steckler, 2002; Oakley et al., 2006). To the reviewers knowledge, only one of the trials included in the review conducted mediator analysis (le Roux et al., 2014) and three moderator analyses (Betancourt et al., 2012; Murray

et al., 2015; Rahman et al., 2012), all having to do with characteristic of participants not programme design or implementation. In order to improve the applicability of evidence, more needs to be done to understand the mechanisms underlying the success of CHW delivered interventions. In addition to this lack of process data, three issues with the trial design of studies in the systematic review limit the applicability of the data.

Firstly, over three quarters of included interventions were delivered in an efficacy setting where CHWs are more likely to have been carefully selected; managers more motivated; remuneration more available due to research funding; training and supervision more intensive; and recipient groups curated (Lewin et al., 2010; van Ginneken et al., 2013). For example, several trials in the sample excluded residents of precisely the type of hard-to-reach geographies that CHWs were designed to serve (World Health Organization, 1978); Jaffar et al. (2009) excluded those living on far-away islands, Kramer et al. (2014) those living in villages “considered by local leaders to be too difficult to reach in the rainy season” (p. 5323), and Ohnmar et al. (2012) those living in areas with security concerns or small populations. These exceptional inputs and exclusions, likely made the interventions easier to implement than they would be in non-experimental settings. Further research is required on how high quality, sustainable operationalisation of these interventions is feasible at scale in remote settings.

Secondly, the diverse set of interventions reported across studies meant there were few data for each outcome of interest, including cost. Exacerbating this diversity is sometimes poor consistency in outcome selection and definition. For example, in Comparison 1, CHW delivered interventions for maintenance of antiretroviral therapy, virological failure was defined slightly differently (>400 copies/mL in Chang

et al. 2010 and > 500 copies/mL in Jaffar et al. 2009), precluding meta-analysis. Moreover, with three exceptions (Gardner et al., 2003; Grantham-McGregor et al., 1991; Sankaranarayanan et al., 2005), most trials measured outcomes in the short-term after programme implementation—whether the effects are sustained in the long-term remains to be determined. That said, given that a longer duration of follow up may been needed to show the benefits of psychosocial interventions (Chatterjee et al., 2014) and behaviour change interventions (Colbourn et al., 2013), the short duration of follow-up of such trials may also have resulted in an under-estimation of effect.

Thirdly, less than a quarter of included trials reported economic data. Of these papers, less than two thirds were full economic evaluations. While several trials failed to include all relevant costs (e.g. training, supervision), four factors may have meant reported costs of evaluated interventions were overstated (Nkonki et al., 2014; Pitt et al., 2016). Firstly, most interventions were geographically limited and so could not benefit from potential economies of scale. Secondly, the CHWs in most trials delivered a single intervention. Such specialist CHWs, whether delivering a curative or preventative service, have been found to be expensive compared to those that cover multiple interventions (Makan, 1997). Diversifying CHW tasks, with due attention not to overload them (Jaskiewicz & Tulenko, 2012) could result in unit cost reductions. This would require trials to report what portion of a CHWs time is spent on the intervention and the possibilities of integration with other activities, which <5% did. Either way, it would not be appropriate to increase the workload without increasing remuneration and this must also be taken into consideration. Thirdly, most evaluated programmes had been designed for the purpose of the trial; Makan (1997) found that CHWs and managerial staff became more efficient over time thus it is possible that costs might fall and effectiveness could potentially increase as

management strategies were optimised and became more efficient over time—though their effects may attenuate over this timeframe as well. Fourthly, more than two thirds of the costing exercises were conducted from the provider’s perspective (i.e. only costs incurred by the provider were considered). Population level savings (e.g. relating to having a healthier workforce) that might be relevant to policymakers were not included. Given that cost of care is a major burden for people living in the rural areas CHWs typically serve (Jaffar et al., 2009) evaluations that take a broader societal perspective are warranted.

While very limited data were available on unit cost and resource requirements, these studies generally highlight the feasibility and potential cost effectiveness of CHWs in providing various tasks, particularly in comparison to facility-based care. That said, one third of studies did not perform sensitivity analysis, nearly half did not properly discount, and no studies adequately addressed ethical and distributional implications; as such, these data should be interpreted with caution.

4.4.4 Quality of the Evidence

In the GRADE system, well-conducted randomised controlled trials are initially categorised as high quality and observational studies are initially categorised as low quality. The quality of the evidence can subsequently be rated down based on five criteria (risk of bias, inconsistency of effect, indirectness, imprecision, publication bias) or rated up based on three criteria (large effect size, dose-response gradient, plausible confounding) (Guyatt et al., 2008b). The review includes 86 RCTs covering a wide range of interventions and settings. For studies included in the meta-analyses the evidence ranged from very low to high quality as assessed using GRADE; these assessments are explained in the Summary of Findings Tables included in the results section.

While many meta-analyses were performed and so, using the conventional significance level of 5%, it is expected that 1-2 of the tests will be statistically significant due to chance even when there is no true difference between the interventions being compared, each subgroup was likely an independent hypothesis (Higgins & Green, 2011). While the results were fairly consistent across meta-analyses in showing improvements in favour of CHW delivered interventions, for some outcomes there was important variation in the reported effects that could not be explained—these should be interpreted with caution.

These assessments are made with the caveat that—as had been noted elsewhere (Langendam, 2012; Rehfuss & Akl, 2013)—there are issues with the applicability, reproducibility, and clarity of GRADE in the field of public health. Where the results of a systematic review are highly heterogeneous and effect estimates rely on single studies or cannot be pooled, as in this review, applying GRADE criteria is challenging (Movsisyan, Melendez-Torres, & Montgomery, 2015). Even the selection of the appropriate body of evidence can be difficult to do systematically and transparently; in order to make reviews policy-relevant heterogeneous measures are, as in this case, sometimes grouped under multifaceted “umbrella outcomes” (Rehfuss & Akl, 2013, p. 6). Finally, GRADE does not integrate specifications that are particularly salient for assessing the effectiveness of complex interventions, for example, intervention delivery, economics, context and equity (Waters et al., 2011).

4.4.5 Potential Biases in the Review Process

As explained in the introduction, this systematic review was originally registered as an overview (PROSPERO 2015:CRD42015016638). While this change in methodology necessitated changes to the original protocol all amendments are

documented in the methods section and were agreed a priori. Nonetheless, the review has some limitations.

Firstly, CHW RCTs remain poorly indexed in electronic databases (Grayson & Gomersall, 2003; Lewin et al., 2010). As noted in Chapter 1, a large number of terms are used in the literature to describe CHWs and while nearly 70 synonyms for CHW were used in the electronic search for this review, several of the terms for CHWs used in the included studies (e.g. CBDOT observers, cluster facilitator etc.) were not among them. Similarly, though no language restrictions were imposed on the search and several non-English terms for CHW were included in the electronic search strategy, few non-English studies were found. It is therefore possible that, despite the extensive searches undertaken, some relevant RCTs were not identified.

Secondly, as has been noted elsewhere (Lehmann & Sanders, 2007; Lewin et al., 2010; Love et al., 1997; Witmer et al., 1995) there is no widely accepted definition for this cadre of health worker. Moreover, as many studies do not adequately describe the characteristics of included health workers, it is difficult to apply any definition. As such, exclusions may be debated.

Thirdly, despite an extensive search for unpublished and ongoing trials (which included searching grey literature, handsearching conference abstracts, and contacting experts), the studies located may be unrepresentative of all unpublished studies. Most of the meta-analyses included too few studies to allow for a meaningful assessment of asymmetry that might potentially suggest publication bias. That said, the included trials were conducted by a wide range of researchers and many reported results that were statistically non-significant, so there is no specific reason for suspecting a high risk of publication bias.

Fourthly, while randomised trials have a unique value in assessing intervention effects and the weaknesses of non-randomised studies have been described, the exclusion of all non-randomised studies without further consideration places a zero weighting on evidence from such studies. As discussed in section 4.4.1, this may mean that CHW delivered interventions aimed at the population level or social determinants of health could be under-represented (Petticrew & Roberts, 2008). Updates of this systematic review should consider including such evidence.

Finally, limitations of the search strategy aside, poor reporting and nonresponse meant the reviewer was unable to obtain relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some studies.

To ensure quality in conduct and reporting, the reviewer followed the Methodological Expectations of Cochrane Intervention Reviews (MECIR) and PRISMA statement for reporting systematic reviews (see Appendix 4e).

4.4.6 Agreements and Disagreements With Other Studies or Reviews

In an effort to determine what types of healthcare provision can be safely task shifted, a number of reviews have been carried out on CHW effectiveness. This systematic review, however, arose from the reality that an overview of these reviews would have been at risk of bias and would not have provided a comprehensive view of the evidence. Of 58 reviews related to CHW efficacy identified in Chapter 3, only 13 were of high quality (a score ≥ 8 on AMSTAR) (Arrowsmith et al., 2012; Bee et al., 2014; Bhutta et al., 2008b; Christopher et al., 2011; Kredo et al., 2014; Lassi & Bhutta, 2015; Lee et al., 2014; Lewin et al., 2010; Mangham-Jefferies et al., 2014; Okwundu et al., 2013; Sunguya et al., 2013; van Ginneken et al., 2013; Volmink & Garner, 2007). Of these, only five were specifically focused on community health workers or had results grouped by provider (Christopher et al., 2011; Kredo et al.,

2014; Lewin et al., 2010; Okwundu et al., 2013; Volmink & Garner, 2007); all save Kredo et al. (2014) and Okwundu et al., 2013 carried out searches more than seven years ago and are thus highly likely to be out of date. Because Kredo et al. (2014) looked only at initiation and maintenance of antiretroviral therapy and Okwundu et al., (2013) only at malaria, a broader review was warranted.

The outcomes examined in this review extend beyond those examined in previous reviews. That said, where overlap exists, the conclusions of this review largely align with those of the five older, high quality, CHW-focused reviews (Christopher et al., 2011; Kredo et al., 2014; Lewin et al., 2010; Okwundu et al., 2013; Volmink & Garner, 2007). As with Kredo et al. (2014), this review found moderate quality evidence that there is probably no difference in mortality or loss to follow-up when doctors deliver HIV care in the hospital or CHWs workers provide home-based care. In line with Christopher et al. (2011), Okwundu et al. (2013), and Lewin et al. (2010), this review found low quality evidence that CHWs reduce child mortality. In addition, the review echoes the Lewin et al. (2010) findings of moderate quality evidence that CHWs promote immunisation uptake, the initiation of breastfeeding, and exclusive breastfeeding and low quality evidence that CHWs reduce child morbidity. Additional studies conducted in the years since Lewin et al. (2010) was published led the current review to upgrade the body of evidence that CHWs reduce child morbidity from low to moderate.

While this review echoes the Volmink et al. (2007) conclusion that there is no significant difference between DOT provided by a family member and DOT provided by a community health worker; trials conducted since Lewin et al. (2010) and Volmink et al. (2007) carried out their searches indicate that CHW delivered interventions can improve TB completion rates versus facility-based care.

By incorporating more recent evidence on CHW efficacy for maternal, paediatric and infectious disease outcomes and reviewing the evidence on injury prevention, NCDs, healthy child feeding practices, child cognitive development, and mental health outcomes for the first time, this review confirms and extends the knowledge base developed by these earlier, high quality reviews.

4.5 Conclusion

Identifying and reviewing existing information is the first step in a phased approach to intervention development that has its conceptual roots in the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011) and the MRC Framework for Developing and Evaluating Complex Interventions (Craig, Dieppe, Macintyre, & Michie, 2008). While CHWs typically undertake multiple medical tasks (Paintain et al., 2014; Puett, Coates, Alderman, Sadruddin, & Sadler, 2012), first ascertaining whether and how they can do each safely and efficaciously is important for intervention design and optimisation. The systematic review identified 155 papers reporting 86 trials, of which nearly 40% were published in the past five years. As discussed above, the recent increase in rigorous testing of what tasks can be undertaken by CHWs yielded novel findings and reinforces the importance of a broad, up-to-date systematic review in this area. The implications for the design of these programmes and future research are discussed below.

4.5.1 Implications for Practice

While the heterogeneity of included studies reflects the enormous diversity of activities carried out by CHWs and the variety of contexts in which they function, careful grouping of studies by PICOS characteristics has allowed for insight into specific areas of CHW effectiveness. Moreover, the choice to conduct a broad review

has allowed for: (a) the creation of a “one-stop,” comprehensive overview of answering the overarching question of whether evidence exists that CHWs can perform various tasks. Nearly 40% of studies included in this review were published in the past five years; such a user-friendly front end is particularly useful for central Ministry of Health planning purposes and the uptake of relevant information by policymakers in an area where evidence has been changing rapidly and is scattered across many sources. (b) The identification of gaps in knowledge that are unlikely to have been uncovered by reviews focusing on CHW delivered interventions for a specific health issue. (c) The bringing together of insights about CHW characteristics and programme design that are likely to be applicable for interventions across health issues. (d) The identification of the most promising interventions for CHWs to deliver (Higgins & Green, 2011).

For comparisons addressing mental health outcomes, included analyses showed moderate quality evidence that CHW delivered interventions can reduce the prevalence of depression, improve PTSD symptoms, and reduce functional impairment related to CMD. There was also low quality evidence that CHW-led interventions can lessen the severity of common mental disorders and very low quality evidence that CHW delivered interventions may have a small to moderate effect on reducing adolescent depressive symptoms.

For comparisons addressing maternal and paediatric health outcomes, the included analyses showed high quality evidence that CHW delivered interventions reduce perinatal mortality and improve WAZ. They showed moderate quality evidence that CHW delivered interventions reduce child morbidity, reduce neonatal mortality, improve WHZ, and increase exclusive breastfeeding and the speed of its initiation. They also showed moderate quality evidence of no difference in infant

mortality, stillbirths, maternal mortality, and maternal morbidity between care provided by CHWs and usual, usually facility-based, care. There is low-quality evidence that CHW delivered interventions may reduce child mortality, increase the uptake of clean delivery kits, improve infant cognitive development, and produce no difference in HAZ. Finally, there is very low quality evidence that CHWs interventions have a large effect on maternal knowledge of infant development, may reduce infant morbidity, and may increase facility-based delivery.

For comparisons assessing infectious disease outcomes, included analyses showed high quality evidence that CHW delivered interventions can improve TB completion rates versus facility-based care. They also showed moderate quality evidence that CHW DOTS is not superior to family member DOTS, and that facility-led maintenance of ARVs is no better than CHW-led maintenance in terms of loss to follow-up or death. There is low-quality evidence that CHW delivered interventions may improve TB cure rates versus facility-based care. Studies too diverse for meta-analysis also suggest that CHW-led interventions can improve retention in pre-ARV care and that CD4 count may be similar in home-based care provided by CHWs and facility-based care.

For comparisons assessing public and preventative health outcomes, the included analyses showed low quality evidence that CHW delivered interventions may reduce the presence of home hazards and increase the coverage of mass treatment or preventative therapy. Studies too diverse for meta-analysis suggest that home visitation to educate or deliver vaccines, whether by CHWs or other health professionals, is likely to be useful in improving the coverage of a range of vaccines for children and that CHW-led interventions with appropriate clinical support can lower systolic blood pressure.

Poor reporting of CHW characteristics, CHW programme features, and regional or political context unfortunately made it difficult to explore heterogeneity stemming from such differences and, consequently, to make inferences regarding how to implement and optimise CHW programmes in different settings. That said, while the magnitude of the effect of CHW led-interventions varied, the results indicate that for no outcomes did CHWs cause harm or perform worse than usual care. Moreover, there is mixed quality but consistent evidence that, given reduced training and remuneration costs, CHWs could potentially reduce the direct and indirect costs of health care.

Given this, health policymakers should consider expanding the remit of existing CHWs to include these areas, while using rigorous designs and careful measurement to evaluate programmes and document implementation along the way. As discussed in section 4.4.1, CHW-led health delivery is a complex intervention, and this review does not answer all relevant questions pertaining to operationalisation of these programmes.

4.5.2 Implications for Research

For health issues in which evidence is currently of low quality or does not exist, new trials assessing the effectiveness of CHW delivered interventions are warranted.

For health issues where CHW delivered interventions demonstrate benefits, the focus of new research ought, in keeping with the Multiphase Optimisation Strategy (MOST) introduced in Chapter 1, to shift to optimising public health impact—i.e. understanding what CHW or programme characteristics enable or mediate effects (Collins et al., 2011). Such information on the relationship between outcomes and individual intervention components would help those designing and

implementing CHW programmes and ensure that resources are being used effectively (Craig et al., 2008; Glasziou, Meats, Heneghan, & Shepperd, 2008; Glenton, Lewin, & Scheel, 2011a; Shepperd et al., 2009). As has been noted elsewhere (Lewin et al., 2010) head-to-head comparison trials of the same intervention delivered by CHWs with different characteristics or types of support will contribute to disaggregating these components and better addressing this question. For this reason, existing trials of this nature were identified, critically appraised, and summarised in a systematic review of interventions to improve CHW performance, reported in Chapter 6. The results of Chapter 6 pertain directly to objective three of this thesis.

In new trials of either type, greater attention must be paid to study design. Trialists ought to specify a primary outcome and, unless testing independent mechanisms, adjust for the measurement of multiple outcomes. Adverse effects should be measured and the equity impacts of these interventions investigated. Such trials, as much as possible, should be conducted in programme settings and should, under no circumstances, exclude the hard-to-reach communities to which the CHW cadre was designed to extend services. As explained above, process evaluations and economic studies should also accompany such trials.

Likewise, costs and cost-effectiveness are important information for health policy planning; researchers should include these outcomes in their trials and facilitate cost-effectiveness meta-analysis by collecting and reporting cost-effectiveness data in a standardised format (e.g. costs per lives saved or disability-adjusted life years (DALYs) averted) (Lassi, Haider, & Bhutta, 2010).

Greater attention must also be paid to the quality of reporting (see: Montgomery et al., 2013a). Trialists ought to describe more thoroughly the intervention under consideration (e.g. design, delivery, uptake context) as well as

provide a complete description of the characteristics of CHWs deployed (e.g. recruitment process; duration, frequency, intensity, and assessment of training; type and frequency of supervision; type and amount of incentivisation etc.). While this review was able to calculate average effects of interventions, a better understanding of context, implementation, and optimisation was hindered by lack of clarity in intervention description. As has been noted by others (Lewin et al., 2010; van Ginneken et al., 2013), there is a need to develop a comprehensive taxonomy of CHW and programme design characteristics in order to enable better reporting. The development of such a taxonomy is the subject of Chapter 5. The results of Chapter 5 pertain directly to objective two of this thesis.

As explained in the discussion, randomised trials have a unique value in assessing intervention effects. By answering the question “can the intervention work?” a comprehensive systematic review of RCTs is an essential first step in furnishing policymakers with the necessary information for decision-making. Without it, there is little point in asking questions of sustainability, transferability, financing, and optimisation. That said, as has been pointed out elsewhere these are nonetheless policy-relevant questions that RCTs do not answer (Walt, 2005). As such, further systematic reviews drawing on a range of study designs (e.g. process evaluations, ethnographic work) to answer such questions would complement this work.

Chapter 5: Toward a Taxonomy of Community Health Worker Characteristics and Programme Features

5.1 Introduction

In line with aim two of this thesis, the objectives of this chapter are to (a) generate a formal community health worker taxonomy that comprises a comprehensive classification of distinct (i.e. non-overlapping and non-redundant), precisely defined properties that can be used with ease and confidence to describe CHW characteristics and programme design features and (b) to use this taxonomy as the basis to recommend a parsimonious set of characteristics for intervention description reporting in future evaluation research.

5.1.1 Conceptualising Community Health Workers

While enormous intellectual energy has been spent testing tasks to be shifted (Callaghan et al., 2010; Fulton et al., 2011; Joshi et al., 2014; Kredo et al., 2014; Mdege et al., 2013), less energy has been expended on issues related to systematically disaggregating the types of personnel involved. Consider the CHW definitions provided in two studies on providing primary care to newborns and children:

“illiterate... householders...[trained] to recognise symptoms in their under-5 children [and give treatment]... in each [area] one [CHW] was chosen to coordinate all other [CHWs and] collected the monthly reports” (Kidane & Morrow, 2000, p. 552).

“recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders... received a combination of classroom-based and apprenticeship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change management, and trust-building... after training, suitable candidates were closely mentored and supervised by a regional programme supervisor (n=4) responsible for 6–7 [CHWs], for an additional week before final selection was made... Performance assessment of [CHWs] included feedback from community members, spot checks by their supervisors during home visits and community meetings to assess their level of community engagement, and monitoring by the supervisors of whether targets for home visits and community meetings were being met” (Kumar et al., 2008, p. 1154).

The two definitions differ markedly in their content and detail. In the first study, the CHWs are illiterate residents recruited in an unspecified manner provided with training of an unspecified length with minimal supervision by their peers. In the second study, CHWs are high school graduates selected through a competitive process based on performance in a weeklong classroom and field based training module, supervised in a 1:6-7 ratio by a superior who performed audits and on-the-job training.

There is a distinct lack of standardised language to describe CHWs. As noted in Chapter 1, two recent systematic reviews of CHW delivered interventions (Bhutta et al., 2010; Gilmore & McAuliffe, 2013) identified nearly 70 unique terms (ranging from “barefoot doctor” to “lady health worker”) used worldwide for community health workers (see Appendix 1a for a listing). Even with suitable search strategies, this variation hinders identification of relevant research, hampering synthesis and knowledge transfer (Glasziou et al., 2008).

This failure to identify research and disaggregate types of CHWs may mean that particular characteristics (literacy, CHW remuneration) that distinguish between interventions remain unidentified. If such unreported or “unseen” differences are

“active ingredients” associated with effectiveness, the design of optimally effective interventions (Collins et al., 2011) and the uptake and replication of such interventions will be impeded. The first systematic review included in this thesis (Chapter 4) found 86 trials of CHW delivered interventions; while such evaluations are important for providing evidence of programmes’ efficacy, they must be accompanied by analyses that identify the conditions necessary for effective and sustainable programme implementation, particularly given existing qualitative documentation suggesting that CHW programmes have not always been successful (Kadiyala, 2004). Unfortunately, reporting across this body of evidence was not sufficiently detailed or consistent to allow these analyses.

There has been a recent uptick in the use of taxonomies to organise other areas of medical knowledge such as social media use in chronic disease research (Patel, Chang, Greysen, & Chopra, 2015), medicine use by consumers (Lowe, Ryan, Santesso, & Hill, 2011), determinants of healthcare professional practice (Flottorp et al., 2013), and patient complaints in healthcare systems (Reader, Gillespie, & Roberts, 2014). The need for a similar classification system for CHWs is well recognised; two highly cited Cochrane reviews of CHW delivered interventions called for a “comprehensive” taxonomy (van Ginneken et al., 2013, p. 49) that “might allow [CHW] interventions to be conceptualised in terms of their lay component rather than in terms of the specific health issues [tasks] on which they focus” (Lewin et al., 2010, p. 42).

The roles and features of nursing were specified in a taxonomy almost 25 years ago (Iowa Intervention Project, 1992). While researchers and health systems continue to experiment with what roles CHWs can perform (van Ginneken et al., 2013) the

time has come to apply the same conceptual rigour in defining, measuring and specifying the precise features of health workers ascribed the term “CHW.”

As Berman et al. (1987) note in their seminal paper on CHWs, “CHW programmes represent a mode for the organisation of services rather than a type of intervention” (p. 445). While descriptions of tasks carried out by CHWs exist, in part, elsewhere (Crigler, Hill, Furth, & Bjerregaard, 2011; Glenton et al., 2011b) and the range of interventions they carry out continues to be tested and expanded (see Chapter 4), this study will bring much needed clarity and precision to the delivery agent itself. This paper describes the development of a taxonomy to describe CHW characteristics and programme features with the aim of improving the design and reporting of clinical trials and component selection experiments, synthesis of evidence from these trials, and replication of efficacious interventions.

5.1.2 Reporting Guidelines and Intervention Description

The challenge of reporting complex interventions has long been recognised. In 2000, the Medical Research Council (MRC) published a framework for reporting and evaluating complex interventions (Medical Research Council, 2000). Though the MRC framework stresses the importance of complete and accurate reporting, it provides little operational guidance.

Reporting guidelines published thereafter, such as Consolidated Standards of Reporting Trials (CONSORT) (Moher, Schulz, Altman, & Group, 2001; Schulz, Altman, & Moher, 2010), have significantly improved the completeness of reporting in randomised controlled trials (Turner et al., 2012). Their impact, however, has been differentially realised. While the number of studies specifying certain items—for example, allocation concealment, sample size, scientific rationale and background,

method used for sequence generation, and aggregate score over reported CONSORT items—has increased, the reporting of intervention descriptions has not been similarly affected (Kok et al., 2015a).

Two official CONSORT extensions relevant to the area of health promotion have since been published: one extending the CONSORT statement to randomised trials of nonpharmacologic treatments (Boutron, Moher, Altman, Schulz, & Ravaud, 2008) and one extending item five of the CONSORT statement: the Template for Intervention Description and Replication (TIDieR) (Hoffmann et al., 2014). An additional reporting guideline extending CONSORT to social and psychological interventions is now being developed (Montgomery et al., 2013a).

The first extension calls for the reporting of eligibility for care providers and centres and “a careful description” of care providers “(case volume, qualification, expertise, etc.)” and centres (volume) in each group (p. 296, 298). The checklist points out that a nonpharmacologic treatment might be found to be safe and effective in a trial performed in high-volume centres by high-volume care providers, but could have different results in low-volume centres (Halm, Lee, & Chassin, 2002). TIDier echoes this guidance, calling for the “expertise, background, and any specific training given” to be reported for each category of intervention provider (Hoffmann et al., 2014, p. 10).

Given the number of variables along which CHWs can vary (e.g. paid/unpaid, literate/illiterate, pre-existing/recruited for the study), however, following this new guidance may still frequently yield descriptions where salient features go unreported. Moreover, CHW delivered interventions are typically provided outside a health facility (Perry et al., 2014; WHO Study Group, 1989). As such, the emphasis on volume of care—a shorthand for expertise and support—must be translated to the

community settings where CHWs provide care (e.g. by detailing supervision, the ratio of CHWs to patients etc.).

A similar phenomenon is evident in reporting standards for other types of trial designs. The Transparent Reporting of Evaluations with Non-randomised Designs (TREND) guidelines call for even less information about the delivery agent, asking only “who delivered the intervention?” and “how was the content given?” (Des Jarlais, Lyles, & Crepaz, 2004, p. 363).

The consequences of this are far-reaching. Without sufficient detail, Ministries of Health may be unable to replicate effective interventions. The absence of standardised definitions for community health worker characteristics and programme features means that systematic reviewers develop their own systems for classifying providers and synthesising study findings (e.g. Lewin et al., 2010; van Ginneken et al., 2013) creating confusion and the risk of effort duplication. All of this points to the urgent need for consensus. While existing guidelines have greatly improved the completeness and transparency of task reporting, more must be done to ensure completeness and consistency with regards to adequately describing the delivery agent, namely CHW characteristics and programme features.

While it would be possible to include tasks in this taxonomy, effectiveness studies have often failed to separate the health impact of a specific task (e.g. a vaccine), from the delivery strategy used to reach the target population (Victora, 2004). As such, the author felt it was important to restrict the focus of the taxonomy to CHW characteristics and programme features in order to more clearly separate the *task* (e.g. malaria management) from the *delivery agent* (e.g. CHWs). Moreover, as explained above, existing reporting guidelines already do a fine job of delineating the

salient features of tasks, but have thus far failed to do so adequately for this highly heterogeneous class of delivery agent.

5.1.3 An Introduction to Taxonomies

The word taxonomy comes from the Greek *taxis*, meaning arrangement or order, and *nomos*, meaning law or usage (Jones, Liddell, & Scott, 1996). In contemporary information management, taxonomy is understood both in a narrow sense, as a hierarchical categorisation system, and in a broad sense, as the “rules or conventions of order or arrangement” (Hedden, 2010; Lambe, 2014, p. 4). Lambe (2014) lists three attributes of a taxonomy: (i) a knowledge map that displays a domain in a particular form (e.g. tree, list), (ii) a classification system that groups together related things, and (iii) a semantic representation made from a controlled vocabulary.

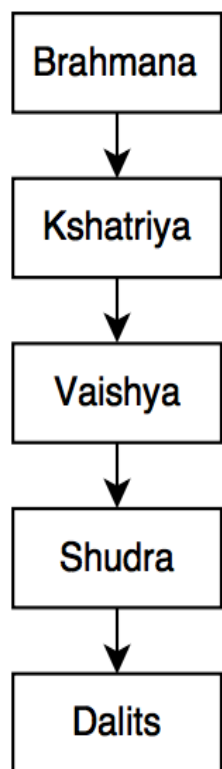


Figure 5.1 Indian caste system

While the narrower definition of a taxonomy as a hierarchical structure has its origins in the biological sciences (Linnaeus, 1758), definitions in several fields now commonly ascribe to taxonomies the quality of being “hierarchical” (Rosenfeld & Morville, 2002). Yet, defining taxonomies as hierarchies without further defining what is meant by that term can be confusing as the concept implies different things in different disciplines (Lambe, 2014). For example, in pure sciences, hierarchies are defined as having the characteristics of *inclusiveness* (the top category “includes” all subordinate categories), *relational consistency* (the kind of relationship between each level in a hierarchy is the same), *inheritance* (subordinate categories in a

hierarchy inherit all of the attributes of superordinate categories), and *mutual exclusivity* (an entity can belong in only one place in the hierarchy) (Kwasnik, 2000).

Yet in humanities or social sciences, hierarchies typically signify nothing more than a distribution of power. For example, the Indian caste system is social hierarchy in which those in higher strata have more power and privilege (Fig 5.1). Yet a scientific understanding of such a hierarchy, applying the concepts of inclusiveness and inheritance, would mistakenly suggest that Dalits are “part of” Brahmana or share some of their characteristics (Lambe, 2014).

Moreover, to suggest that taxonomies are inherently hierarchical ignores the many ways taxonomies have been used in the natural, medical and social sciences. Stavri and Michie (2012) have identified six different classification systems across different domains: (i) nomenclatures, (ii) ordered sets, (iii) hierarchical classification, (iv) matrix classification, (v) faceted classification, and (vi) social categorisation systems. Each classification system is described in turn; their characteristics are summarised in Table 5.1.

(i) Lists and nomenclatures

In his book *The Checklist Manifesto*, Gawande (2009) chronicled the important efficiency, consistency, and safety outcomes achieved in the medical field through the use of checklists. The power and impact of such checklists indicates that they are an important mechanism for information classification and knowledge transfer, despite their simplicity. Lambe (2014, p. 14) identifies lists—defined as a “collection of related things”— as the most basic form of taxonomy.

Stavri and Michie (2012), however, begin with a nomenclature: an “unstructured system of terms,” in which classification depends on the precise meaning of the terms (p. 114). They provide an example from astronomy: in 2006, Pluto was de-classified as a planet because it did not meet the third criterion of the International Astronomical Union’s newly adopted definition for such celestial bodies

(International Astronomical Union, 2008). As Lambe (2014), notes, nomenclatures can also be defined as “a list of all the parts that make up a whole” and are sometimes depicted in “system maps,” visual representations of a knowledge domain where “proximity and connections between entities” express their taxonomic relationships to real-world analogues (e.g. an artery diagram or conceptual map) (p. 42-43).

(ii) Ordered sets

An ordered set involves placing classified objects in a logical order so as to communicate information about the relationships between the objects (Stavri & Michie, 2012). For example, Bloom’s Taxonomy (Bloom, 1956) of educational objectives is divided into three domains—cognitive, affective and psychomotor—that contain levels ascending from the lowest level of processing (e.g. knowledge, comprehension) to the highest (e.g. synthesis, evaluation) (Fig 5.2). Though statistical tests of the models have not found conclusive support for the nature and order of the levels of processing—path analysis suggesting, for example, that the order of ‘analysis’ and ‘synthesis’ may be reversed—Bloom’s original intention reflects the characteristics of an ordered set (Krathwohl, 1994; Kreitzer & Madaus, 1994).

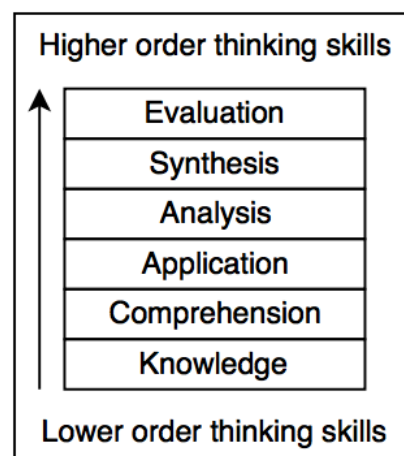


Figure 5.2 Bloom's Taxonomy (1956)

(iii) Hierarchical classification

While such ordered lists are valuable, when classifying a large number of objects it is desirable to group items together; such data reduction has been shown to improve processing of large quantities of information (Miller, 1956). Hierarchies are organisations of objects into “groups based on similarity, using superordinate and subordinate categories” (Stavri & Michie, 2012, p. 116). Members of each taxon at each level are on the average more related to each other than they are to members of other taxa at corresponding levels (Sokal, 1974). When delineating the features of a hierarchy, Stavri and Michie (2012) reference the same scientific criteria as Lambe (2014): inclusiveness, relational consistency, inheritance, and mutual exclusivity (Kwasnik, 2000); this is the definition that will be employed in this paper. In other words, while hierarchies can be presented in a list, they differ from ordered lists in that they are cumulative structures and often have several branches (Stavri & Michie, 2012). While tree structures represent “the transition from general specific or from whole to part” and are not necessarily hierarchies, hierarchies are typically tree structures (Lambe, 2014, p. 14). The most famous hierarchical classification was developed by the Swedish botanist Carolus Linnaeus in the 18th century to standardise the naming system for plants and animals (Linnaeus, 1758; McKercher, 2016). It used seven tiers—Kingdom, Phylum, Class, Order, Family, Genus and Species—with each level reflecting a greater degree of specificity. Commonly cited (Lambe, 2014; Stavri & Michie, 2012) examples of hierarchical classifications from other disciplines include the Dewey Decimal System for library classification (Broughton, 2004), The World Health Organisation’s International Classification of Diseases (ICD-10) (World Health Organisation, 2008), and corporate organisational charts—though the latter do not always meet all four criteria (Lambe, 2014). For such

cases, Lambe (2014) suggests the term “polyhierarchy”—hierarchies in which items that belong in more than one class or can be categorised in multiple ways have their relationship mapped to more than one superordinate concept. While polyhierarchies are useful where there is more than one organising principle at work, Rosenfeld and Morville (2002) note that they are problematic for taxonomy representation. If too many hierarchies are cross-connected, the structure—i.e. the very thing hierarchies are meant to convey and reinforce—becomes harder to discern.

(iv) Matrix classification

A more useful way of dealing with such complexity is by using matrices and facets. As opposed to trees, which subdivide groups on the basis of only one dimension at a time, matrices allow for the categorisation of content along multiple dimensions (Lambe, 2014). A well-known example of matrix taxonomy is the Periodic Table, which uses atomic weight and chemical similarities to place chemical elements into a grid (Stavri & Michie, 2012). This arrangement helped explain behaviours (e.g. the ‘periodicity’⁴ of the elements) and gave predictive power (e.g. Mendeleev was able to accurately predict the properties of scandium, germanium and gallium before their discovery) (Bensaude-Vincent & Stengers, 1996; Lambe, 2014). Above three dimensions, however, matrices are no longer visually easy to comprehend or navigate and it becomes difficult to represent taxonomies in this way.

(v) Faceted classification

In such situations, objects might better be classified according to a number of unrelated dimensions or “facets.” Lambe (2014) defines a facet as “a base taxonomy” comprising “only one of the fundamental dimensions” along which content can be analysed (p. 37). Though Indian librarian S.R. Ranganathan invented facets in the

⁴ The observation that elements exhibit similar properties at regular intervals of atomic mass increase.

1930s (Ranganathan, 1967), due to their multiple dimensions they are not well suited to locating books within physical collections (as opposed to alternative, one-dimensional enumerative schemes like the Dewey Decimal System - DDS). As result, facets were not widely used until the Internet age (Lambe, 2014; Stavri & Michie, 2012). Facets are, for example, uniquely well positioned for use with metadata, defined as “the collection of structured information about a document or piece of content” (e.g. title, author, keywords, or price, size, brand) (Lambe, 2014, p. 37). Each piece of metadata can be represented as a separate facet, and each made searchable. Facets overcome the constraints of a single classification scheme and, because they are post-coordinated (as opposed to pre-coordinated like DDS) they can expand as new fields of knowledge emerge (Lambe, 2014). Though expressive, facets do not map well.

(vi) Social categorisation systems

Finally, Stavri and Michie (2012) have identified social categorisation systems (i.e. user-tagging), in which participants organise items (e.g. tweets, images) using tags. While such tags have been found to conform to official guidelines for constructing controlled vocabularies—the tags are typically nouns and tend to avoid jargon—they do not aim for objectivity as by nature tags are an expression of personal opinion (Peterson, 2006; Spiteri, 2007). Such “folksonomies” have a ‘flat structure and do not describe relationships between categories (Lambe, 2014; Mathes, 2004).

Table 5.1 Characteristics of classification systems from the natural, medical and social sciences (Adapted from Stavri & Michie, 2012)

	List or Nomenclature	Ordered set	Hierarchy	Matrix	Facet	Social
Example	Planets	Bloom's taxonomy	Dewey Decimal System	Periodic Table	Product metadata	Tagging photos
Class. schemes	One	One	One	Two- three	Multiple	Multiple
Hierarchical	No	No	Yes	No	Multiple hierarchies possible	No
Hospitality*	Poor: requires changing definitions	Poor: sets are fixed	Good: within the existing classifica- tion scheme	Good: within classif- cation schemes	Excellent: post- coordinated, new classifica- tion schemes can be added	Excellent: post- coordinated, new classificat- ion schemes can be added
Enables prediction	No	Yes	Yes	Yes	No	No

* The degree to which the system is able to change or expand to accommodate new subjects

Additional considerations

Despite deriving from different epistemological and methodological bases, the terms taxonomy and typology are often erroneously used interchangeably (McKercher, 2016). While taxonomic systems are empirically based and classify items using observable and measurable characteristics, typologies are theoretically driven classification schemes based on ideal types that may accentuate characteristics not found empirically (McKercher 2016, Lamb et al., 2011). This paper aims to develop a taxonomy.

It is also worth distinguishing between coding taxonomies and intervention taxonomies, as each type has been optimised for different uses (Kok et al., 2015a). In a coding taxonomy, items are mutually exclusive so as to enable reliable classification, whereas in a development taxonomy, items may overlap and even encompass each other as the goal is to enable planners to most efficiently select the

features that fit their circumstances (Kok et al., 2015a). Taxonomies developed primarily for the purpose of categorising intervention content are not a good basis for intervention development because, for the sake of completeness, they will contain characteristics associated with ineffective, or even counter-effective interventions. The converse is also true—because the categories in development taxonomies are not mutually exclusive they are not well suited to coding. The taxonomy in this chapter is intended for use in coding. It was conceived to aid in component research, not, as in other recent papers (Galbraith et al., 2011), to catalogue core elements presumed to be responsible for intervention efficacy.

There are at least four potential benefits of developing a taxonomy of CHW characteristics:

1. First, a taxonomy will define the full range of CHW characteristics available in order to improve sector-wide awareness, understanding, and communication (Catterall, 1970; Lowe et al., 2011).
2. Second, controlled vocabulary of a taxonomy will facilitate conceptual mapping of different types of CHWs. It will clarify the key features of interventions to aid in evaluation and allow for improved characterisation and consistent reporting of interventions in trials (Lamb et al., 2011; Michie et al., 2013; Willis et al., 2013).
3. Thirdly, a taxonomy will enable systematic reviews to use a reliable method for extracting information about intervention content. This will aid in the identification of areas where research is strong and where gaps exist, the reliable synthesis of evidence, and the exploration of discrete, replicable characteristics (or combinations of characteristics) as potential mechanisms of action associated with intervention effectiveness (Michie et al., 2013).

4. Fourthly, this enhanced classification and characterisation will allow for the accurate replication of interventions. This enables the investigation of generalisability across behaviours, populations, and settings and for the operationalisation of the practice standards that permit the faithful propagation of interventions found to be efficacious or effective (Abraham & Michie, 2008; Michie et al., 2013).

The methods by which this process was undertaken are described in the next section.

5.2 Methods

5.2.1 Theoretical Underpinnings

Taxonomic development can take place either inductively or deductively (Elo & Kyngäs, 2008). The inductive process is atheoretical and begins with examination of individual cases of the phenomenon of interest in order to distinguish the criteria to be used to establish the taxa (DeJong, Horn, Gassaway, Slavin, & Dijkers, 2004). A deductive approach derives a taxonomy from an existing theory or theories by beginning with the a priori establishment of taxa criteria and proceeding to grouping and labelling (DeJong, Horn, Gassaway, Slavin, & Dijkers, 2004). Because it was desired that the classification system map easily onto real-world programmes, an empirical, inductive approach to developing a taxonomy of CHW characteristics and programme features was adopted.

No specific theory or single approach grounds classification methods (Lamb et al., 2011). Instead, taxonomists typically take a varied approach that is ‘fit for purpose’ (Bradley, Curry, & Devers, 2007; McKercher, 2016)—employing, at times, a systematic review or overview (e.g. Patel et al., 2015), consensus methods (e.g. Lamb et al., 2011), the study of representative cases or influential studies (e.g. Chun, Golomb, & Turk-Browne, 2011), or some combination of these methods (Flottorp et

al., 2013). Several recent taxonomies in the medical field have been constructed solely on the basis of existing literature (e.g. Lowe et al., 2011; Patel et al., 2015; Reader et al., 2014); in order to ensure a comprehensive mapping of existing CHW types, this same systematic approach is employed in this paper. While subjecting the resulting literature-based taxonomy to a consensus process would likely be helpful in further refining the taxonomy and possibly increasing uptake, resource constraints made it not possible to undertake at this time. A consensus process will be noted as a potential future project; until then, the systematic methods employed in this paper allow the resulting taxonomy to stand on its own merit without this additional step.

As described earlier, the ultimate form taken by the taxonomy has also been debated: the hierarchic, mutually exclusive, monothetic classification ideal has been questioned and the potential for overlapping clusters of polythetic taxa explored (Sokal, 1974). Despite this methodological pluralism, a number of criteria have been proposed to ensure the soundness and utility of a generated taxonomy (DeJong et al., 2004; Flottorp et al., 2013; Lamb et al., 2011; Michie et al., 2013). Building on these suggestions and observations, the taxonomy described in this chapter will be iteratively considered against the following criteria:

1. The taxonomy is jointly exhaustive (i.e. all examples of phenomena should be capable of being assigned to one of the taxa) yet parsimonious.
2. It is feasible to report the required information in published studies.
3. The included characteristics can be observed and tested empirically.
4. The included characteristics are conceptually unique (i.e. not redundant or overlapping).
5. The included characteristics are defined in a way that is accurate and easily understood.

6. The included characteristics are grouped in a clear and meaningful manner.
7. Future CHW characteristics and programme features can be included in the taxonomy.

5.2.2 Data Sources

The first step in building the taxonomy involved developing as comprehensive a list of CHW characteristics and programme features as possible using:

1. Existing classification schemes
2. Primary randomised controlled trials of CHW delivered interventions and connected process evaluations and/or qualitative studies of CHW experience
3. Influential studies (>100 citations) in the CHW literature

As noted in Chapter 4, randomised trials do not typically describe CHW characteristics and programme features in great detail. The reason for their inclusion alongside the more intuitive sources used for the taxonomy (i.e. existing classification schemes, process evaluations, qualitative studies, and influential studies in the CHW literature) is largely opportunistic: much of the coded data was already available from the two systematic reviews conducted in this dissertation and its inclusion posed no obvious harm.

Existing classification schemes

Existing classification schemes were identified in two ways. First, Medline (1946 to February Week 3 2016), AMED (1985 to February 2016), Global Health (1973 to Week 07 2016), EMBASE (1974 to 2016 February 29), and PSYCHINFO (1967 to February Week 4 2016) were searched on the Ovid platform and POPLINE on K4Health for articles providing a CHW classification scheme with the following search strings:

OVID databases:

1. Community Health Workers/
2. ((community or village) adj2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)).ti.
3. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) adj2 (worker)).ti.
4. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "colaborador voluntario" or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or "drug-kit manager" or anganwadi* or "barefoot doctor" or "distributeurs communautaires" or "relais communautaires" or "agentes polivalentes elementares").ti.
5. ((saude or salud) adj2 (agente or voluntario or tecnicos or promotores)).tw.
6. (health* or medical*) adj3 (auxiliar* or attendant* or village or volunteer)).ti.
7. 1 or 2 or 3 or 4 or 5 or 6
8. (matrix or taxonomy or synthesis or typology or model* or framework or tool or realist or design or interpret* or performance or mechanism*).ti.
9. 7 and 8

Popline:

((((Keyword:COMMUNITY WORKERS OR Keyword:BAREFOOT DOCTORS OR Keyword:UNTRAINED PERSONNEL)) AND ((Title:matrix OR Title:taxonomy OR Title:synthesis OR Title:typology OR Title:model OR Title:framework OR Title:tool OR Title:realist OR Title:design OR Title:interpret OR Title:performance OR Title:mechanism))))

Second, the following grey literature sources were searched: CDC, CHW Central, CORE group, Frontline Health Workers Coalition, Global Health Workforce Alliance, One Million Community Health Workers, and USAID, and WHO. In both cases, documents containing CHW classification schemes were selected.

Primary trials and connected process evaluations or qualitative studies

The primary trials were identified using comprehensive searches in the context of two systematic reviews: (a) CHW delivered interventions for primary health care outcomes and (b) interventions to improve CHW performance. The search methods are detailed in Chapter 4 and Chapter 6, respectively. The titles of all primary trials included in these reviews were individually entered in to Scopus and citing articles screened for process evaluations and/or connected qualitative studies providing further details on trial community health workers. For registered trials, this process was repeated using the trial registration number.

Influential studies in the CHW literature

Because not all CHW delivered interventions have been evaluated in trials we also searched for influential studies in the CHW literature to ensure the characteristics of CHWs that have not been evaluated but who are working in the field were captured in the taxonomy. The limits of citation analysis have been well documented; for example, a high number of citations may indicate a controversial, rather than authoritative work and citation indexes may not cover citations books or non-English language journals (Cleary & Hunt, 2010; Diodato & Gellatly, 2013). Moreover, several indicators for assessing the impact of a given publication exist, each with their own advantages and drawbacks (Cleary & Hunt, 2010). This paper has followed guidance from the Leiden Manifesto for research metrics (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015) by using a simple indicator (citation count), a transparent process (triangulation from three leading citation indices), and using quantitative evaluation to support qualitative assessment (screening). The three leading sources of bibliometric data (Cleary & Hunt, 2010)—Web of Science, Scopus, and Google Scholar (using Publish or Perish software, Harzing, 2007)—were searched using the following search strings:

Scopus:

TITLE (((community OR village) PRE/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*)) OR ((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR extension OR community) PRE/2 (worker)) OR (accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR monitora* OR promotora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "colaborador voluntario" OR feldsher OR "shastho shebika" OR "shastho karmis" OR sevika OR asha OR "drug-kit manager" OR anganwadi* OR "barefoot doctor" OR "distributeurs communautaires" OR "relais communautaires" OR "agentes polivalentes elementares") OR ((saude OR salud) PRE/2 (agente OR voluntario OR tecnicos OR promotores)) OR ((health OR medical) W/3 (auxiliar* OR attendant* OR village OR volunteer)))

Web of Science:

(TI=((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*)) OR TI=((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or community) near/2 (worker* or volunteer*)) OR TI=(accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or colaborador voluntario or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or drug-kit manager or anganwadi* or "barefoot doctor" or distributeurs communautaires or relais communautaires or agentes polivalentes elementares) OR TI=((saude or salud) NEAR/2 (agente or voluntario or tecnicos or promotores)) OR TI=((health or medical*) NEAR/3 (auxiliar* or attendant* or village or volunteer)))

Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH
Timespan=All years

Google Scholar:

Four separate queries were run through Publish or Perish (Harzing, 2007): (1) “community health worker,” (2) “community health workers,” (3) “lay health worker,” (4) “lay health workers.”

Given that Google Scholar tracks citations in sources that Scopus and Web of Science do not (e.g. books, book chapters, dissertations, theses, working papers, reports, conference papers); better aggregates minor variations of the same title; and tends to have better coverage of non-English sources, it is a valuable complement to the two more traditional databases (Harzing, 2013). Studies listed as having more than 100 citations in any database were screened. If CHWs in LMICs were the primary focus of the article, they were included.

5.2.3 Data Selection and Extraction

The definition of “CHW characteristic or programme feature” used in this study is any quality inherent to the CHW or programme design attribute related to the intervention’s delivery agent (as opposed to the task itself or the context in which it takes place). Such characteristics or programme feature must meet the following criteria adapted from Michie et al. (2013): (a) could be an “active ingredient” of the

larger intervention, (b) is the smallest component compatible with retaining the proposed active ingredient, (c) is observable and replicable, (d) may or may not have an established empirical evidence base.

When searching for existing information classification systems, any framework that incorporated CHW characteristics or intervention design factors was included, even if that is not the main focus of the framework.

For primary trials, connected studies, and influential papers, any listed CHW characteristic and programme feature and its definition (if provided) were extracted. For existing classification systems, the following information was extracted:

- a. The logic underlying how the characteristics were selected, described and grouped
- b. The names and definitions of the classification system's groupings
- c. Any information provided regarding strengths or weaknesses of the classification system

5.2.4 Data Synthesis

Inductive thematic analysis was used to identify, describe and categorise the full range of CHW characteristics and programme features. Efficacy studies and related studies were coded in alphabetical order by author; influential studies were then coded by year, from earliest to latest. During this process codes were added and refined. Codes were then verified against a list of characteristics and features extracted from included information classification systems. All codes were re-checked against the definition for a CHW characteristic or programme feature and assessed for redundancy or overlap. The final codes were then nested into a series of progressively broader and more general categories, forming the taxonomy (Hedden, 2010).

The end result was a prototype faceted classification system and an accompanying glossary defining operational terms. The process of development, involving repeated testing and refinement, from the initial stages of design through to the final version of the taxonomy is described in the results.

5.3 Results

5.3.1 Results of the Search: Primary Studies and Influential Papers

Two hundred and fifty-three records were included in taxonomy construction (Fig 5.3). Electronic searches retrieved 22 process evaluations or qualitative studies connected to the 155 records in Chapter 4 and zero such documents connected to the 14 performance trials in Chapter 6. Five hundred and twenty influential records were retrieved via citation analysis. Excluding duplicates, a total of 352 records were screened for inclusion. 125 full text papers were considered for inclusion in the review; 62 records met the inclusion criteria. Influential records were excluded for having non-LMIC countries as their subject (21), already being included as a trial in one of the systematic reviews (10), or not having CHWs as their subject (32).

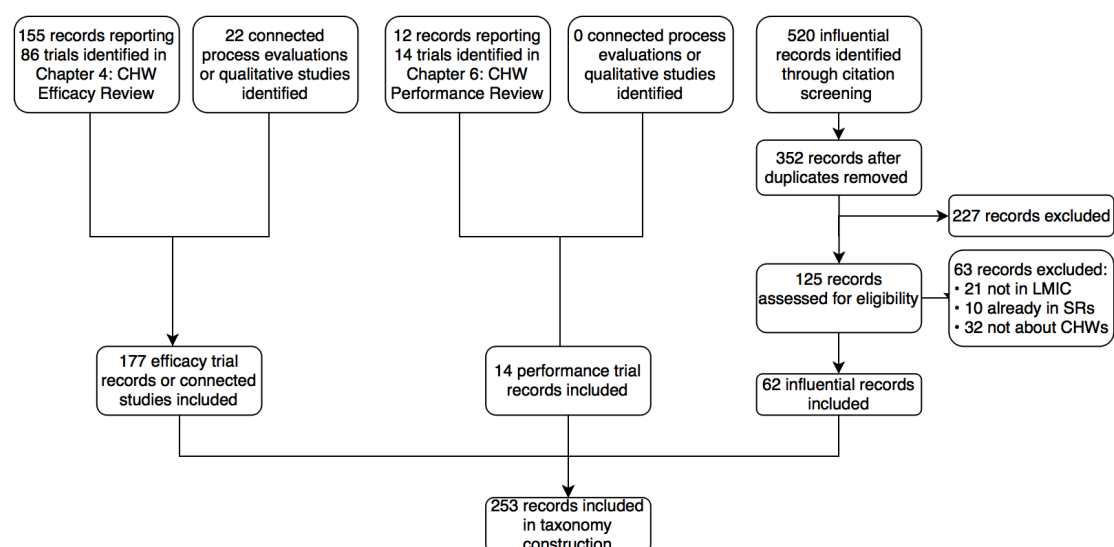


Figure 5.3 Identification of records for taxonomy construction

The articles identified in the review and their key characteristics are listed in Appendix 5a. Among trials, publication years ranged from 1986 to 2016; among influential studies the range was 1981 to 2013. Of the 191 papers reporting 100 trials included, twenty-four studies⁵ (24%) were conducted in ten upper-middle-income countries: Brazil (2), China (2), Iran (1), Iraq (1), Jamaica (4), Mexico (1), Paraguay (1), South Africa (11), Thailand (1), and Turkey (1). Forty-eight studies (48%) were conducted in fifteen lower-middle-income countries: Bangladesh (8), Cameroon (1), Ghana (6), Guatemala (1), India (12), Kenya (2), Myanmar (2), Nigeria (1), Pakistan (8), Philippines (1), Senegal (1), Swaziland (1), Vietnam (2), and Zambia (2). Twenty-nine studies (29%) were conducted in nine low-income countries: Burkina Faso (1), Burundi (1), Ethiopia (3), Malawi (2), Mali (1), Nepal (2), Uganda (11), Tanzania (7), and The Gambia (1).

5.3.2 Results of the Search: Existing Attempts to Classify CHWs

Electronic searches for existing CHW information classification systems retrieved 576 records; searches in the grey literature produced an additional 19 records. After duplicates were removed, 524 records were screened for inclusion. Of these, 31 full text records were assessed. Nine records reporting 8 existing information classification systems met the inclusion criteria (Fig 5.4). Records were excluded for containing no framework (6), having non-LMIC countries as their subject (5), not having a framework that contained CHW or programme characteristics (6), or not having CHWs as their subject (5).

⁵ One study was conducted in both an upper-middle income (China) and lower-middle income (India) country and so is counted twice in this list.

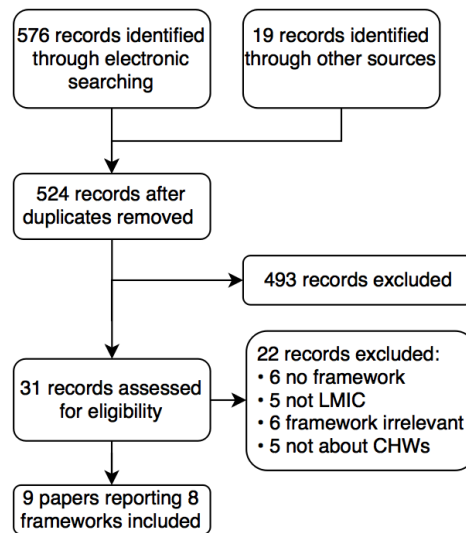


Figure 5.4 Identification of existing information classification systems

Existing information classification systems related to CHWs and their key characteristics are listed in Table 5.2. The information classification systems varied in their objectives, complexity, domains, and groupings. The frameworks were based on literature reviews (Bhutta et al., 2010; Glenton et al., 2013a; Glenton et al., 2011b; Kok et al., 2014; Naimoli, Frymus, Wuliji, Franco, & Newsome, 2014), programme data and professional experience (Palazuelos et al., 2013), or on unspecified methods (Crigler & Hill, 2009; Crigler et al., 2011).

Table 5.2 Existing CHW information classifications systems

#	Paper ID	Name and Type of Framework	Objective	Method	Number and Names of Items in the First Tier
1	Crigler 2009, Crigler 2011	CHW-AIM: CHW-PFA Facets	Help Ministries, donors and NGOs assess and strengthen their CHW programmes	Identified "15 proven best practices" in unspecified way	<u>Fifteen:</u> (1) CHW Role; (2) Advancement; (3) Individual Performance Evaluation; (4) Recruitment; (5) Initial Training; (6) Continuing Training; (7) Supervision; (8) Programme Performance Evaluation; (9) Equipment and Supplies; (10) Incentives; (11) Community Involvement; (12) Documentation, Information Management; (13) Referral; (14) Linkages to Health System; (15) Country Ownership
2	Glenton 2013	Unnamed Facets	Explore factors affecting the implementation of CHW programmes for MCH	Cochrane systematic review of qualitative evidence. Findings organised in a logic model using framework thematic synthesis	<u>Five:</u> (1) Integration of the LHW programme with other health services; (2) Working conditions, training and supervision; (3) User fees; (4) LHW selection criteria; (5) Visible support and participation from credible sources
3	Glenton 2011	A Typology of LHW Programmes for Immunisation Facets	Assess effects of CHW delivered interventions on the uptake of childhood immunisation and develop typology of intervention models	Systematic review of RCT, NRCT, CBA, and ITS studies of any intervention delivered by CHWs and designed to increase childhood immunisation uptake	<u>Four:</u> (1) Organisation of care, (2) LHW tasks, (3) LHWs selected on the basis of being peers, (4) LHWs given some form of remuneration
4	Kok 2014	Conceptual framework of factors influencing CHW performance Facets	Contribute to the evidence base needed by policymakers in the development of CHW programmes	Systematic review to identify intervention design related factors influencing performance of CHWs	<u>Six:</u> (1) Nature of tasks and time spent on delivery; (2) Human resource management; (3) Quality assurance, (4) Community links, (5) Health system links; (6) Resources and logistics.
5	Naimoli 2015	Community Health Worker generic logic model Facets	Promote early and continuous causal thinking as decision makers design, implement, scale up, and evaluate CHW	Review of the existing research on community and health system support for enhanced CHW performance plus a facilitated discussion of review findings with experts	<u>Six:</u> (1) Written policies and programmes; (2) People; (3) Funding; (4) Organisations and facilities; (5) Material and equipment; and (6) Time

Continued on next page

#	Paper ID	Name and Type of Framework	Objective	Method	Number and Names of Items in the First Tier
<i>Continued from previous page</i>					
6	Palazuelos 2013	5-SPICE Facets	Structure how implementers and researchers understand, study and improve their respective programmes	Programme experience of I-NGO used to identify essential ingredients "whose components and interactions should be considered in unison when structuring or analysing a CHW program"	<u>Five</u> : (1) Supervision, (2) Partners (3) Incentives, (4) Choice, (5) Education
7	Bhutta 2010	CHW Core Typology Facets	Assess the evidence base of the impact /effectiveness of global experience of CHWs in delivering care related to health and nutrition MDGs	Constructed based on a global systematic review and eight country case studies, taking into consideration expert opinion	<u>Nine</u> : (1) Recruitment; (2) Educational criteria; (3) Training content, duration & role (initial & ongoing); (4) Certification process; (5) Monitoring supervision & evaluation; (6) Volunteer/ salaried; (7) General or Performance based incentive; (8) Career pathway & development; (9) Referral system
CBA : controlled before–after studies; CHW AIM : Community Health Worker Assessment and Improvement Matrix; CHW-PFA : The CHW Programme Functionality Assessment Tool; HPN : Health, Population and Nutrition; ITS : interrupted time series; LHW : Lay Health Worker, alternative name for CHW; MCH : Maternal and child health; MDGs : Millennium Development Goals; NRCT : non-randomised controlled trials; RCT : randomised controlled trials; USAID : United States Agency for International Development					

No study had the primary aim of conducting a coding taxonomy and so none of the resulting information classifications were optimised or operationalised for trial reporting or coding. Two studies produced logic models in order to identify causal pathways to improved CHW performance (Glenton et al., 2013a; Naimoli et al., 2014). While some components (e.g. CHWs have a career pathway) meet the definition of CHW characteristic or programme feature, others (e.g. CHWs see their incentives as...appropriate and fair in relation to their tasks and level of training) are second-order concepts upon which it would be difficult or impractical to classify CHWs. Similarly, the Community Health Worker Assessment and Improvement Matrix (CHW-AIM) was designed to assess the functionality of CHW programmes as opposed to classify the various types of CHW that exist (Crigler & Hill, 2009; Crigler et al., 2011). While several of the assessed domains would be relevant to a CHW

taxonomy (training, supervision, incentives) they do not seem to have been assembled systematically—the authors specify no methods—and each is defined with suggestions for a particular configuration (e.g. “Regular supervision visits every 1–3 months”) rather than delineated in a way that allows for all possible configurations to be captured.

One was a taxonomy of types of immunisation interventions (Glenton et al., 2011b); while it incorporated CHW characteristics (e.g. remuneration) the primary focus of the taxonomy was on the intervention being delivered, not the CHW delivering it.

Three studies had outputs more closely related to a CHW taxonomy. Both Palazuelos et al. (2013) and Kok et al. (2014) created conceptual frameworks of CHW programmes and Bhutta et al. (2010) created MDG-specific CHW “types.” Though the 5-SPICE framework put forth by Palazuelos et al. (2013) contains relevant components (e.g. supervision, incentives) these are too loosely defined and not sufficiently detailed to be useful in coding. For example, supervision was described only as “management plans and structures.” Moreover, the components were primarily assembled based on the professional experience of collaborators associated with one NGO and so may not be comprehensive. While Kok et al. (2014) undertook a systematic search, because their objective was to isolate factors influencing the performance of CHWs (as opposed to classify all existing CHW types), their list of characteristics and programme features may likewise not be comprehensive. As explained in section 5.1.3, frameworks designed—as this one was—to aid in the development of programmes are not well suited to coding. Indeed, even Kok et al. (2014) note that “a general vocabulary related to different types of CHWs” would be helpful (p. 1222). Like CHW-AIM, Bhutta et al. (2010) make recommendations for

ideal CHW programme configurations. Though their MDG-specific “types” contain information on relevant categories (e.g. educational criteria, training duration), it is not clear how these categories were derived and the framework itself is not presented in a useable format for coding. This is demonstrated clearly when the country case studies included in the review are summarised using a reductive descriptor of the format “[long OR short] duration training programmes, with [description] tasks for CHWs, with [strong OR weak] supervision activities, and within a [strong OR weak] health system.”

Thus, while these classification systems are relevant to the objective of this paper, none have shared this objective or produced the desired output. This paper aims to build on the insights of these information classification systems by using systematic, transparent, fit-for-purpose methods to produce a clear, usable coding taxonomy of CHW characteristics and programme features.

5.3.3 Construction of the Taxonomy

Refining codes

The evolution of the thematic analysis as data from the different types of studies was extracted and coded is charted in Fig 5.5. Efficacy studies and related studies were coded in alphabetical order by author and produced the set of codes in the box to the left (1.0). Influential studies were then coded by year, from earliest to latest. During this process codes were added and refined; this is reflected in the list of codes in the box to the right (2.0). New codes from the influential papers were higher-order programme organising principles (whether CHWs were generalists or specialists, whether they were intended to be agents of political reform or health extension workers) and longer term issues (opportunities for advancement) that may not be of immediate concern in a trial setting but are relevant to generalising findings

for the purpose of programme organisation, sustainability, and scale. Finally, codes (2.0) were verified against a list of 37 characteristics and features extracted from included information classification systems; no new codes were added. Theoretical saturation—the point at which characteristics were no longer emerging from the data and all extracted features fit into categories already devised—was confirmed (Bowen, 2008).

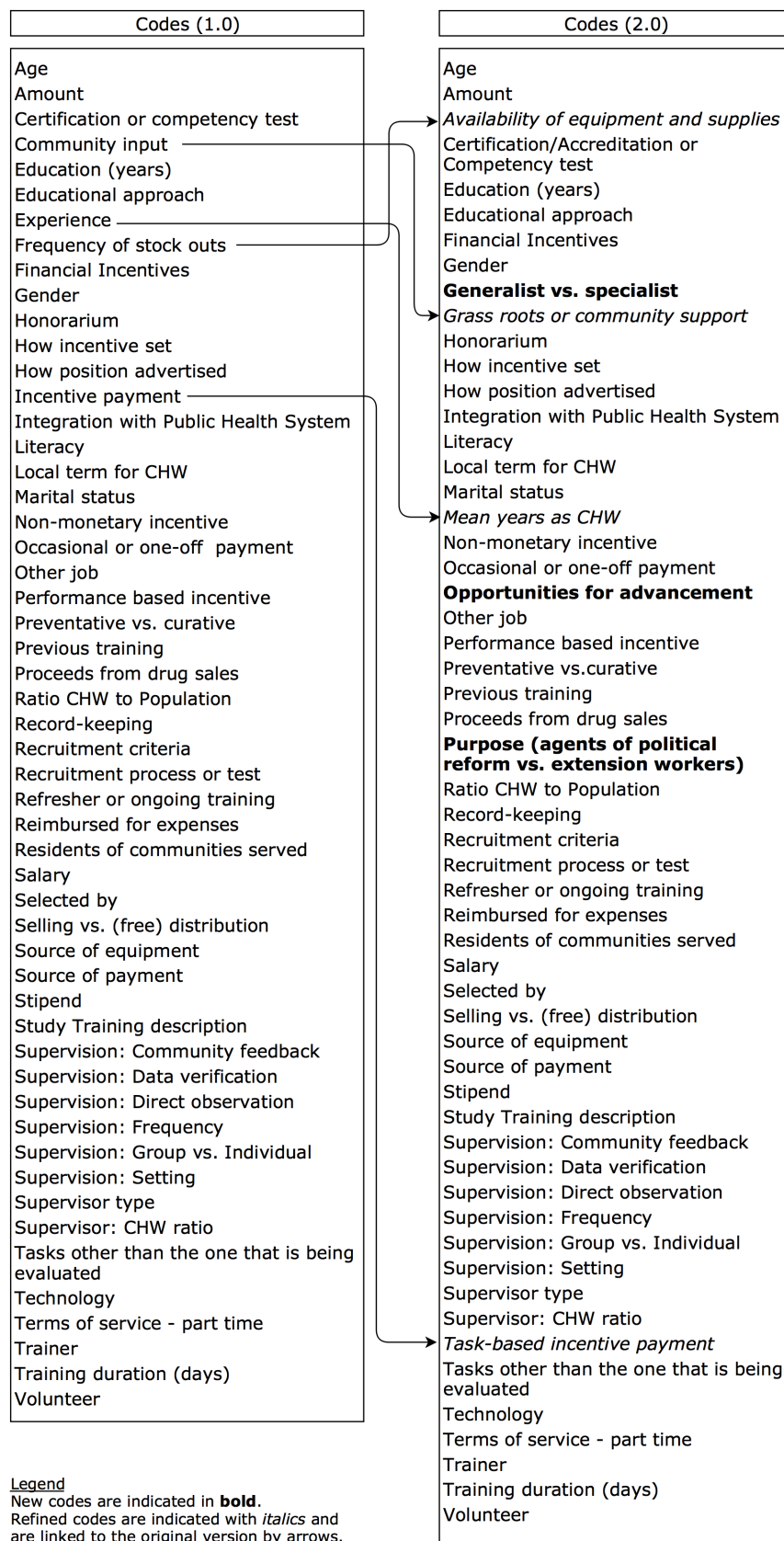


Figure 5.5 Refining of codes

Refining the taxonomy

All codes in codes (2.0) were re-checked against the definition for a CHW characteristic or programme feature provided in section 5.2.3. Three codes were eliminated for being related to the task being performed rather than the delivery agent (preventative vs. curative care, selling vs. free distribution of services, technology employed). Codes were also checked for redundancy or overlap; three codes (honorarium, salary, stipend) were combined into one concept. These changes are reflected in the final list of codes (3.0) in Fig 5.6.

Codes (3.0)	
Age	Record-keeping
Amount	Recruitment criteria
Availability of equipment and supplies	Recruitment process or test
Certification/Accreditation or Competency test	Refresher or ongoing training
Education (years)	Reimbursed for expenses
Educational approach	Residents of communities served
Financial Incentives	<i>Salary/stipend/honorarium</i>
Gender	Selected by
Generalist vs. specialist	Selling vs. (free) distribution
Grass roots or community support	Source of equipment
How incentive set	Source of payment
How position advertised	Study Training description
Integration with Public Health System	Supervision: Community feedback
Literacy	Supervision: Data verification
Local term for CHW	Supervision: Direct observation
Marital status	Supervision: Frequency
Mean years as CHW	Supervision: Group vs. Individual
Non-monetary incentive	Supervision: Setting
Occasional or one-off payment	Supervisor type
Opportunities for advancement	Supervisor: CHW ratio
Other job	Task-based incentive payment
Performance based incentive	Tasks other than the one that is being evaluated
Preventative vs. curative	Technology
Previous training	Terms of service - part time
Proceeds from drug sales	Trainer
Ratio CHW to Population	Training duration (days)
Record-keeping	Volunteer

Eliminated codes are indicated with ~~strike through~~ text. Combined codes are indicated with *italics*.

Figure 5.6 Final list of codes

The breadth of the coded CHW characteristics and programme features meant they did not lend themselves to being summarised using a single classification

dimension and so a faceted taxonomy was developed. After grouping codes into like categories, two categories and six dimensions emerged (see Fig 5.7). These were used as the organising schema of a faceted CHW taxonomy; because facets do not map well, the taxonomy is represented as a coding and reporting form in Fig 5.8.

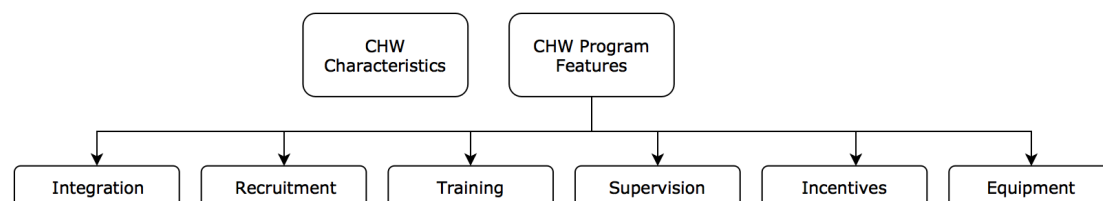


Figure 5.7 Dimensions along which CHW characteristics and programme features are categorised in the taxonomy

A link to download the full taxonomy in FileMaker Pro format (for ease of extraction and reporting) is provided in Appendix 5b. The definition of each facet is provided below along with an example. Citations for examples used to illustrate each facet are listed in Appendix 5c.

5.3.4 Explanation and Elaboration of The Taxonomy

CHW Characteristics

Because CHWs can vary considerably in their ascribed attributes, a brief description of the CHWs' characteristics enables easy identification of the type of health worker being assessed. It should be noted that the intention of this section is to capture *actual* characteristics; required characteristics should be listed under facet 16 'Recruitment Criteria.'

1. **Local term for CHW:** The term used to refer to CHWs in the areas they serve (e.g. "Saksham Sahayak").
2. **Age (mean):** The average age of evaluated CHWs (e.g. "41 years").

CHW Taxonomy v.1			
CHW Characteristics			
1. Local term for CHW		2. Age (Mean)	
3. Gender (%) Male Female		4. Literacy ☐ Yes ☐ No	
5. Terms of service: ☐ Full Time ☐ Part Time		6. Locally resident ☐ Yes ☐ No	
7. Years experience (Mean)		8. Years education (Mean)	
Programme Features			
Integration			
9. CHWs recognised as part of the formal health system		☐ Yes	☐ No
10. Clinics cooperate w/CHWs in patient care (e.g. referral)		☐ Yes	☐ No
11. Role played by community in supporting CHWs			
Recruitment			
12. Position advertising		13. Ratio CHW: Population	
14. Recruitment process			
15. Selection by:		16. Recruitment criteria (list)	
☐ Community ☐ Research Team			
☐ Health System ☐ Other			
Training			
	Pre-study		Study
17. Duration			
18. Format (descr.)			
19. Content (descr.)			
20. Trainer			
21. ☐ Specialist ☐ Generalist		List tasks not under eval:	
	Yes	No	Describe
22. Continuous training	☐ Yes	☐ No	
23. Certification/Test	☐ Yes	☐ No	
Availability of equipment and supplies			
24. Source of equipment		25. Stockouts inhibit work ☐ Yes ☐ No	
Supervision			
26. Frequency	27. Setting	28. Supervisor type	
29. Procedure:	Yes	No	Yes No
Verify reports	☐ Yes	☐ No	Observe CHW activities ☐ Yes ☐ No
Verify/restock supplies	☐ Yes	☐ No	Provide additional training ☐ Yes ☐ No
30. ☐ Individual ☐ Group		31. Perf evaluation (descr)	
Incentives			
32. Type:		Non-financial (Descr., Frequency)	Financial (Amount, Frequency)
☐ Task/Perf-based			
☐ One-off			
☐ Salary			
☐ Sale proceeds			
☐ Reimbursement			
☐ None			
33. Source of payment			
34. How incentive set			

Figure 5.8 Faceted CHW Taxonomy

3. **Gender (%)**: The gender breakdown of evaluated CHWs (e.g. “BRAC [decided] to switch to an exclusively female cadre” would be inputted as “0%” Male, “100%” Female).
4. **Literacy**: Whether evaluated CHWs were able to read (e.g. “mother coordinators [were] ...illiterate” would be inputted by selecting the “no” radio button).
5. **Terms of service**: Whether being a CHW is considered a full time or part time occupation (e.g. “[CHWs] were expected to work four hours a day” would be inputted by selecting the “part time” radio button).
6. **Locally resident**: Whether evaluated CHWs were from the communities they serve (e.g. “The intervention was delivered by four women, resident in [name of town] would be inputted by selecting the “yes” radio button).
7. **Years of experience (Mean)**: The average number of years evaluated CHWs have been serving in the community (e.g. “5 years”).
8. **Years of education (Mean)**: The average number of years evaluated CHWs have spent in school (e.g. “10 years”).

Programme Features

Integration

The support CHWs receive from surrounding health facilities and the community they serve may have a bearing on their ability to carry out certain health tasks. These linkages are coded in the following way:

9. **CHWs recognised as part of the formal health system**: Whether CHWs are employed by and/or recognised by the government (e.g. “[CHWs] are

employees of the Ghana Health Services” would be inputted by selecting the “yes” radio button).

10. Clinics cooperate with CHWs in patient care: Whether CHWs received support from local health centres in terms of integrated training, referral, or other provisions or the strength of the facility-based care had a bearing on the outcome of the trial (e.g. “The interaction between the main effects of GP training and [CHW-led home health education] on the primary outcome approached significance” would be inputted by selecting the “yes” radio button).

11. Role played by community in supporting CHWs: Description of the support, if any, offered to evaluated CHWs (e.g. “active involvement of communities...to promote use of community workers by community members”).

Recruitment

CHW selection is not standardised and pre-requisites vary widely. Given the ongoing debate about the best way to recruit, it is important to describe the process.

12. Position advertising: How potential CHWs were made aware of the position (e.g. “Women were informed about the training programme by word of mouth and by distribution of leaflets”).

13. Ratio CHW: Population: The size of the population to which each CHW was responsible for providing care (e.g. “one female community health worker for...about 4000 people” would be inputted as 1:4000).

14. Recruitment process: Describe how CHWs were chosen for the role (e.g. “Formal tests and an interview were conducted prior to an observation during

role plays...We trained more people than required for the intervention and selected those who performed best to undertake the work”).

15. Selection by: Indicate by whom CHWs were selected (e.g. “[CHWs] were nominated by fellow patients at each clinic site with final approval by [clinic] staff if they met all qualifications” would be inputted by checking the “Community” and “Health system” boxes).

16. Recruitment criteria (list): List CHW recruitment criteria (e.g. “the minimum requirements for being a [CHW] were knowledge of the local language, being literate, preferably passed higher secondary school, and motivated to be involved in the community care of older people”).

Training

Unlike other cadres of health workers (e.g. doctors), CHW training and licensing is not always nationally defined and consistently enforced. It is thus important to provide adequate detail regarding how CHWs are prepared for their roles.

17. Duration: Length of training evaluated CHWs received prior to the study and as part of study (e.g. “All study CHWs had previously undergone 6 wk. training before becoming community health workers. As part of the study, they participated in an additional 5-d training workshop” would be inputted as Pre-study: “6 weeks”, Study: “5 days”).

18. Format: Describe how training provided prior to the study and as part of study was delivered (e.g. “[CHWs] received a combination of classroom-based and apprenticeship-based field training”).

19. Content: Describe the content area covered in the training provided prior to the study and as part of study (e.g. “Training on basic HIV pathogenesis,

prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form”).

20. **Trainer:** Identify who delivered the training provided prior to the study and as part of study (e.g. “the principal investigator”).
21. **Specialist vs. Generalist:** Indicate whether CHWs provided care for more than one ailment, regardless of whether outcomes for all ailments were evaluated in the trial. List these additional tasks (e.g. “we integrated a cognitive behaviour therapy-based intervention into the routine work of community-based primary health workers” would be inputted by selecting the “Generalist” radio button and providing a description of the pre-existing tasks).
22. **Continuous training:** Indicate whether CHWs received ongoing training and, if yes, describe (e.g. “received refresher training every 3 to 4 months during the implementation of the intervention”).
23. **Certification/Test:** Indicate whether CHWs had to pass a test before interacting with patients and, if yes, describe (e.g. “an assistant quizzed each CHW using photographs to ensure all participants could distinguish between strong positive, faint positive, negative, and invalid [rapid diagnostic malaria test] results”).

Availability of equipment and supplies

Given the complexity of supplying community-based programmes and the frequency with which stock-outs are highlighted as barriers to programme implementation, it is important to describe the source and availability of CHW equipment and supplies.

24. **Source of equipment:** Describe who supplies the drugs and equipment (e.g. “drugs...must [typically] be purchased from a government pharmacy...for the

period of the study...[CHWs] were supplied with antimalarials by the project”).

25. **Stockouts inhibit work:** Indicate whether stock-outs inhibit work (e.g. “[CHWs] frequently run out of stock” would be inputted by selecting the “yes” radio button).

Supervision

Given the range of supervisory support, it is important to adequately describe how CHWs’ work is observed and directed.

26. **Frequency:** The frequency with which supervision takes place (e.g. “Fortnightly”).
27. **Setting:** Where supervision takes place (e.g. “during home visits and community meetings” vs. “in a health facility”)
28. **Supervisor type:** Describe the qualifications of the supervisor (e.g. “World Vision Uganda staff with prior [interpersonal psychotherapy in a group format] experience”).
29. **Procedure:** Indicate what activities constitute supervision (e.g. “supervisors...visited them at least twice a month, observed meetings, and provided training” would be inputted by checking the “yes” radio button next to “Observe CHW activities” and “Provide additional training”).
30. **Individual vs. Group:** Whether supervision took place one-on-one or in a group context (e.g. “Peer support, through group meetings, may also make an important contribution to morale and motivation” would be inputted by selecting the “Group” radio button).
31. **Performance evaluation:** Describe how CHW performance is evaluated, if at all (e.g. “Performance assessment of Saksham Sahayaks included feedback

from community members, spot checks by their supervisors during home visits and community meetings to assess their level of community engagement, and monitoring by the supervisors of whether targets for home visits and community meetings were being met”).

Incentives

Describe how CHWs are motivated in or compensated for their work.

32. **Type:** Indicate the type of incentives that were provided and provide a brief description in the appropriate box (e.g. “Each supervising mother received, as payment, approximately US\$4 for each home visit” would be inputted by checking the box labelled “Task-based” and inputting the amount and frequency into the “Financial Incentive” column. “Provided occasional incentives (i.e., bicycles, umbrella, and stationery) when resources were available” would be inputted by checking the box labelled “One-off” and inputting the description into the “Non-financial incentive” column “A monthly salary of 3800 Pakistani rupees (about US\$45)” would be inputted by checking the box labelled “Salary” and inputting the amount and frequency into the “Financial Incentive” column. “Peer counsellors were not paid a salary, but received local transport costs during training and home visits” would be inputted by checking the box labelled “Reimbursement” and inputting the description into the “Financial Incentive” column. “No financial or other incentives were provided to patients or staff” would be inputted by checking the box labelled “None”).
33. **Source of payment:** Indicate who supplies the incentive (e.g. “Supported by their community and do not receive any payment from the Ministry of Health”).

34. **How incentive set:** Describe how the type and level of the incentive was chosen (e.g. “the equivalent of a minimum wage in Brazil at the time” or “All CHWs were paid a salary in line with existing South African government stipends for CHWs”).

The focus of this taxonomy is on reporting details of the CHW characteristics and programme features. This taxonomy is designed to be used in conjunction with other tools appropriate for coding and reporting various study designs (e.g. CONSORT and TIDieR in the case of randomised trials, TREND in the case of nonrandomised trials etc.).

5.4 Discussion

5.4.1 Summary of Main Results

Systematic searches and citation screening identified 253 CHW-focused trials, connected process evaluations, and influential papers. These, along with nine records reporting eight existing information classification schemes identified through systematic searches, were included in taxonomy construction.

Inductive thematic analysis was used to identify, describe, and categorise the full range of CHW characteristics and programme features. Trials, related studies, and influential papers were coded for CHW characteristics and programme features; codes were added and refined throughout this process (Fig 5.5). These codes were verified against characteristics and features extracted from included information classification systems, checked for redundancy or overlap, and re-checked against the definition for a CHW characteristic or programme feature provided in the methods section; theoretical saturation was confirmed and a final list of codes produced (Fig 5.6). Like codes were then grouped into two main and six sub-categories (Fig 5.7) to create a

comprehensive taxonomy that enables thorough, standardised reporting and coding of CHW characteristics and programme features (Fig 5.8).

The breadth of the CHW characteristics and programme features did not lend themselves to being summarised using a single classification dimension and so a faceted taxonomy was developed. While this does not allow for the creation of a tidy “vocabulary” of CHW types as has previously been requested (Kok et al., 2014) a faceted taxonomy better represents the complex nature of health delivery and the vast array of possible CHW configurations. The taxonomy is complementary to but does not replace the CONSORT or TREND guidance; it does not include, for example, categories for context or types of interventions CHWs deliver.

5.4.2 Strengths and Contributions of the Taxonomy

In order to ensure its reliability and utility, the taxonomy was iteratively considered against seven literature-derived quality criteria during its construction. Explained in section 5.2.1, the criteria specified that: the taxonomy (i) be jointly exhaustive yet parsimonious and (ii) feasibly used to report the required information in published studies; included characteristics be (iii) observable and testable empirically, (iv) conceptually unique, (v) defined in a way that is accurate and easily understood, and (vi) grouped in a clear and meaningful manner; and (vii) future CHW characteristics and programme features can be included in the taxonomy.

The taxonomy meets these strength criteria in the following way: (i) In order to ensure the taxonomy is jointly exhaustive, a comprehensive, systematic search was undertaken to identify published and unpublished descriptions of CHW characteristics and programme features from academic studies, national governments, aid agencies, and non-governmental organisations. The use of facets enables the taxonomy to capture the vast array of possible CHW configurations without oversimplification into

a single classification dimension. Moreover, because this is a coding taxonomy, developed primarily for the purpose of categorising intervention content, items were included whether or not they have an established empirical evidence base—this further ensures the taxonomy’s comprehensiveness. Nonetheless, this thoroughness is balanced with parsimony: items were included only if they could be an “active ingredient” of the larger intervention.

(ii) Despite a relatively high number of facets, the taxonomy is structured to ensure ease of complete reporting: balancing hospitality with standardisation and efficiency, over 80% of the items in the taxonomy are radio buttons, checklists, and fill-in-the-variable questions. Assuming a conservative response rate of 25 seconds per question, the taxonomy should take less than 15 minutes to complete. Moreover, the taxonomy is presented in Filemaker format to make it easy for both (a) information to be extracted from studies and (b) taxonomy facets to be filled in and included as an appendix in newly published studies.

(iii) That all included characteristics could be observed and tested empirically was part of the definition for included components; items were checked and re-checked against this criterion to ensure its application.

(iii) In order to ensure conceptual uniqueness, the included characteristics were also checked and re-checked for redundancy and overlap, with several codes being eliminated or combined over the course of the coding process.

(iv) In order to ensure accuracy and relevance, definitions were derived from those commonly used in the included records. Different descriptions of the same concept were distilled in a final definition thought to best promote ease of understanding.

(vi) Facets are grouped into two broad categories (CHW Characteristics and Programme Features) and six sub-categories (Integration, Recruitment, Training, Availability of equipment and supplies, Supervision, and Incentives). These organising principles were derived directly from programme descriptions and so ought to meaningfully reflect the range of CHW experience. Moreover, because items were only included if they were the smallest component compatible with retaining the proposed active ingredient, facets are easy to understand and interpret.

(vii) The decision to use facets allows for post-coordinated indexing. This means the taxonomy can accommodate new information regarding the meaningful ways CHWs can vary as it emerges. This hospitality ensures that future CHW characteristics and programme features can be included in the taxonomy. As a result of these strengths, the taxonomy makes several important contributions.

Firstly, it provides a foundational step toward standardisation and accuracy of reporting CHW characteristics and programme features, as required by CONSORT and TREND. As explained in the introduction, TiDiER (Hoffmann et al., 2014), the extension of CONSORT to randomised trials of nonpharmacologic treatment (Boutron et al., 2008), and TREND (Des Jarlais et al., 2004) require only that the “case volume,” “qualification,” and “specific training” of providers be described. Given the range of vastly different health workforce configurations routinely classed as “CHW programmes,” adhering to guidance from these extensions may still frequently yield descriptions interventions where salient features go unreported. This taxonomy clarifies and defines relevant CHW characteristics and programme features and provides standardised language with which to describe them. Moreover, it is presented in an easy-to-use form that trialists can complete and include as a

descriptive appendix in new studies, thereby facilitating reporting that is consistent and complete.

Secondly, the taxonomy adds much-needed operational nuance to CHW delivered intervention components. For example, several trials in the CHW efficacy review described in Chapter 4 identified “supervision,” “training,” and “selection” as critical to the success of the CHW-delivered intervention they were testing, but none defined the terms (Agrasada et al., 2005; Haider et al., 2000; Rahman et al., 2008; Rotheram-Borus et al., 2014). Rather than treat these as monolithic concepts, the taxonomy delineates the components of each. For example, supportive supervision is broken into four discrete tasks (verifying reports, restocking supplies, observing home visits, and collecting community feedback) that may all be separate active ingredients in the delivery agent. Similarly, given the lack of standardisation across the CHW experience, different aspects of training (e.g. whether it is continuous, how CHWs are evaluated at its conclusion) and selection (e.g. what criteria were established, who makes the final decision) were drawn out to help identify what may be critical differences. This conceptual mapping enables (i) relevant, precise reporting of CHW programme features often cited as important but rarely adequately described and (ii) the definitional clarity required to run component selection experiments, identified by the Multiphase Optimisation Strategy (MOST) as helpful in identifying the optimal combination (and dose) of intervention components (Collins et al., 2011).

Thirdly, this will enable systematic reviews to use a standardised method for extracting information about CHW delivered intervention content. This will aid in (i) the identification of areas where research is strong and where gaps exist, (ii) the reliable synthesis of evidence, and (iii) further exploration of discrete, replicable

characteristics (or combinations of characteristics) as potential mechanism of action associated with intervention effectiveness (Michie et al., 2013).

Fourthly, this enhanced classification and characterisation will allow for more accurate replication of CHW programme models. Not only does this enable (i) the investigation of generalisability across behaviours, populations, and settings, but (ii) the operationalisation of the practice standards that permit the faithful propagation and uptake of interventions found to be efficacious or effective (Abraham & Michie, 2008; Michie et al., 2013).

5.4.3 Weaknesses of the Taxonomy and Future Work

The study has three primary weaknesses. Firstly, it is difficult to draw the line between what counts as a feature of the delivery system and what counts as a component of a particular intervention. While the author laid out a clear definition for CHW characteristic and programme feature in the methods section and attempted to apply it consistently through the extraction, coding, and taxonomy construction process, it is possible that there may be disagreements regarding included facets. Secondly, there was only one coder. Multiple independent coders would have improved the validity of the facet generation process (Buscemi, Hartling, Vandermeer, Tjosvold, & Klassen, 2006). Thirdly, the search strategy relied heavily on electronic resources; as such, the taxonomy may only reflect those programmes that have been documented online.

Future work could further strengthen this taxonomy. Given the rigorous, systematic methods employed in this taxonomy's construction, it joins several recent high-quality taxonomies in the medical field that have been constructed solely on the basis of existing literature (e.g. Lowe et al., 2011; Patel et al., 2015; Reader et al., 2014). That said, subjecting this taxonomy to revision through a consensus process

would (a) help refine it, (b) ensure its relevance to those who would be employing it, and (c) possibly help with uptake (Grant et al., 2013). As the checklist gets applied over time, it may also be possible to reduce the number of facets via use of factor analysis to find correlated variables (Child, 2006).

The taxonomy has also not been validated as a coding mechanism. Empirical testing by health practitioners who have not been involved in development of the taxonomy is needed to provide evidence of validity and test-retest reliability. Undertaking these three processes will, in turn, allow the strengths listed above to be further realised.

5.5 Conclusion

Werner (1981) first bemoaned the “confusing array” of CHWs over thirty-five years ago (p. 48). More recent papers noting the difficulty of precise classification given “the wide range of [CHWs]” indicate that confusion persists to this day (Haines et al., 2007, p. 2122). While enormous intellectual energy has been spent testing ‘tasks’ to shift to CHWs (cf. Chapter 4), less energy has been expended on issues related to systematically disaggregating the types of personnel involved. This failure to identify research and disaggregate types of CHWs may mean that particular characteristics (literacy, CHW remuneration) that distinguish between interventions remain unidentified. If such unreported or “unseen” differences are “active ingredients” associated with effectiveness, the design of optimally effective interventions (Collins et al., 2011) and the uptake and replication of such interventions will be impeded. In answer to repeated calls for a comprehensive taxonomy that allows for the conceptualisation of CHW delivered interventions “in terms of their lay component rather than in terms of the specific health issues [tasks] on which they focus” (Lewin et al., 2010; van Ginneken et al., 2013), systematic

review methods were used to develop a faceted taxonomy of CHW characteristics and programme features.

Such a taxonomy allows for consistent, precise, and complete reporting of relevant CHW-related information as required by the CONSORT statement and TREND. It provides the definitional clarity required for optimisation of CHW programmes through component selection experiments and systematic review subgroup analysis and enables the reliable identification and synthesis of evidence required to identify research gaps. The clear delineation of relevant CHW characteristics and programme features allows for more accurate replication of CHW programme models and, in turn, investigation of generalisability and the operationalisation of the practice standards that permit the faithful propagation and uptake of efficacious interventions.

While consensus methods may be used in future to further refine the taxonomy, it remains a long overdue, foundational first step in the comprehensive and standardised description of CHW characteristics and programme features.

Chapter 6: A Systematic Review of Interventions for Improving the Performance of Community Health Workers

6.1 Introduction

In line with aim three of this thesis, the objective of this chapter is to assess the effects of interventions to improve the performance of community health workers for community-based primary health care outcomes in low- and middle-income countries.

6.1.1 The Enigma of CHW Performance

As noted in Chapter 1, many countries initiated national CHW programmes or scaled-up local initiatives to the national level following the 1978 Alma Ata Conference (Bhattacharyya et al., 2001). A series of reviews in the late 1980s and early 1990s, however, found that these large-scale community health worker programmes failed to replicate the success of smaller community-based programmes (Berman et al., 1987; Frankel, 1992; Gilson et al., 1989; Heggenhougen et al., 1987; Walt, 1988, 1990). As a result, enthusiasm for CHWs began to wane among donor organisations and ministries of health (Campbell & Scott, 2011) and by the end of the 1990s, most national CHW programmes had been disbanded (Abbatt, 2005; Lehmann & Sanders, 2007; Schneider et al., 2008; Segall, 2003; Standing & Chowdhury, 2008).

These programmes fell victim to extrinsic structural and economic factors and

intrinsic institutional factors (Standing & Chowdhury, 2008). Extrinsically, health sector reform stemming from the global economic recession in the 1980s, including cuts to public health systems, undermined community level service delivery (Cueto, 2004; Hall & Taylor, 2003; Walt, 1990). Intrinsically, inadequate training, logistical support, incentivisation, supervision, and integration with the wider health system led several reviews to argue that CHWs had not failed, but rather, their potential had not been realised (Berman et al., 1987; Frankel, 1992; Gilson et al., 1989; Walt, 1988, 1990).

Prompted by the HIV/AIDS pandemic, the slow pace of progress in reaching the health-related MDGs, and questions about how to improve coverage and equity of care in the face of a global health workforce shortage, the early 2000s saw a renewed interest in the use of CHWs (Christopher et al, 2011).

During this period, rigorous evidence accumulated on the efficacy of CHWs to deliver assorted health interventions (Kredo et al., 2014; Lassi et al., 2010; Lewin et al., 2010; Okwundu et al., 2013). As demonstrated in Chapter 4, the largest systematic review of CHW efficacy and effectiveness RCTs to date, there is high quality evidence that CHW delivered interventions reduce perinatal mortality, improve child nutritional status (i.e. WAZ), and improve TB completion rates versus facility-based care. There is also moderate quality evidence that CHW delivered interventions reduce the prevalence of depression, improve PTSD symptoms, reduce functional impairment related to CMD, reduce child morbidity, reduce neonatal mortality, increase exclusive breastfeeding and the speed of its initiation, and that facility-led maintenance of ARVs is no better than CHW-led maintenance in terms of loss to follow-up or death.

Yet the most recent evaluations of national-scale CHW programmes—

published in the *American Journal of Tropical Medicine and Hygiene* in March—remain unfavourable (Amouzou et al., 2016a; Amouzou et al., 2016b; Munos et al., 2016). A dose-response analysis of health surveillance assistants in Malawi, a randomised controlled trial of health extension workers in Ethiopia, and a controlled evaluation of agents de santé à base communautaire [community-based health workers] in Burkina Faso all found no improvement in their primary outcomes of interest. While Munos et al. (2016) noted that it was not possible to draw conclusions about the effectiveness of the integrated community case management (iCCM) strategy evaluated given implementation shortcomings, Amouzou et al. (2016a); Amouzou et al. (2016b) documented strong implementation yet found no improvements in care seeking or mortality.

As explained in Chapter 1, statistical significance and optimisation are distinct and independent concepts and methods in intervention science have largely been developed to measure the former (Collins et al., 2011). Thus, without more insight into intervention components, it is difficult to assess whether a given large-scale CHW programme did not achieve a statistically significant effect (i) because it does not work, (ii) because it was not implemented properly (type III error), or (iii) because it was not yet optimised (Basch, Sliepcevich, Gold, Duncan, & Kolbe, 1985; Borrelli, 2011; Collins et al., 2011; Hasson, 2010). Given that qualitative assessments as far back as Berman et al. (1987) suggest it is often the latter, a thorough investigation of the relationship between outcomes and individual intervention components is warranted. Unfortunately, it remains the case that, as Frankel (1992) noted two decades ago, our knowledge of the mediators and moderators of CHW performance is limited (Global Health Workforce Alliance, 2012; Naimoli, Frymus, Quain, & Roseman, 2012).

At the time this chapter was written, no systematic review had yet employed rigorous evidence to address the relative effectiveness of different types of CHW characteristics or programme features and their effect on CHW performance (Lassi et al., 2010; Swider, 2002). While history has demonstrated that broad contextual factors related to the health system or the political and economic climate can have an impact on community health worker performance, this review will be focused on modifiable aspects of the intervention itself, namely: CHW characteristics and programme features. By identifying head-to-head comparison trials of the same health-related tasks delivered by CHWs with different characteristics or types of support, this chapter will seek to disaggregate the impact of different intervention components (Collins et al., 2011). In keeping with MOST and building on the results from Chapter 4, the question of how to optimise CHW performance—via different approaches to recruitment, training, supervision, incentivisation or other factors—is addressed here.

6.1.2 How the Intervention Might Work

There is little consensus about how an ideal CHW programme should be structured or what elements are most crucial for its effectiveness (Palazuelos et al., 2013). While researchers and aid agencies have developed a number of tool kits to improve programmes (Crigler et al., 2011; The University of Arizona Rural Health Office and College of Public Health, 1998), a widely recognised conceptual framework for CHW programme design and analysis does not exist. In answer to repeated calls for a “comprehensive” taxonomy that allows for the conceptualisation of CHW delivered interventions “in terms of their lay component rather than in terms of the specific health issues [tasks] on which they focus” (Lewin et al., 2010, p. 42; van Ginneken et al., 2013, p. 49), systematic review methods were used in Chapter 5

to develop a faceted taxonomy of CHW characteristics and programme features. Two categories and six dimensions emerged (Fig 6.1).

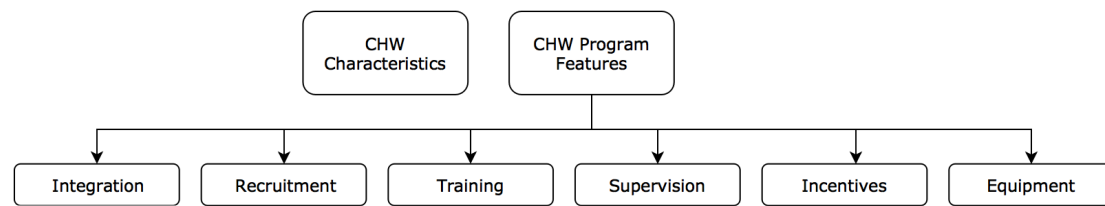


Figure 6.1 CHW taxonomy categories from Chapter 5

While the Chapter 5 taxonomy provides a useful classification system for the types of interventions that may strengthen CHW performance, the interventions considered for this review were not restricted to those targeting the elements identified. Any program-level intervention aimed at helping health workers better perform their existing tasks was included. For example, observational work has suggested interventions targeting, for example, data management or workload may also help CHWs perform better (Baawo et al., 2014; Jaskiewicz & Tulenko, 2012; Strachan et al., 2012).

Qualitative evidence on CHW performance indicates that poor outcomes are associated with low levels of CHW motivation (De Zoysa & Cole-King, 1983); poor retention (Alam, Tasneem, & Oliveras, 2012b); few opportunities for career advancement (Chevalier, Lapo, O'Brien, & Wierzba, 1993); poor community relations (Glenton et al., 2013a); and negative interactions with managers (Glenton et al., 2013). The ways in which CHW characteristics and programme features highlighted in the taxonomy might alter these modifiable factors are explained below.

CHW Characteristics

Several papers have noted that ascribed, rather than achieved, characteristics may have a bearing on CHW performance. Observational studies and narrative

reviews have identified characteristics like village origin, sex, and age as possible mediators of CHW performance (Bhattacharyya, 2001; López Quiñones, 1999; Ofosu-Amaah, 1983). CHWs from inside the community, for example, may be more invested in their surroundings and so more likely to be retained in employment.

CHW Programme Features

Integration: Strong partnerships—between CHW and community as well as CHW and the formal health system—are associated with CHW performance: respectful, regular contact with the healthcare system is correlated with strong programme outcomes (Gilson et al., 1989; Haines et al. 2007; Lehmann et al. 2004) and observational evidence suggests programmes developed and managed by communities address perceived local needs more closely than those developed by outsiders (Lehmann & Sanders, 2007). Letting communities design the terms of reference for their CHW, for example, may improve access to care.

Recruitment: Several reviews cite appropriate CHW selection as a precursor to success, noting that selecting workers who are not respected by the community tends to result in low community engagement with the programme (Bhattacharyya et al., 2001; Campbell & Scott, 2011; Lehmann & Sanders, 2007). While the importance of recruitment is, in principle, nearly universally acknowledged, uncertainty remains about how best to operationalise the process; some reviewers assert that a balance must be struck between community selection and input from the health system (Ofosu-Amaah, 1983) while others maintain that solely the community ought to have an input (Bhattacharyya et al. 2001, Lehmann et al., 2004). An appropriate recruitment strategy, for example, might ensure candidates are widely respected in their communities and thus better able to improve behavioural health outcomes.

Training: Across reviews, there is support for the implementation of ongoing

training, including multiple refresher and advancement sessions (Lehmann et al., 2004). Consistent with the evolution of the conceptualisation of the CHW role, biomedical, skill-based competency training has gradually replaced training focusing on the social aspects of health (Campbell & Scott, 2011; Gilroy & Winch, 2006; Gilson et al., 1989). Specifics about where training ought to take place, who ought to conduct it, how it should be designed etc. remain unclear. More rigorous post-training assessments, for example, might reduce CHW mistakes and thus improve physical outcomes for patients.

Supervision: Because CHWs are situated on the periphery of the health system, supervision has been cited as an important way to maintain quality and engagement (Campbell & Scott, 2011; Lehmann & Sanders, 2007; Stinson, Malianga, Marquez, & Madubuike, 1998). While policy documents note supervisors should be “competent and have appropriate supervisory skills,” the specifics of supervision are poorly defined: questions of who should supervise (e.g. peers vs. nurses), how often, and what should occur on such visits, remain unresolved (World Health Organization et al., 2008, p. 4). Regular supervision, for example, might increase the accountability of CHWs and thus increase quality of care.

Incentives: The incentivisation of CHWs has been much debated, with entire narrative reviews being devoted to the topic (Bhattacharyya et al., 2001). While there is disagreement in the literature about the ideal form of compensation (i.e. monetary vs. non-monetary) that CHW motivation must be addressed is not in dispute (World Health Organization, 2007). The provision of monetary incentives, for example, might increase retention among workers who were previously volunteers.

Equipment: Cross-sectional surveys of CHWs indicate that they are often limited in their work by irregular supplies (Puett, Alderman, Sadler, & Coates, 2013).

Interventions to decrease the number of stock-outs, for example, may improve patient cure rates.

6.1.3 Why it is Important to do this Review

While narrative reviews have thus far isolated principles (e.g. “appropriate supervisory skills”) associated with CHW performance, none have identified actionable strategies to improve performance (e.g. “weekly supportive supervision by a nurse composed of supply audits, patient checks, and on-the-job training”). At the time of its writing, this was the first systematic review to focus entirely on the effectiveness of program-level interventions to improve the performance of CHWs.

6.2 Methods

All stages of this systematic review were undertaken by the DPhil candidate alone, except where indicated. While systematic reviews are typically conducted in teams, this approach was taken to demonstrate for the purposes of university examination the DPhil candidate’s own and sole expertise in systematic reviews. Additional expertise will be sought to review the analysis ex post and prior to publication.

6.2.1 Criteria for Considering Studies for this Review

Types of Studies

Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, randomised controlled trials, cluster-randomised trials, crossover trials, and factorial trials were eligible for inclusion. Based on guidelines from the Cochrane EPOC Review Group, if fewer than two controlled trials had been found, data from well-designed non-randomised intervention studies such as interrupted time series with at least three time points before and after

intervention, cohort, and case-control observational studies would have been included (Effective Practice and Organisation of Care Group, 2015).

Types of participants

Community health workers:

Several definitions for and variants of the term community health worker have been employed in the literature. For the purpose of this review, the definition provided in Chapter 1 will be used. Community health workers are any lay health workers who:

- Live in the catchment they serve
- Are primarily based in the community (as opposed to a health facility)
- Belong to the formal health system (i.e. are managed by the government or an implementing NGO)
- Perform tasks related to healthcare delivery
- Have received organised training but have no formal or paraprofessional certification or tertiary education degree

Although CHWs have been used in high-income countries, differences in healthcare budgets, education levels, access to care, and disease profile make it unproductive to include such actors in the comparison.

Recipients of care:

Aside from residence in a low- or middle-income country as classified by The World Bank Group (listed in Appendix 3a), there were no restrictions on the types of patients for whom data were extracted.

Types of Interventions and Comparators

Any intervention designed to improve the performance of CHWs, compared to CHWs who did not experience the intervention. Studies were only included where the

primary objective, or one of the primary objectives, was to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs. This includes studies like ‘head-to-head’ comparisons of CHWs delivering the same task with different types of support that have to date been excluded from reviews of CHW effectiveness (e.g. Lewin et al., 2010) and were excluded from the systematic review reported in Chapter 4.

Studies that were designed to examine the efficacy of a particular CHW activity (e.g. zinc tablets for diarrhoea), as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity were excluded. Interventions of any duration or follow up were included; had sufficient interventions been identified, they would have been stratified by length of follow-up (e.g. 6 months, 1 year, >1 year) during the analyses.

Also excluded were interventions in which:

- Healthcare was provided to a subset population (i.e. in a school, in a church etc.) and not the wider community. Such interventions are qualitatively different from other CHW delivered interventions as the population and care-taking environment may not be representative.
- The CHW was a family member trained to deliver care and support only to members of his or her own family (i.e. the CHW did not provide care or service to others in the community). Such interventions are qualitatively different from other CHW delivered interventions as they are precipitated on a bond between caregiver and care receiver that could affect the process and/or effects of the intervention.
- Lay people were trained to manage their own health (i.e. lay self-management) rather than the health of their community.

Types of Outcome Measures

A number of processes and outcomes might be affected by interventions that aim to improve the performance of CHWs. Reported outcomes were extracted and categorised as follows; listed outcomes of interest were not used as criteria for including studies.

- Primary:
 - Objective physical outcomes for patients
 - Behavioural health outcomes for patients
- Secondary
 - Utilisation of services indicators (e.g. mean number of consultations/month)
 - Quality of care provided by CHW
 - CHW retention

6.2.2 Search Methods for Identification of Studies

Electronic searches

A comprehensive search strategy was used to identify all relevant studies regardless of language, year, or publication status (published, unpublished, in press, and in progress). The following strategy was refined in consultation with a Social Sciences Data Librarian/Information Specialist and used to search MEDLINE:

1. ((community or village) adj2 (agent\$ or aid\$ or promot\$ or mobili?er\$ or distribut?r\$ or worker\$)).ti,ot,ab,kf.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary\$ or extension) adj2 (worker\$ or volunteer\$)).ti,ot,ab,kf.
3. (accompanier\$ or accompagnateur\$ or activista\$ OR animatrice\$ OR

brigadista\$ or kader\$ or monitora\$ or promotora\$ or sevika\$ or fhw\$ or
 chw\$ or lhw\$ or vhw\$ OR chv\$ or "shastho shebika" or "shasto karmis" or
 anganwadi\$ or "barefoot doctor" OR "agente comunitario de salud" or "agente
 comunitario de saude").ti,ot,ab,kf.

4. (performance or effective\$ or skill\$).ti,ot,ab,kf.
5. 1 or 2 or 3
6. 5 and 6

Terms 1-3 search for the population of interest and term 4 searches for the interventions evaluated. In order to maximise sensitivity, no randomisation filter was included (Khan, Ter Riet, Glanville, Sowden, & Kleijnen, 2001). The strategy was adapted for other databases using the appropriate controlled vocabulary, as applicable,

Databases searched:

1. AMED, Ovid
2. BiblioMap: EPPI-Centre database of health promotion research
3. British Library for Development Studies at IDS
4. CINAHL, Ebsco
5. Cochrane Central Register of Controlled Trials (CENTRAL)
6. Database of Promoting Health Effectiveness Reviews (DoPHER)
7. EMBASE, Ovid
8. Global Health, Ovid
9. IDEAS database of unpublished working papers
10. Impact Evaluation Database (3ie)
11. IPA (Innovations for Poverty Action) Database
12. Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index - Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities)
13. J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database
14. MEDLINE, Ovid
15. MEDLINE In-Process & Other Non-Indexed Citations, Ovid
16. POPLINE
17. Pubmed

18. WHOLIS

19. World Bank's JOLIS search engine

Searches were conducted May 4, 2014 and June 30, 2014 and were updated on May 11, 2016 and July 12, 2016. Precise dates of coverage, date last searched, and search strategies used for the databases listed above can be found in Appendix 6a.

Searching other resources

Additional studies were identified through:

1. Handsearch of (a) seven journals from 1960 (or first online issue) to May 2014 (*Health Policy and Planning*, *Journal of Community Health*, *Social Science and Medicine*, *Human Resources for Health*, *BMC Public Health*, *Global Health Action*, and *Global Public Health*) and (b) conference proceedings from the American Society of Tropical Medicine and Hygiene.
2. Review of websites of nineteen leading global public health research institutions, donors, implementers, and technical agencies: International Child Health Review Collaboration; World Health Organization; United Nations Children's Fund; United Nations Development Programme; United Nations Population Fund; Joint United Nations Programme on HIV/AIDS; United States Agency for International Development; UK Department for International Human Development Resource Centre; Partners in Health; the Global Fund to Fight AIDS, Tuberculosis and Malaria; Bill & Melinda Gates Foundation; Clinton Foundation; Centre for Global Development; University Research Co-Center for Human Services (URC-CHS); CARE, Frontline Health Workers Coalition; One Million Community Health Workers Campaign; Global Health Workforce Alliance; and CHW Central.

3. Search of grey literature databases relevant to public health: Open GREY, New York Academy of Medicine Grey Literature Report, BIOSIS, and Proquest Dissertations & Theses.
4. Search of trial registries for on-going research: WHO International Clinical Trials Registry Platform, the ISRCTN Register, and ClinicalTrials.gov.
5. Direct contact with study authors and leading experts in the field of community health to solicit potentially relevant unpublished papers, on-going research, and suggestions for other contacts (see Appendix 6a for full list).
6. Examination of the reference lists of related reviews and primary studies to identify other studies that meet the inclusion criteria.

6.2.3 Data Collection and Analysis

Selection of studies

Titles and abstracts were independently screened by two authors in order to identify studies that met the eligibility criteria. Both authors then retrieved the full text of studies selected as meeting or possibly meeting the criteria and verified them again in order to produce a final list of included studies. Any discrepancies were discussed until agreement was reached.

Data extraction and management

One reviewer extracted data on study design, participants, interventions, outcomes and methodological quality, using a pre-designed data collection form. The extraction sheet was informed by the Cochrane EPOC Group data collection checklist, the Template for Intervention Description and Replication, and the Oxford Implementation Index; it was piloted on the first two studies (Effective Practice and Organisation of Care Group, 2015; Hoffmann et al., 2014; Montgomery et al., 2013b). In order to improve data quality, a second reviewer audited extracted data.

Interventions were expected to vary widely (e.g. increased supervision, monetary and non-monetary incentives etc.) and so special care was taken to note all relevant features so as to allow a careful consideration of the heterogeneity of the data set. Where reports were uncertain or included summary measures, authors were contacted for clarification. Any consistently underreported data were noted and highlighted along with future research needs in the discussion. Data were entered and managed using Excel; a copy of the data collection sheet is included in Appendix 6b.

Assessment of risk of bias in included studies

Risk of bias for all RCTs was assessed using The Cochrane Collaboration's tool for evaluating the risk of bias. This approach involves assessing the risk of bias across six domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential biases (Higgins & Green, 2011). As per guidance in the Cochrane Collaboration Handbook, additional information on (i) recruitment bias, (ii) baseline imbalances, (iii) loss of clusters, (iv) unit of analysis issues and (v) comparability with individually randomised trials was collected for cluster randomised trials. A judgment of low risk of bias, high risk of bias, or unclear risk of bias was made for each domain based on definitions in the Cochrane Handbook (Higgins & Green, 2011). Multiple-intervention studies were also checked to ensure data were presented for each of the groups to which participants were randomised. Had there been crossover trials, information on design suitability, the presence of a carry-over effect, whether only first period data are available, analysis issues, and the comparability of results with those from parallel-group trials would have been collected.

The quality of evidence across a body of evidence (i.e., multiple studies with similar interventions and outcomes) was assessed using the GRADE approach

(Guyatt et al., 2008b), defining the quality of evidence for each outcome as “the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest” (Higgins & Green, 2011, section 12.2.1).

Measures of treatment effect

To investigate the feasibility of a meta-analysis, all eligible studies were summarised in Excel to the fullest possible extent. Where possible, standardised measures of effect (risk ratios for dichotomous outcomes, standardised mean differences for continuous outcomes) have been provided to aid comparison. Three authors were contacted for additional data; none were able to supply it.

Unit of analysis issues

Cluster-randomised trials were verified to ensure appropriate analysis had been done (i.e. adjustment for clustering in calculating confidence intervals or p-values). Failure to account for clustering is likely to overestimate the precision of the study; if such an analysis had not been done, the necessary intraclass correlation coefficients (ICCs) were extracted or obtained from the investigators and results reanalysed (Higgins & Green, 2011). The reviewer reduced the size of each C-RCT to its ‘effective sample size’ using the design effect $(1 + (M-1) \times ICC)$, where M is the mean cluster size and ICC is the intraclass correlation coefficient (Rao & Scott, 1992). Where no information on the intra-cluster correlation coefficient (ICC) was reported in any of the cluster RCTs included in the analysis group, the reviewer assumed an ICC of 0.02 typical of primary and community care interventions (Campbell, Grimshaw, & Steen, 2000b). Where an ICC was reported among the studies in a group, this ICC was used for the adjustments to other studies; if multiple studies in a group reported ICCs the most conservative ICC was used. Where results were presented as cluster means the number of events in each group was estimated as

($N \times \text{cluster mean} / 100$). If this was not possible, point estimates, but not the reported confidence intervals or p-values, were reported. Had there been sufficiently similar studies to conduct meta-analyses, a sensitivity analysis would have been undertaken using imputed ICCs based on data from the other included studies (Campbell et al., 2000b).

For factorial studies, all comparisons were included. Had meta-analysis been conducted, all eligible intervention groups in studies with multiple eligible intervention groups would have been combined to estimate an average treatment effect compared with the control or, if intervention groups were sufficiently different that an average treatment effect would be meaningless (i.e. one group testing incentives and one testing supervision), the 'shared' group would have been split into groups with smaller sample size, and included in (reasonably independent) comparisons. For crossover trials, data would have been extracted and analysed only from the period before the crossover.

Dealing with missing data

Authors of each of the included studies were contacted to obtain missing statistics (i.e. unreported data, data necessary to standardise effect measures) and attrition was assessed as a form of bias. Unfortunately, few authors were able to provide the requested data.

Assessment of reporting biases

Selective outcome reporting was assessed using the approach described in the Cochrane Handbook (Higgins & Green, 2011). Had there been sufficient studies, funnel plots (plotting effect size against standard error) would have been used to detect potential bias. Because asymmetry can be due to clinical and methodological heterogeneity in addition to publication bias, these sources of heterogeneity would

have been examined as possible explanations (Peters et al., 2008).

Data synthesis

Given that the question asked in the review is broad with respect to the type of intervention and outcome considered, data were grouped by the type of intervention (i.e. recruitment, training, supervision, integration, incentives) and the outcomes measured. Had interventions and outcomes been sufficiently similar across the different studies, data would have been pooled and summary effect sizes estimated using random-effects models. Because no meta-analysis was conducted, a structured synthesis is provided. Where possible, standardised measures of effect (risk ratios for dichotomous outcomes, standardised mean differences for continuous outcomes) have been provided to aid comparison. Due to missing data, over a third of outcome measures were not possible to standardise and are reported as in the original trial papers.

Assessment of heterogeneity

Had meta-analyses been conducted, the extent of heterogeneity across comparable studies would have been assessed using forest plots, the I^2 statistic and the Chi^2 test.

Sensitivity analysis

A sensitivity analyses would have been performed by excluding studies with a high risk of bias or imputed data for any outcome for which more than one comparable study exists.

6.3 Results

6.3.1 Description of Studies

Results of the search

Excluding duplicates, a total of 6401 records were screened for inclusion. Of these, 6331 were recovered from electronic database searches ending in June 2014, 1 was from handsearching, 5 were from personal contacts, 3 were from cross-referencing, 32 were from grey literature databases, and 29 were from registers of ongoing studies. After screening the titles and abstracts, 40 records were deemed potentially relevant. Full-text copies (or sufficient information) for 39 of these citations were obtained. The paper not obtained is an unpublished Spanish-language report that could not be acquired from the authors; it was unclear from the title whether it met the inclusion criteria (Carrón, Melián, & León, 1993).

Of the 40 records, 33 were excluded for reasons detailed below (Fig 6.2); 7 records reporting 9 studies met the inclusion criteria (Ayele, Desta, & Larson, 1993; Bailey & Coombs, 1996; Carrón, Melian, & Leon, 1994; Chang et al., 2011; Derenzi et al., 2012; Omer, Mhatre, Ansari, Laucirica, & Andersson, 2008; Winch et al., 2003).

The search was updated in May and July of 2016. A total of 3717 documents were retrieved: 3713 from electronic databases and 4 from cross-referencing. After duplicates were removed, the abstracts and titles of 1681 studies were screened and 17 records deemed potentially relevant. Full-text copies for these citations were obtained; 12 were excluded for reasons detailed below (Fig 6.2) and 5 records reporting 5 studies met the inclusion criteria (Andreoni, Callen, Khan, Jaffar, & Sprenger, 2016; Ashraf, Bandiera, & Lee, 2015; Bossuroy, Delavallade, & Pons, 2016; Gautham, Iyengar, & Johnson, 2015; Ginsburg et al., 2014).

A total of 12 records reporting 14 studies were included in the updated systematic review. Seven studies were RCTs (Andreoni et al., 2016; Bailey & Coombs, 1996; Derenzi et al., 2012ABC; Gautham et al., 2015; Ginsburg et al., 2014) and seven studies (Ashraf et al., 2015; Ayele et al., 1993; Bossuroy et al., 2016; Carrón et al., 1994; Chang et al., 2011; Omer et al., 2008; Winch et al., 2003) were cluster-randomised trials with similar units of allocation (district, district, CHW catchment, health zone, group of districts, clinic catchment area, and health zone, respectively).

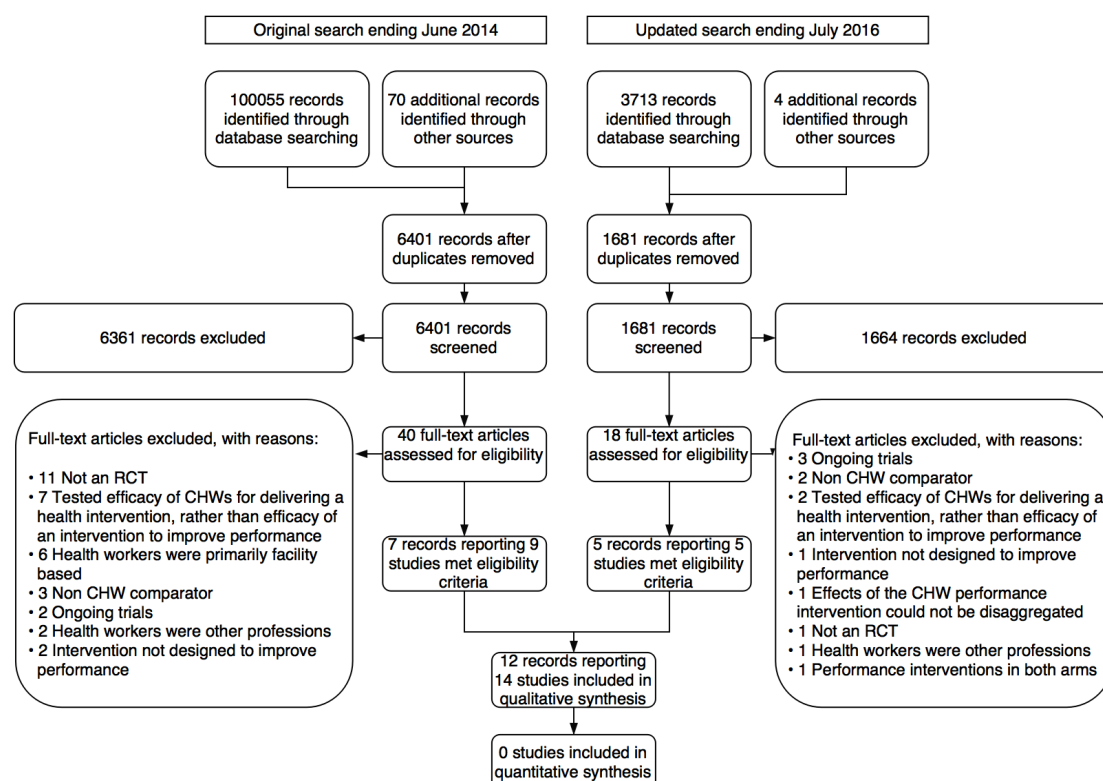


Figure 6.2 PRISMA study flow diagram

Included studies

Setting and intervention characteristics

While detailed information about each intervention is included in the Characteristics of included studies tables in Appendix 6d, salient features of each are provided below in summaries guided by the Template for Intervention Description

and Replication (TIDieR) checklist (Hoffmann et al., 2014) and summarised in Table 6.1.

Table 6.1 Brief Study Characteristics

Study	Setting/ Country (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
Andre- oni 2016	Rural & Urban Pakistan (2x2 factorial RCT)	Female CHWs, age not reported, $\bar{x}$ education 10 yrs ($I_{a+c}=85$, $I_{a+d}=84$, $C_{b+c}=88$, $C_{b+d}=80$). Patients were children under 5 years ($N=$ unspecified)	<u>Incentives</u> I: a+c; a+d C: b+c; b+d (a) Employment contracts tailored to individual discounting preferences (b) Untailored contracts (c) Advanced allocation of daily work targets (d) Day-of allocation	<u>CHWs:</u> Closeness to vaccine distribution policy objective <u>Patients:</u> None (Immediate)
Ashraf 2015	Rural Zambia (C-RCT)	CHWs, 30% female, $\bar{x}$ age 27 yrs., min. education 12 yrs. ($I=149$, $C=149$). Patients were women, mean 4.2 years education, & children under 5 years ($N =$ unspecified)	<u>Recruitment</u> I: Recruitment materials emphasising career possibilities C: Recruitment materials emphasising benefits to community	<u>CHWs:</u> Home visits, patients seen, comm. mtgs. organised <u>Patients:</u> Use of services (18 months)
Ayele 1993	Rural Ethiopia (C-RCT)	Male CHWs, $\bar{x}$ age 32 yrs, $\bar{x}$ education 5 yrs ($I=50$, $C=50$). Patient characteristics not reported ($N=$ unspecified)	<u>Training-Supervision</u> I: 5 day in-person refresher course and monthly supervision C: No refresher, no supervision	<u>CHWs:</u> Composite functional status score <u>Patients:</u> None (6 months)
Bailey 1996	Rural Guatemala (RCT)	CHW characteristics not reported ($I=22$, $C=27$). Patient characteristics not reported ($N=$ unspecified)	<u>Training</u> I: Interactive, tiered training with visual cue cards C: Didactic training	<u>CHWs:</u> Knowledge score <u>Patients:</u> None (Immediate)

Study	Setting/ Country (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
<i>Continued from previous page</i>				
Bossu- roy 2016	Urban India (C-RCT, 4 arm)	CHWs 72% male, 50% some university, age not reported (N=78). Adult patients 57.7% male, 57.6% no school or some primary (N = 2500)	<u>Incentives</u> I _a : Financial incentives based on patient detection for six months then based on treatment adherence I _b : Financial incentives based on patient detection for six months then a fixed salary I _c : Fixed salary for six months then incentives based on treatment adherence C: Fixed salary	<u>CHWs:</u> TB case detection <u>Patients:</u> TB default rate (6 months)
Carrón 1994	Rural Paraguay (2x2 factorial C-RCT)	Female CHWs, $\bar{x}$ age & education not reported (I _a =45, I _b =45, I _{a+b} =45, C=45). Patient characteristics not reported (N=unspecified)	<u>Training-Supervision</u> I: a; b; a+b (a) Trained using interactive service algorithm (Decision tree) (b) Role-play based supervision (Behavioural feedback instrument) C: Usual training and supervision	<u>CHWs:</u> Quality of care provided <u>Patients:</u> None (12 months)
Chang 2011	Rural Uganda (C-RCT)	Female CHWs, $\bar{x}$ age & education not reported (I=13, C=16). Adult HIV+ patients; 66% female, $\bar{x}$ age 34 yrs (I=446, C=524)	<u>Integration-Equipment</u> I: A text message reporting system and health centre warmline C: No mobile health intervention (usual care)	<u>CHWs:</u> None <u>Patients:</u> Virologic failure, adherence (2 years)
DeRenzi 2012A	Urban Tanzania (Pilot RCT)	Female CHWs, age 23-55 yrs., $\bar{x}$ education not reported (I=8, C=7). Patient characteristics not reported (N= unspecified)	<u>Supervision</u> I: Escalating text message reminder system and follow-up for underperforming CHWs C: Usual care	<u>CHWs:</u> None <u>Patients:</u> Change in closed referrals (24 days)

Study	Setting/ Country (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
<i>Continued from previous page</i>				
DeRenzi 2012B	Urban Tanzania (RCT)	CHWs, age 23- 63 yrs., $\bar{x}$ education not reported (I=34, C=40) Chronic care patients (N=unspecified)	<u>Supervision</u> I: Escalating text message reminder system & follow-up for underperforming CHWs C: Usual care	<u>CHWs:</u> None <u>Patients:</u> Mean time overdue at clinic (40 days)
DeRenzi 2012C	Urban Tanzania (RCT)	CHWs, age 23- 63 yrs., $\bar{x}$ education not reported (N=87) Chronic care patients (N=unspecified)	<u>Supervision</u> I: Escalating text message reminder system only C: Usual care	<u>CHWs:</u> None <u>Patients:</u> Mean time overdue at clinic (3 months)
Gauth- am 2015	Rural India (RCT)	CHWs, 50% female, $\bar{x}$ 37 yrs, >10 yrs education (I=8, C=8). Paediatric (45.2%) and adult (54.8%) patients, 35.7% male 64.3% female (I=65, C=61)	<u>Equipment</u> I: Mobile phone-based, media-rich procedural guidance applications C: Paper guidelines	<u>CHWs:</u> Protocol compliance <u>Patients:</u> None (2 months)
Gins- burg 2014	Rural Bangla- desh (3- arm RCT)	Female CHWs, age not reported, high school education (I _a =10, I _b =10, C=10). Female patients, $\bar{x}$ 38 yrs (I _a =246, I _b =240, C=70)	<u>Equipment (I_a)/</u> <u>Training-Equipment (I_b)</u> I: a; b (a) Smart phone with applications to guide interview, report data, show motivational video, and offer an appointment (b) As in (a) plus additional "patient navigation training." C: Paper-based tools	Patient adherence with referral (1 month)
Omer 2008	Urban Pakistan (C-RCT)	Female CHWs, $\bar{x}$ age & education not reported (I=52, C=unspecified). Mothers and their children (I=529, C=541)	<u>Integration-Training</u> I: Indigenous, evidence- based communication tool C: Usual care	<u>CHWs:</u> Home visits <u>Patients:</u> Uptake of protective health practices (10 months)

Study	Setting/ Country (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
<i>Continued from previous page</i>				
Winch 2003	Rural and Urban Mali (C-RCT)	CHW characteristics not reported (N=~102). Children <14 yrs. (N=152, C=134)	<u>Integration -Training- Reporting-</u> I: Visual aids, refresher training, a counter referral recordkeeping notebook, a community meeting C: Standard training	<u>CHWs:</u> Proportion of children treated at home <u>Patients:</u> Closed referrals

$\bar{x}$: mean; **C**: control; **Comm**: community; **C-RCT**: cluster RCT; **I**: intervention; **N**: intervention/control participant breakdown unspecified, total number of participants; **mtg**: meeting; **RCT**: randomised controlled trial;

While the taxonomy presented in the background section is helpful for classifying intervention *components* it is not as helpful for classifying studies themselves; nearly half of the interventions tested were compound and simultaneously targeted several of the listed CHW characteristics and programme features—integration, recruitment, training, supervision, incentives, and equipment—hypothesised to affect CHW performance. Nonetheless a categorisation of interventions by components is as follows:

Recruitment Interventions

One study, conducted in rural Zambia, reported on a recruitment intervention (Ashraf et al., 2015). CHWs in intervention clusters (N = 149) were recruited using posters that made career opportunities salient (e.g. “Become a community health worker to gain skills and boost your career!”). CHWs in control clusters (N = 149) were recruited using posters that made benefits to the community salient (e.g. “Want to serve your community? Become a community health worker!”). CHWs in both groups then attended a one-year residential training to learn basic medical practices, were paid 290 USD/month, were supervised by unspecified actors at an unspecified frequency, and provided basic primary care in their communities. No information

about implementation fidelity or the total number of patients treated was provided.

Training Interventions

One study, conducted in rural Guatemala, reported on a training intervention (Bailey & Coombs, 1996). In Bailey & Coombs (1996), CHWs treated diarrhoea; recruitment was not specified. Intervention CHWs (N = 22) were trained using the 'kader' method, which employs (i) tiered training whereby trainers teach trainees exactly as the trainers themselves have been taught, (ii) interactive role-playing by trainees using "Counselling Cards" to practice counselling or teaching, and (iii) visual cue-cards featuring diagnostic algorithms. Control CHWs (N = 27) received usual didactic training. CHWs in both arms provided diarrhoea case management. All CHWs were volunteers and were supervised by total health workers. No information about implementation fidelity was provided.

Supervision Interventions

Three trials reported on supervision interventions (De Renzi et al. 2012ABC). All three trials were conducted in urban Tanzania. In all three studies, CHWs had been working in their communities prior to the study and using a mobile form platform for longitudinal client tracking, 'CommCare'. Training was not described in any of the studies.

In DeRenzi et al. (2012A) all CHWs collect routine health information and have the option to refer sick clients to a health facility. CHWs in the intervention group (I = 8) continued to use CommCare but also received an SMS message two days after they had reported a referral, reminding them of the follow-up visit. Daily SMS messages were sent until the follow-up was recorded. Control CHWs (C = 7) continued to use CommCare and did not receive SMS messages. No monetary incentives were provided.

DeRenzi et al. 2012C followed directly after DeRenzi et al. 2012B; the context and intervention used in the two studies are largely similar so the interventions are described together. Both DeRenzi et al. (2012B) and DeRenzi et al. (2012C) studies used the same CHWs ($N = 87$); these CHWs provide home visits to chronic care patients. In 2012B, intervention CHWs received an escalating reminder system that sent SMS reminders directly to the CHW before notifying the CHW's supervisor after several overdue days. Control CHWs received no additional input. In 2012C, CHWs were re-randomised. Intervention CHWs received the same intervention as in 2012B. Control CHWs received an SMS-only condition in which CHWs were provided with proactive and overdue reminders, but escalation to the supervisor never occurred. No monetary incentives were provided to the CHWs in either study.

No information about implementation fidelity or the total number of patients treated was provided for any of the three studies.

Incentives Interventions

Two trials reported on supervision interventions (Andreoni et al., 2016; Bossuroy et al., 2016). Andreoni et al. (2016) was conducted in rural and urban areas of Pakistan and Bossuroy et al. (2016) was conducted in urban areas of India. Prior to the intervention, all CHWs in Andreoni et al. (2016) were provided a smart-phone equipped with a reporting application to track when and where vaccinations were conducted; CHWs in Bossuroy et al. (2016) received no baseline input.

In Andreoni et al. (2016) structural estimates of time preferences were used to customise incentives for existing CHW vaccinators based on the preferences measured during a series of government-organised polio vaccination drives. CHWs were offered a fixed bonus of 1000 rupees (~10 USD) for completing a total of 300

vaccinations over a two-day drive. Either in advance of the first day of the drive or on the first day of the drive, vaccinators set daily targets for vaccinations on day 1 (v_1) and day 2 (v_2) of the drive. If either of the vaccination targets, v_1 or v_2 , were not met, the bonus was not received. Vaccinators were randomly assigned an interest rate, R , such that a single vaccination allocated to day 2 reduced by R the number of vaccinations required on day 1. The individual discounting parameters observed in the first drive were used to tailor incentive contracts in a follow-up drive. There were two intervention groups: (A) CHW contracts tailored to individual discounting patterns with advance allocation choice ($N = 85$) and (B) CHW contracts tailored to individual discounting patterns with day-of allocation choice ($N = 84$). There were also two control groups: (A) Untailored CHW contracts and advance allocation choice ($N = 88$) and (B) Untailored CHW contracts and day-of allocation choice ($N = 80$). At the end of each day of the drive, CHWs convened with their supervisor at the local clinic.

Bossuroy et al. (2016) assessed whether financial incentives provided to health workers may encourage them to detect new tuberculosis cases and improve treatment adherence. CHWs ($N = 78$) were randomly assigned to receive either a fixed salary or a salary dependent on the number of patients they had detected. In the following six months, they were randomly re-assigned to either a fixed or an incentivised salary scheme, based on the number of defaults. Supervision was not described.

Baseline health training of the CHWs was not described in either study, though CHWs in Andreoni et al. (2016) were trained on the intertemporal bonus contracts and the process by which allocations were made and submitted. No information about implementation fidelity was provided about either of the two trials.

Equipment Interventions

Two trials assessed equipment interventions: Gautham et al. (2015) took place in rural India and Ginsburg et al. (2014) in rural Bangladesh. In Gautham et al. (2015) CHWs had been working in their communities prior to the study, in Ginsburg et al. (2014) CHWs were recruited by the study team.

In Gautham et al. (2015) all CHWs received a 2-day training by physicians in the local language on guideline-based care and in use of the mHealth system. After the training, CHWs in the intervention group ($N = 8$) returned their field sites with the applications on their mobile phones, while CHWs in the control group ($N = 8$) had the data card with the application removed from their mobile phones and used a set of paper guidelines. CHWs in both arms then provided management of fevers, diarrhoea and respiratory problems to paediatric and adult patients ($I = 65$, $C = 61$). Supervision and incentives were not specified.

In Ginsburg et al. (2014) all CHWs received a 3-day training on breast anatomy, physiology, pathology, and clinical breast examination. CHWs in arm A ($I_a = 10$) received a smart phone with applications to guide interview, report data, show a motivational video, and offer an appointment; CHWs in arm B ($I_b = 10$) received the same smart phone application as those in arm A plus an extra day of training in patient navigation. Arm A was thus classified as an equipment intervention and Arm B as a training-equipment intervention. CHWs in the control arm ($C = 10$) received no smart phone and recorded data on paper. CHWs in all arms interviewed adult women aged 25 and older in order to identify women with abnormal clinical breast exam ($I_a = 246$, $I_b = 240$, $C = 70$). CHWs were supervised by local women; incentives were not described. No information about implementation fidelity was provided about either of the two trials.

Compound Interventions

Six studies reported compound interventions. Ayele et al. (1993) and Carrón et al. (1994) assessed training-supervision interventions, Chang et al. (2011) assessed an integration-equipment intervention, Omer et al. (2008) an integration-training intervention, and Winch et al. (2003) an integration-training-reporting intervention. Each will be described in turn. Arm B of Ginsburg et al. (2014) reported a training-equipment intervention; this was described in the previous section.

Ayele et al. (1993) and Carrón et al. (1994) assessed training-supervision interventions. The Ayele et al. (1993) trial was conducted in rural Ethiopia; the Carrón et al. (1994) trial in rural Paraguay. In Ayele et al. (1993), both intervention (N = 50) and control (N = 50) CHWs received a baseline input of registration books, standard Ministry of Health of Ethiopia monthly report forms, and a general meeting in each peasant association to discuss the CHWs' job description. The CHWs in Carrón et al. (1994) received no baseline input.

In Ayele et al. (1993), government CHWs conducted health promotion activities. Intervention CHWs received a five-day in-person refresher course and monthly "supportive" supervision in their villages (Ayele et al., 1993, p. 383). The course, provided by an unspecified actor, reviewed expected CHW healthcare provision functions and responsibilities; no details were given regarding specific supervisory activities undertaken. Control CHWs received nothing after baseline input. No details about incentives were provided.

In Carrón et al. (1994), CHWs recruited by a research organisation distributed family planning materials. Intervention CHWs in the first group (N = 45) received a service algorithm/decision tree ("ABC of the Distributor" - ABC) to use during patient encounters. Intervention CHWs in the second group (N = 45) received supervision by nurse-midwives using a novel tool ("Behavioural Feedback

Instrument” - BFI) designed to diagnose service behaviour of CHW and provide feedback. The BFI consists of role playing exercises in which the nurse-midwife enacts one of four client profiles. The instrument includes detailed instructions about reinforcing or corrective feedback to be given according to the type of response exhibited by the CHW. The instrument was explained to CHWs by nurse midwives in four 30-minute individual meetings. Intervention CHWs in the third group (N = 45) received both the ABC of the Distributor and the BFI. Control CHWs (N = 45) received the usual two-day annual family planning course and yearly two-day refresher training provided by unspecified actors in the capital, Asunción. No details about incentives were provided.

No information about implementation fidelity was provided in either trial. In Carrón et al. (1994), 446 patients were treated in the intervention group and 524 in the control; in Ayele et al. (1993), no information about the total number of patients treated was provided.

The Chang et al. (2011) integration-equipment C-RCT was conducted in rural Uganda. Both intervention (N = 13) and control (N = 16) CHWs received a baseline input of a two-day residential training on HIV provided by an unspecified actor, a bicycle, t-shirts, basic supplies, clinical supervision at unspecified intervals, and a monthly allowance of 12.5 USD. Recruitment was not specified. Intervention CHWs received a mobile phone, a one-day residential training provided by an unspecified actor, and a one hour-long field-based practicum training. They were asked to send a text message reporting adherence and clinical data back to a centralized database immediately after or during each home visit and were encouraged to call a clinical warmline with questions or concerns. Clinic staff receiving CHW texts and calls could decide to provide care instructions, send a higher-level care provider to the

patient, or arrange transport to the health centre. CHWs who did not receive the mHealth intervention continued providing regular adherence monitoring and psychosocial support to patients. No information about implementation fidelity was provided; 446 patients were treated in the intervention group, 524 in the control.

The Omer et al. (2008) integration-training C-RCT was conducted in eight rural and two urban enumeration areas in Pakistan. Government CHWs provided maternal care in their communities. Intervention CHWs (N = 52) received an indigenous, evidence-based communication tool to use with pregnant women (N = 529) during their home visits. The tool—an embroidered cloth developed in consultation with villagers, government health officials, and CHWs—was based on *ajrak*, a traditional Sindh fabric, and depicted maternal practices (e.g. attending and not attending antenatal check-ups, giving colostrum after birth etc.). An unspecified actor provided training in the tool, including role modelling and practice in the community. Control CHWs (N = unspecified) continued their usual work: visiting women (N = 541) before and after childbirth to discuss safe practices in pregnancy based on the “standard” CHW training, details of which were unspecified. Supervisors held monthly meetings at district level in both groups. No information about incentives or implementation fidelity was provided.

The Winch et al. (2003) integration-training-reporting study was conducted in rural and urban areas of Mali. Both intervention and control CHWs (N = 102 total, intervention/control breakdown unspecified) had previously received a thirty-five day literacy and a seven-day drug administration course from unspecified actors. Government CHWs distributed drugs to treat common illness; supervision and incentives were not described. Intervention CHWs received: (a) visual aids to explain to carers of young children (i) how to administer chloroquine and (ii) the symptoms

requiring immediate referral to a health facility. The aids were developed by the National Centre for Information, Education and Communication for Health (CНИЕCS) of Mali. (b) A referral notebook to send with referred children to and from the health centre in which health centre staff were to write the diagnosis, treatment given, and instructions for follow-up. (c) A five-day refresher training course conducted by unspecified personnel covering chloroquine administration, malaria danger signs, and communication skills. (d) A village meeting to explain CHW programme changes. Control CHWs received nothing beyond the basic training described. No information about implementation fidelity was provided; 151 patients were treated in the intervention group, 131 in the control.

Participants

Reporting of CHW characteristics was poor. Winch et al. (2003) and Bailey & Coombs (1996) did not report the sex, age or education level of CHWs. While the remaining twelve studies reported the sex of the workers, only six studies reported CHWs' age (range: 23-63 years) (Ashraf et al., 2015; Ayele et al., 1993; Derenzi et al., 2012ABC; Gautham et al., 2015; Ginsburg et al., 2014) and five the schooling received by the CHWs (range: 5-12 years) (Andreoni et al., 2016; Ashraf et al., 2015; Ayele et al., 1993; Gautham et al., 2015; Ginsburg et al., 2014).

Patient data were similarly poorly reported. Neither Ayele et al. (1993), Bailey & Coombs (1996), Carrón et al. (1994), nor DeRenzi et al. (2012ABC) reported the sex, age, or education level of care recipients. Three studies reported both sex and education: the HIV+ patients in Chang et al. (2011) were 66% female, 34% male and had a median age of 34 years in intervention group and 36 years in control group; the mean age of the women in Ginsburg et al. (2014) was 38 years; patients in Gautham et al. (2015) were 35.7% male and 64.3% female 45.2%, 54.8% were greater than or

equal to 10 years old. Mean education level was not reported in any of the three studies. The mean age of the pregnant mothers in Omer et al. (2008) was not reported, but 45% reported receiving "some formal education" (p. 182). Two studies reported both age and education: patients in Winch et al. (2003) were children under 14 years whose carers had a mean 0.85 years of education; in Ashraf et al. (2015) patients were children under 5 years whose carers had a mean 4.2 years of education. The intervention evaluated by Andreoni et al. (2016) was directed at children under 5 years.

Excluded studies

Forty studies were excluded for the following reasons: 12 did not meet the criteria for study design (Ahmed, Desta, Tekle, & Mweta, 1993; Akogun et al., 2001; Alam et al., 2012b; Avula, Frongillo, Arabi, Sharma, & Schultink, 2011; Bhatt, Maru, & Prabhakar, 1981; Chin-Quee et al., 2016; Curtale, Siwakoti, Lagrosa, LaRaja, & Guerra, 1995; Elder et al., 1992; Findley et al., 2013; Harvey et al., 2008; Pineda, Bertrand, Santiso, & Guerra, 1983; Puett, Coates, Alderman, Sadruddin, & Sadler, 2012); 5 had a non-CHW comparator (Amouzou et al., 2014a; Baqui et al., 2008; Bertrand, Mangani, Mansilu, McBride, & Tharp, 1986; Bojang et al., 2011; Pradeep et al., 2014); 9 tested the efficacy of CHWs for delivering a health intervention, rather than efficacy of an intervention to improve CHW performance (Begum, Bhuiyan, & Asiruddin, 2009; Bhutta et al., 2008a; Boone et al., 2007; Davies-Adetugbo, Fabiyi, Ojoofeitimi, & Adetugbo, 1997; Jintaganont, Stoeckel, & Butaras, 1988; Jung et al., 2009; Morris, 2015; Owais, Hanif, Siddiqui, Agha, & Zaidi, 2011; Turab et al., 2014); 6 tested health workers who were primarily facility-based (Björkman & Svensson, 2009; Brieger, 2000; Florez-Arango, Iyengar, Dunn, & Zhang, 2011; Foreit & Foreit, 1984; Iyengar & Florez-Arango, 2013; Zurovac et al., 2011); 3 tested interventions

that were not designed to improve CHW performance (Kalyango et al., 2012; Sranacharoenpong & Hanning, 2012; Zafar et al., 2016); 1 evaluated an intervention for which the effects of the CHW performance piece could not be disaggregated (Das et al., 2014); 1 had performance interventions in both arms (Joos et al., 2016); and 3 used other professionals as health workers (Afenyadu, Agyepong, Barnish, & Adjei, 2005; Ashraf, Bandiera, & Jack, 2014; George et al., 2012). Studies plausibly expected to be among those included are listed in the Characteristics of excluded studies table in Appendix 6d, along with a reason for exclusion.

Ongoing studies

Information regarding ongoing studies (Hackett & Sellen, 2015; Kallander et al., 2012; Rabbani et al., 2014; Saha et al., 2013; Vedanthan et al., 2014) is provided in the Table of Ongoing Studies in Appendix 6d.

6.3.2 Risk of Bias in Included Studies

Assessments of the risk of bias for included studies are included in the Characteristics of included studies tables in Appendix 6d and are summarised in Fig 6.3 and Table 6.2 below. These assessments were used for interpreting the results and assessing the quality of evidence for specific outcomes. The first six categories apply to all included trials, while the latter four apply only to cluster-randomised control trials.

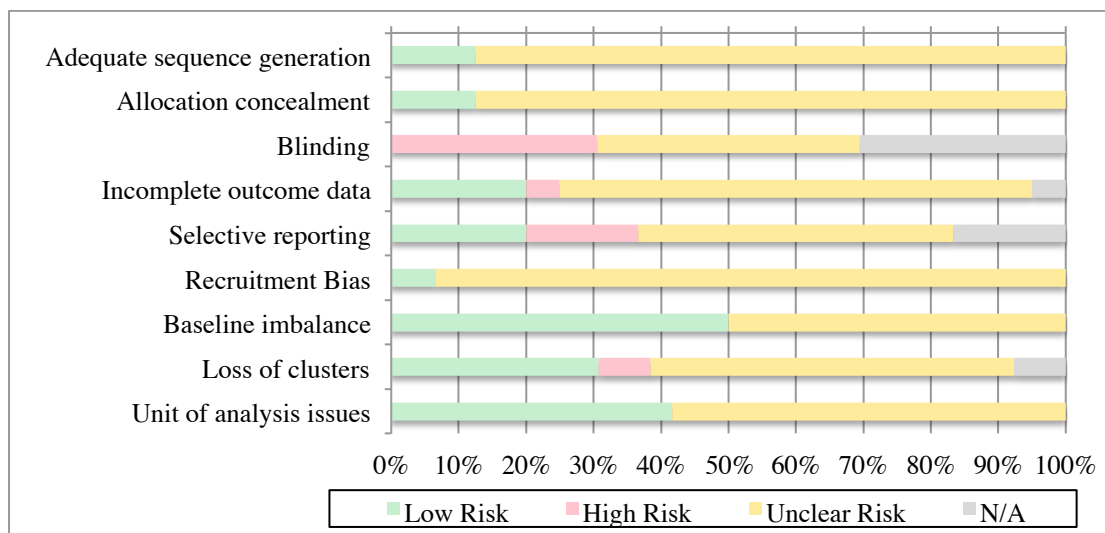


Figure 6.3 Review author's judgement about each methodological quality item presented as percentages across all included studies

	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Recruitment bias	Baseline imbalance	Loss of clusters	Unit of analysis issues
Andreoni (2016)	?	?	?	?	+	?				
Ashraf (2015)	?	?	?	?	-	?	+	?	?	+
Ayele (1993)	?	?	-	+	+	?	+	-	+	-
Bailey (1996)	?	?	-	?	+	?				
Bossuroy (2016)	?	?	-	+	+	?	+	+	?	+
Carrón (1994)	?	?	-	?	-	?	+	+	+	-
Chang (2011)	?	+	-	?	+	+	+	+	+	+
DeRenzi (2012A)	?	?	-	?	-	?				
DeRenzi (2012B)	?	?	-	?	-	?				
DeRenzi (2012C)	?	?	-	?	-	?				
Gautham (2015)	?	?	-	-	?	?				
Ginsburg (2014)	+	+	?	+	?	?				
Omer (1998)	+	?	-	+	+	?	+	?	+	?
Winch (2003)	?	?	-	?	?	?	+	+	+	+

Table 6.2 Methodological quality summary: review author's judgement about each methodological quality item for each included study

Random sequence generation and allocation concealment (selection bias)

Though all trials reported that they were randomised, only one two described the method of randomisation (Ginsburg et al., 2014; Omer et al., 2008). Likewise, only two trials reported a method of allocation concealment (Chang et al., 2011; Ginsburg et al., 2014). As such, the potential for selection bias is a notable weakness of this data set.

Blinding (performance bias and detection bias)

Given the time and human contact required for the interventions, the blinding of participants and intervention staff was not feasible. The impossibility of blinding intervention staff may have led to performance bias. Only three studies (Ayele et al., 1993; Ginsburg et al., 2014; Omer et al., 2008) reported the blinding of outcome assessors. While Chang et al. (2011) measured biological outcomes and Andreoni et al. (2016) used GPS-enabled phones meaning blinding of outcome assessors may not have been entirely crucial, the more subjective outcome collection tools used in the remaining studies (i.e. observational service tests, client questionnaires, household interviews, direct observations of CHW performance, and data collected by CHWs) make detection bias a notable risk.

Incomplete outcome data (attrition bias)

Final attrition rates for the group about which primary outcome data were gathered ranged from 0% (Ayele et al., 1993) to 30% (Derenzi et al., 2012C). Carrón et al. (1994) reported that an unspecified level of provider mortality unevenly affected the 180 CHWs randomised at baseline and attempted to compensate for this by randomly deleting data from arms of the trial with less attrition until group numbers were even. Not only does this process obscure the actual attrition rate, it still leaves the study at high risk of attrition bias: randomly deleting data is not the same as never

having collected it; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups. Three of the studies that collected data from patients did not identify those patients a priori or include data on CHW dropout so rates of attrition are unclear (Gautham et al., 2015; Ginsburg et al., 2014; Winch et al., 2003). One of the studies, however, specified that analysis nonetheless followed the intention to treat principle, in the sense that all women in the intervention communities were analysed as though exposed, whether or not CHWs had used the intervention materials while treating them (Omer et al., 2008).

Selective reporting (reporting bias)

The variety of measurements employed in the 14 trials suggests that no standard set of outcomes exists for the evaluation of CHW performance. Data were reported in several of formats; many were unsuitable for meta-analysis even had it been advisable. Repeated attempts were made to contact trialists for missing statistical data; while many authors were helpful in providing additional data, further information was often unavailable due to data loss or non-response. Of the trials, only two reported having a protocol (Andreoni et al., 2016; Chang et al., 2011). While Chang et al. (2011) reported all pre-specified outcomes, the outcomes listed in the protocol for the Andreoni et al. (2016) study were embargoed. As with Andreoni et al. (2016) the risk of selective reporting among the other eleven trials is thus unclear.

Publication bias

Despite an exhaustive literature search, it is possible that the findings are subject to publication bias. Funnel plots were not used to explore publication bias as there were too few studies to draw conclusions from such analysis.

C-RCTs: Recruitment bias

No C-RCT recruited individuals to clusters post-randomisation, so risk of recruitment bias across trials is low.

C-RCTs: Baseline imbalance

Three studies (Ashraf et al., 2015; Carrón et al., 1994; Omer et al., 2008) did not provide an explicit statement of group equivalence at baseline, though Carrón et al. (1994) pair-matched clusters to reduce the risk of such differences. Two studies reported baseline differences: Ayele et al., (1993) reported baseline differences in CHW age and years since graduation but did not control for them; despite pair-matching clusters. Winch et al. (2003) reported baseline differences in literacy, spouse's education, and carer's education and adjusted for economic status (as measured by presence of a metal roof), education, and literacy in a multivariate random effect logistic regression model. Chang et al. (2011) and Bossuroy et al. (2016) reported that there were no significant baseline differences among participants.

C-RCT: Loss of clusters

Ashraf et al. (2015) and Bossuroy et al. (2016) did not note whether clusters were lost; in the remaining five studies, no clusters were lost.

C-RCT: Unit of analysis issues

Only three studies (Ashraf et al., 2015; Bossuroy et al., 2016; Chang et al., 2011; Winch et al., 2003) explicitly took clustering into account in the analysis, the rest were unclear or did not.

C-RCTs: Comparability with nonrandomised trials

Because no meta-analysis was conducted, comparability with individually randomised trials wasn't relevant.

C-RCT Contamination

Contamination and threats to programme differentiation were unlikely given the geographical separation of the different clusters and the fact that the experimental and control interventions were delivered by different personnel.

6.3.3 Effects of Interventions

The protocol for this review specified that a meta-analysis would be conducted if appropriate. Upon reviewing the data, it was decided that meta-analysis would be uninformative. The 14 studies had dissimilar intervention designs, uptake, context, and outcomes (Montgomery et al., 2013b). This clinical heterogeneity made the programmes unsuitable for aggregation, since a meta-analysis of disparate interventions may have obscured genuine differences in effect and separating studies into sub-groups rendered groupings too small for informative meta-analysis. Another major obstacle to meta-analysis was the poor reporting of primary outcomes. Missing information and unit of analysis errors without reported intra-cluster correlations in several studies (detailed below) meant it would have been potentially misleading to calculate an average effect across studies (Effective Practice and Organisation of Care Group, 2014). Systematic error may have shifted the estimate of the intervention effect in either direction and meant the confidence intervals would have likely underestimated the true uncertainty around the effect. Meta-analysis of this type compounds errors, producing a ‘wrong’ result that may be interpreted as having more credibility (Higgins & Green, 2011; Littell, Corcoran, & Pillai, 2008). In lieu of quantitative synthesis, a structured summary is provided (Popay et al., 2006).

Studies are more easily grouped by their outcomes than the interventions examined. Five studies examined the review’s primary outcomes (Chang et al., 2011; Derenzi et al., 2012BC; Omer et al., 2008; Winch et al., 2003) and nine studies

examined secondary outcomes (Andreoni et al., 2016; Ashraf et al., 2015; Ayele et al., 1993; Bailey & Coombs, 1996; Carrón et al., 1994; Derenzi et al., 2012A; Gautham et al., 2015; Ginsburg et al., 2014; Omer et al., 2008).

Primary outcomes

The primary outcomes specified in the protocol were objective physical outcomes for patients and subjective health outcomes for patients (Table 6.3).

Biological outcomes for patients

One study reported biological outcomes for patients: in their trial of an integration-equipment intervention in Uganda, Chang et al. (2011) found no significant differences in HIV patients' cumulative risk of virologic failure (failure defined as 0.400 copies/mL; RR = 1.17, 95% CI 0.84 to 1.64, P = 0.34) or mortality (RR = 0.82, 95% CI 0.55 to 1.22, P = 0.33). This evidence of no effect was graded high (no reason for downgrade).

Behavioural outcomes for patients

Six studies reported behavioural outcomes for patients. In their trial of incentives in India, Bossuroy et al. (2016) found no significant difference in patient TB default rate (B = -0.04, SE = 0.06, P > 0.05), though CHWs were incentivised based on treatment adherence.

In their trial of a integration-equipment intervention in Uganda, Chang et al. (2011) found no significant differences in HIV patients' <95% pill count adherence (RR = 1.01, 95% CI 0.97 to 1.06, P = 0.59) or loss to follow-up (defined as lack of a pharmacy refill visit in over 90 days, RR = 1.29, 95% CI 0.50 to 13.32, P = 0.60).

Table 6.3 Primary outcomes

Primary outcome measure	Study (N)	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Biological outcomes for patients	Chang 2011 (N=970 patients)	Reporting-Supervision	Virologic failure	RR = 1.17 (0.84–1.64) P = 0.34
			Mortality	RR = 0.82 (0.55–1.22) P = 0.33
	Bossuroy 2016 (N = 2500 patients)	Incentives	TB default rate	B = -0.04 (SE = 0.06) P > 0.05
Behavioural outcomes for patients	Chang 2011 (N=970 patients)	Reporting-Supervision	<95% pill adherence	RR = 1.01 (0.97–1.06) P = 0.59
			Loss to follow-up	RR = 1.29 (0.50–13.32) P = 0.60
	DeRenzi 2012B (N=un-specified patients)	Supervision	Mean days clients overdue	U = 271.00 P < 0.001 r = 0.500
	DeRenzi 2012C (N=un-specified patients)	Supervision	Mean days clients overdue	See text
	Omer 2008 (N=1070 patients)	Integration-Training	Attend at least one prenatal visit	RR = 1.94 (1.56-2.41) P < 0.0001
			Colostrum to newborns	RR = 1.21 (1.12-1.32) P < 0.0001
			Stop routine heavy work	See text
			Exclusive breastfeeding	See text
	Winch 2003 (N=286 patients)	Training-Reporting-Integration	Proportion of malarial children treated by mothers exactly per IMCI norms	RR = 27.76 (0.53-1441.23) P < 0.0001

In their trial of an integration-training intervention comprised of an indigenous, evidence-based communication tool in Pakistan, Omer et al. (2008) found that pregnant women in intervention communities were significantly more likely to attend at least one prenatal check-up (RR = 1.94, 95% CI 1.56 to 2.41, $P < 0.0001$) and give colostrum to their newborn babies (RR = 1.21, 95% CI 1.12 to 1.32, $P < 0.0001$). The trial also reported that the odds of pregnant women stopping routine heavy work during pregnancy were no different in the intervention and control communities and that the odds of pregnant women maintaining exclusive breastfeeding for four months were higher in the intervention than control group, but provided no quantitative data for either outcome. A woman advised by a CHW to reduce routine heavy work, however, was significantly more likely to adhere to the advice than someone who was not so advised or who was advised by others (RR = 1.31, 95% CI 1.20 to 1.422, $P < 0.0001$).

In their trial of an integration-training-reporting intervention comprised of visual aids, refresher training, a counter referral recordkeeping notebook and a community meeting in Mali, Winch et al. (2003) found that the proportion of malarial children who were treated by their mothers, as instructed by CHWs, with a full course of chloroquine in the home exactly per IMCI norms (correct dose, 1 dose per day, administered for the full 3 days) was significantly higher in the intervention group than the control group (RR = 27.76, 95% CI 0.53 to 1441.23, $P < 0.0001$).

In their trial of a supervision intervention in Tanzania, DeRenzi et al. (2012B) found SMS+Supervision (an escalating text message reminder system plus follow-up for underperforming CHWs) to be efficacious compared to usual care for reducing the mean number of days a CHW's chronic care clients were overdue. In DeRenzi et al. (2012C), CHWs were re-randomised to SMS+Supervision or SMS-only interventions.

For SMS-only CHWs who had been in the SMS+ Supervision group during DeRenzi et al. (2012B), removing the follow-up caused performance to worsen relative to those who continued to receive SMS+Supervision ($U = 68.00$, $P = 0.023$, $r = 0.340$). For CHWs who had been in the control group during DeRenzi et al. (2012B), performance improved; however, no statistically significant difference was found between the SMS+ Supervision and SMS-only groups ($U = 101.00$, $P = 0.880$, $r = 0.043$).

Overall, these studies suggest that CHW performance improvement interventions can improve certain behavioural outcomes for patients. The body of evidence for behavioural outcomes for patients was graded moderate (downgraded for design and implementation as allocation concealment and random sequence generation were unclear in the majority of studies).

Secondary outcomes

Secondary outcomes were utilisation of health services, quality of care provided by CHWs, and CHW retention (Table 6.4).

Table 6.4 Secondary outcomes

Secondary outcome measure	Study	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Utilisation of health services	Ashraf 2015 (N = specified in text)	Recruitment	Utilisation of health services	See text
	DeRenzi 2012A (N = 11 CHWs)	Supervision	Closed referrals	SMD = 0.86 (-0.30, 2.03)
	Ginsburg 2014 (N=556 patients)	Equipment (I _a) /Training-Equipment (I _b)	Closed referrals	I _a : 43% I _b : 63% C: 53% P >0.05

Secondary outcome measure	Study	Intervention Type	Variables	Measure of Effect (95% CI or SE)
<i>Continued from previous page</i>				
Utilisation of health services (continued)	Winch 2003 (N=79 patients)	Training-Reporting-Integration	Closed referrals	RR = 0.78 (0.57, 1.07) P = 0.13
	Andreoni 2016 (N = 337 CHWs)	Incentives	Closeness to vaccine distribution policy objective	SMD = - 0.32 (-0.54, -0.1)
CHW quality of care	Ashraf 2015 (N= 298)	Recruitment	Home visits	B = 93.95 (SE = 37.19) P < 0.05
			Patients seen	B = 31.79(SE = 260.4) P >0.05
			Comm. mtgs. organised	B = 17.06 (SE = 0.5.22) P < 0.01
	Ayele 1993 (N=100 CHWs)	Training-Supervision	Completion of government-set CHW activities	See text
	Bailey 1996 (N = 49)	Training	Knowledge score	<u>Combined</u> Intervention: 36.8% Control: 24.9% P > 0.05
	Bossuroy 2016 (N = 78 CHWs)	Incentives	TB case detection rate	See text
	Carrón 1994 (N=180 CHWs)	Training-Supervision	Quality of care score	See text
	Gautham 2015 (N= 16 CHW)	Equipment	Protocol compliance	See text
	Winch 2003 (N=102 CHWs)	Training-Reporting-Integration	Appropriate drug administration	RR = 1.45 (0.99, 2.13) P = 0.04
			Dispensing correct amount of medication	OR = 1.5 (0.52-4.33) P = 0.45

Utilisation of health services

Four studies reported on utilisation; three examined the effect of CHW performance interventions on closed referrals and one on utilisation more generally. Three studies reported on closed referrals (Derenzi et al., 2012A; Ginsburg et al., 2014; Winch et al., 2003). In their trial of a training-reporting-integration intervention in Mali, Winch et al. (2003) found no difference in compliance with referral to the health centre between intervention or control patients (RR = 0.78, 95% 0.57 to 1.07, $P = 0.13$) but that patients in the intervention group were nearly four times more likely to be referred than those in the control group (RR=3.76, 95% CI 2.25 to 6.28, $P < 0.0001$). In their trial of an equipment/training-equipment intervention in Bangladesh, Ginsburg et al. (2014) found similar results: while women in intervention arm B (smart phones plus patient navigation) were significantly more likely to attend for care versus women in intervention arm A (smart phones without navigation; 63% vs. 43%, $P < 0.0001$), adherence in the two intervention arms combined was the same (53%) as that of the control arm. Over the 4-month accrual period, however, control arm CHWs interviewed fewer women than CHWs in either of the two intervention arms and the proportion of study participants found to have an abnormal clinical breast exam was more than 3 times greater in the intervention A (3.1%) and intervention B (3.2%) arms than the control arm (1.0%) ($P < 0.0001$).

DeRenzi et al. (2012A), by contrast, found a large but statistically insignificant increase in the mean percentage of closed referrals in the intervention group (SMD = 0.86, 95% CI -0.30, 2.03). This result should be interpreted with caution, however, as a methodological limitation may explain much of the difference: all participants, not just those in the intervention group, were told at the start of the trial that they would receive SMS reminders when they had pending referrals to close

but only the intervention group received such reminders. This instruction may have negatively influenced the control group's performance by setting a false expectation.

Ashraf et al. (2015) examined utilisation of mother and child health services in the facilities where CHWs operate. Difference-in-difference estimates based on the comparison of intervention and control areas before and after CHWs started working reveal that treatment increased the number of women giving birth at the health centre by 31% ($N = 1269$ patients, $B = 14.68$, $SE = 6.32$, $P < 0.05$), the number of children under 5 undergoing health checks by 24% ($N = 1618$ patients, $B = 318.1$, $se = 98.05$, $P < 0.05$), being weighed by 23% ($N = 1610$ patients, $B = 284.3$, $SE = 110.2$, $P < 0.05$) and number of children under 1 receiving immunization against polio by 20% ($N = 1530$ patients, $B = 14.98$, $SE = 4.803$, $P < 0.05$). No statistically significant difference was seen in the number of postnatal (0-6 week) visits, children under 1 year receiving BCG vaccinations, and children under 1 receiving measles vaccinations.

Overall, these studies suggest that CHW performance improvement interventions can improve utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services. The body of evidence for utilisation of services was graded moderate (downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors as unclear in the majority of studies).

CHW quality of care

Eight studies reported on CHW quality of care using a range of measures (Andreoni et al., 2016; Ashraf et al., 2015; Ayele et al., 1993; Bailey & Coombs, 1996; Bossuroy et al., 2016; Carrón et al., 1994; Gautham et al., 2015; Winch et al.,

2003). Four trials reported on CHWs completing (Ashraf et al., 2015; Ayele et al., 1993; Bossuroy et al., 2016) or having the knowledge necessary to complete (Bailey & Coombs, 1996) prescribed activities. In their trial of an training-supervision intervention comprised of (a) a refresher course and (b) monthly supervision in Ethiopia, Ayele et al. (1993) calculated the scores for the 13 CHW government-set activities:

- | | |
|---|---------------------------------------|
| 1. Outreach (days/mo.) | 7. Deaths (# registered/mo.) |
| 2. Health education (sessions/mo.) | 8. Curative (pts./mo.) |
| 3. Environmental health (activities/mo.) | 9. Reports (#/mo.) |
| 4. Maternal & child health
(activities/mo.) | 10. School health (activities/mo.) |
| 5. Expanded programme on
immunisation (activities/mo.) | 11. Home visits (#/mo.) |
| 6. Births (# reg./mo.) | 12. Referrals (#/3mo) |
| | 13. Epidemic control measures (#/3mo) |

They found that 10 of 13 activity scores (similar at baseline, maximum score not stated) were significantly higher ($P < 0.05$) among intervention group CHWs at three and six months following the onset of the trial and that the composite functional status score was also significantly higher among intervention group CHWs ($P < 0.0001$). Though the composite functional status score was defined in the paper as the sum of the 13 activity scores, it is not; the authors of the paper were contacted but were unable to supply clarifying data.

In their trial of a recruitment intervention in Zambia, Ashraf et al. (2015) found that CHWs recruited with career possibilities salient conduct 29% more household visits ($B = 93.95$, $SE = 37.19$, $P < 0.05$) and organise over twice as many community meetings (43 vs. 22, $B = 17.06$, $SE = 0.5.220$, $P < 0.01$) than those recruited with community benefits salient, while the difference in the number of

patients seen at the health post is also positive but not significant ($B = 31.79$, $SE = 260.4$, $P > 0.05$).

In their trial of an incentives intervention in India, Bossuroy et al. (2016) found that, on average, the number of new TB detections increased by 2.18 (33.2%) each month among intervention CHWs during the period said CHWs were incentivised based on case detection ($B = 2.18$, $SE = 0.95$, $P < 0.05$). The number of defaults, however, was significantly larger among intervention CHWs during this period ($B = 0.08$, $SE = 0.04$, $P < 0.05$).

In their trial of a training intervention in Guatemala, Bailey & Coombs (1996) found that CHWs in the intervention group demonstrated a significantly greater ability to correctly diagnose diarrhoea of varying type and severity (Mean intervention score: 77.3%, mean control score: 43.1%, $P < 0.05$). Though it is worth considering the clinical salience of this result given that only non-significant differences favouring the intervention group were found in the CHWs' general knowledge regarding diarrhoea, referral recommendations, and treatment practices.

Two studies reported on treatment protocol compliance (Gautham et al., 2015; Winch et al., 2003). Winch et al. (2003) found that intervention CHWs were significantly more likely to perform all three tasks necessary for appropriate drug administration (telling the carer (i) the dose to give the child each time and (ii) the number of days of treatment, and showing the carer (iii) how to administer the drug) ($RR = 1.45$, 95% CI 0.99 to 2.13, $P = 0.04$). The study also reported that more intervention CHWs (58.9%) sold the correct amount of medication than control CHWs (48.1%). Using the ICC provided (0.210), a risk ratio was approximated as per the process described in the methods section ($RR = 1.22$, 95% CI = 0.69 to 2.18, $P = 0.47$). The p-value calculated by Winch ($P = 0.042$) indicates statistical significance;

three emails attempting to resolve the discrepancy received no response.

In their trial of an equipment intervention in India, Gautham et al. (2015) calculated mean protocol adherence for the intervention and control groups across a variety of stratifications of interest. Controlling for other predictors, the intervention group's protocol compliance exceeded that of control RHPs (+26.8 percentage points, $P < 0.05$) for initial male patients. The intervention group exhibited greater protocol compliance across patient order and across sexes than the control group. The initially high protocol compliance superiority for both male and female patients of the intervention group diminished over time. Gautham et al. (2015) were contacted for the unstratified mean protocol compliance for the intervention and control group over the entire study but were not able to provide it.

Two trials reported on observed differences in care provided (Andreoni et al., 2016; Carrón et al., 1994). In their trial of an training-supervision intervention comprised of a service algorithm and a novel supervision feedback instrument in Paraguay, Carrón et al. (1994) found quality of care provided (measured by an observational service test) improved in all four groups and the pre-test-post-test differences were both statistically ($P < 0.05$) and clinically significant, representing an increase in quality of 25-40%. The largest quality gain was observed in CHWs who received ABC, though differences between groups were not statistically significant. Results were presented in a graph without measures of variability so it is impossible to report numerical values for each outcome. Quality was also assessed using a client questionnaire measuring information provision, interpersonal relations, provider technical competence, and follow-up mechanisms. None of the comparisons between groups yielded statistically significant differences in ANOVA tests. Raw gains were presented in a non-numerical table in which a plus sign (+) indicated a study group

that improved more or showed a smaller quality loss than the control group. It is thus impossible to report numerical values for each outcome.

In their trial of tailored incentives in Pakistan, Andreoni et al. (2016) found that, when outliers were excluded, CHWs who receive tailored contracts were a small to moderate amount closer to the policy objective than their untailored counterparts (SMD = -0.32, 95% CI -0.54 to -0.1). There were differential effects of tailoring across the advance and day-of allocation groups. Untailored day-of choice was associated with significantly larger distance to the policy objective measures. Tailoring for day-of choice significantly reduced these distances: when outliers were excluded, tailoring in day-of choice reduced distance from equal provision by around one-half ($B = 0.130$, $SE = 0.037$, $P < 0.01$). Tailoring in advance choice directionally reduced distance, but the effect was not significant, potentially due to the relatively small average distance measure identified in untailored advance choice. Taken together, outcomes in Andreoni et al. (2016) suggest that for a given interest rate, matching this interest rate to individual preferences generates significantly smoother service provision, reducing distance measures by around one-third on average and around one-half for day-of allocation.

Overall, these studies suggest that CHW performance interventions can improve CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided. The body of evidence for CHW quality of care was graded moderate (downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors as unclear in the majority of studies).

No studies reported on CHW retention or adverse events.

Additional Outcomes

Carrón et al. (1994), Chang et al. (2011), and Derenzi et al. (2012) reported cost outcomes. Though cost was not prespecified in the review protocol as a variable about which information would be collected, it is important in assessing the relative merits of various combinations of inputs to improve CHW performance and so it is reported here (Walker & Jan, 2005). Because the data on resource use and costs were very heterogeneous, meta-analysis was not appropriate and available data are presented in Table 6.5.

The economics component of this review was conducted according to current guidance on the use of economics methods in the preparation and maintenance of Cochrane reviews (Higgins & Green, 2011). The reviewer classified the included economic evaluations based on an established system (Drummond et al., 2005). Briefly, full economic evaluations (e.g. cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis) compare the costs (resource use) associated with one or more alternative interventions (e.g. intervention X versus comparator Y) with their consequences (outcomes, effects) whereas partial economic evaluations (e.g. cost analyses, cost-description studies, and cost-outcome descriptions) do not make explicit comparisons between alternative interventions in terms of both costs (resource use) and consequences (effects) (Higgins & Green, 2011).

Table 6.5 Summary of cost data

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study pop.	Intervention vs. Comparator	Economic results (Currency & price year)
Carrón et al. 1994	CA (Provider, 1 year)	CHWs	I _a = Decision tree I ^b = Behavioural feedback instrument I _{a+b} = I _a + I ^b C = Usual care	I _a was the least expensive (29.42 USD/CHW) followed by I ^b (78.45 USD/CHW), the I _{a+b} (106.80 USD/CHW), and C, the traditional training course (166.17 USD/CHW in groups of twenty-five, 118.40 USD/CHW in groups of fifteen or fewer) (1993 USD)
Chang et al. 2011	CA (Provider, 2 years)	HIV positive adults	I = A text message reporting system and health centre warmline C = Usual care	The total cost of the mHealth intervention was \$2,353.24. The yearly cost was \$1045.88, resulting in a yearly additional cost per patient of \$2.35. (2006-08 US)
DeRenzi et al., 2012B	CA (Provider, 1 year)	Chronic care patients	I = Escalating text message reminder system & follow-up by supervisor C = Usual care	The intervention adds an estimated \$0.84 per client per year of running the ICT4CHW programme with automated escalating reminders (Year unspecified USD)

CA: Cost analysis; **USD:** US dollars.

These studies generally highlight the low incremental costs or potential cost savings of these CHW performance interventions. The quality of the economic evaluations was assessed using the Cochrane-recommended CHEC list (Evers et al., 2005), a consensus criteria list that incorporates items from Drummond and Jefferson (1996) and other systematically-identified guidelines. These assessments are reported in a table in Appendix 6d. Because CHEC was designed to assess full economic evaluations, items not applicable to the partial economic evaluations included are

shaded grey. There was some risk of bias across the studies: not all relevant alternatives and costs (e.g. training, supervision) were reported in all studies, studies did not perform discounting, and no studies adequately addressed ethical and distributional implications.

6.4 Discussion

6.4.1 Summary of Main Results

This review identified 14 trials evaluating the effects of interventions for improving the performance of community health workers for community-based primary healthcare in low- and middle-income countries. The heterogeneity and poor reporting of included studies precluded meta-analysis.

A structured summary laid out moderate quality evidence of the efficacy of CHW performance interventions in (i) improving certain behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided (e.g. bedside manner, technical competence, achieving policy objectives). The evidence for biological outcomes was graded high, though the lack of effect on virologic failure and mortality may not be indicative of all possible biological outcomes. None of the studies included in this review reported on possible adverse effects of these interventions, either to patients, CHWs, or the health system. Evidence is insufficient to draw conclusions regarding the effects of such interventions on retention or to specify which performance improvement intervention strategies are likely to be most effective. Furthermore, the nearly half of studies were

compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components.

That said, there is moderate quality evidence that the following practices improve behavioural outcomes for patients, utilisation of services, and/or CHW quality of care: (1) When recruiting CHWs, emphasising career possibilities rather than benefits to the community (Ashraf et al., 2015). (2) When supervising CHWs, providing escalating reminders for tasks that are overdue and following-up with underperforming CHWs (Derenzi et al., 2012ABC). (3) When incentivising CHWs, tailoring incentives to measured individual preferences—but only for CHWs performing a single repetitive task, not for CHWs who must perform multiple or more complex tasks (Andreoni et al., 2016, Bossuroy et al., 2016). (4) When equipping CHWs, using mobile phone-based procedural guidance applications (Gautham et al., 2015; Ginsburg et al., 2014).

While it is important to recognise that variations in recruitment, training, supervision, incentivisation, integration, and equipment can affect outcomes, the four strategies listed above ought to be interpreted with caution. As explained in Chapter 1, CHW programmes can be conceptualised as complex interventions that unfold in complex systems; intervention effects may not be simply the sum of their components but emergent trajectories that result from dynamic processes—including interaction with the context in which they take place and the way in which they were implemented (Craig, Dieppe, Macintyre, & Michie, 2008; Hawe et al., 2009). As a consequence, it is important to critically interrogate the internal and external validity of the studies that produced these results. This will be done in the following two sections.

6.4.2 Internal validity

Methodological quality of the 14 included studies was difficult to assess due to incomplete reporting of key methodological and clinical features; contacting the trialists did not always yield sufficient additional information to make such judgments. As such, reporting deficiencies may obscure important strengths or weaknesses in the trials.

The selection and validity of outcome measures is also a concern across studies. For example, discordances between qualitative and quantitative evidence in Chang et al., (2011) (i.e. high enthusiasm for the use of mobile phones among CHWs and clinic staff in contrast to lack of significant quantitative benefits), may indicate either that “flashy” technology was perceived positively by participants despite not having quantifiable improvements or, alternatively, that other metrics (e.g. access to care) may have been worth measuring in addition to more downstream events such as mortality. Additionally, there were a wide range of measurement instruments employed to assess “quality of care.” Some were more valid than others: the “functional status score” employed in Ayele et al. (1993), for example, measures activity, not quality. While CHWs must perform their assigned functions as a prerequisite to providing quality care, the two are not the same and should not be confused. Similarly, while Winch et al. (2003) and Omer et al. (2008) relied on patient report to assess the impact of CHWs, more rigorous, direct assessments of performance—like the service tests used by Carrón et al. (1994) and Gautham et al. (2015)—paired with biological measures would have been preferable.

Missing data, unit of analysis, and power problems compounded these issues, making the size (Omer et al., 2008) and precision (Ayele et al., 1993; Carrón et al., 1994; Winch et al., 2003) of the treatment effect difficult to ascertain for several outcomes. Multiple studies (Derenzi et al., 2012ABC; Gautham et al., 2015) were

underpowered to detect meaningful clinical difference.

Finally, it seems reasonable to consider that most of the interventions produced a considerable Hawthorne effect. A number of studies documented that existing CHWs were under-supported in their work, and the intervention is likely to have been a motivational factor (Derenzi et al., 2012ABC; Gautham et al., 2015). In Chang et al. (2011), the only study to demonstrate a null effect for its primary outcome, attention could have worked the opposite way: by virtue of the mobile phones, patient care in the intervention arm may have become easier for clinic staff and resulted in a shift in their efforts toward the control arm (i.e. indirect contamination). Future studies would do well to adjust for the attention effect by, for instance, ensuring some type of regular, non-supportive contact between CHWs and health centre staff, supervisors, or trainers or—as Gautham et al. (2015) did—introducing “placebo” technology to the control CHWs.

6.4.3 Overall Completeness and Applicability (External Validity) of Evidence

The objective of this review was to determine the effectiveness of interventions to improve the performance of CHWs for community-based primary healthcare in low- and middle-income countries. The evidence gathered by this search is somewhat able to address this objective, but several factors limit the applicability of these results (Schünemann et al., 2006).

Diverse outcomes across studies meant there were few data for each outcome of interest, including cost. Both the CHWs and the patient population represented are diverse (e.g. malarial children, HIV patients, pregnant mothers) and insufficient information was provided on relevant socio-demographic characteristics and baseline biologic variation to assess whether populations in a proposed implementation setting

might be comparable. Because no studies reported adverse effects it is difficult to draw conclusions regarding the trade-off between benefits and harms (Lavis et al., 2009a).

Similarly, trials often failed to report relevant information regarding context. This is true both of the cultural context in which CHWs operate and CHW programme characteristics (e.g. recruitment, incentivisation etc.) that remained unmodified in the interventions. While it is possible that the different components of a given CHW delivered intervention work independently, it is likewise possible that, as with most complex interventions, they interact with contextual factors and each other (Craig et al., 2008; Montgomery et al., 2013b). Using a systems thinking perspective (see Chapter 1), Shelley (2015) identified several potential feedback loops in a hypothetical CHW system. The dynamism and complexity of CHW systems is reflected in the results of Ginsburg et al. (2014): women in intervention arm B (smart phones plus patient navigation) were significantly more likely to attend for care versus women in intervention arm A (smart phones without navigation; 63% vs. 43%, $P < 0.0001$), suggesting that equipment may not be as valuable in the absence of adequate training on bedside manner. Additional information on baseline programme characteristics would have provided greater assistance in understanding other threshold effects, if any, of assessed interventions for CHW performance. Without better reporting of culture and context it is difficult to assess applicability of these strategies elsewhere.

As explained in Chapter 4, both Lewin et al. (2010) and Lavis et al. (2009) have delineated factors about a potential implementation setting that ought to be considered by decision-makers when assessing whether CHW programmes are likely to be successful. These are likewise useful in aiding with the interpretation of

potential program-level performance improvement interventions:

1. Whether there are important differences in health system arrangements that may alter the feasibility and acceptability of the given intervention. For example, in Zambia, the country in which the Ashraf et al. (2015) recruitment intervention took place, CHWs are part of the civil service. As a result, they are eligible for government sponsored “in-service training” that allows them to advance to higher-ranked and better paid cadres. In countries where CHW programmes are less well resourced and similar opportunities do not exist, it would not be possible to emphasise the potential for career advancement at the recruitment stage. Resource limitations may also inhibit the time-consuming process of estimating CHW preferences and tailoring incentives as described in Andreoni et al. (2016). Indeed, as noted in Chapter 4, there exist many programmes where CHWs receive no incentives at all (Jaffar et al., 2009; Newell et al., 2006; Pachón et al., 2002).
2. Whether there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting. Equipping CHWs with mobile phone-based procedural guidance applications as described in Gautham et al. (2015) and Ginsburg et al. (2014) would not be possible, for instance, in remote areas without mobile phone reception. Similarly, it would not be possible to have a supervision system premised on following-up with underperforming CHWs as described in Derenzi et al. (2012ABC), without a system to track CHW performance in the first place.
3. Whether there are important differences in the baseline health conditions between where the studies were done and the implementation setting. For

example, the tailoring of incentives described in Andreoni et al. (2016) worked well for a programme in which CHWs had a single categorical objective (vaccinate every child under 5 years in a given area) that only required one repetitive behaviour to achieve (provide oral polio vaccine). Where CHWs have more complex objectives (e.g. provide primary healthcare to a community) that require multiple behaviours (e.g. provide preventative care, curative care, conduct referrals etc.) performance-based incentives would be more difficult to calibrate and may distort behaviour in unintended and undesirable ways (Benabou & Tirole, 2003). This was seen in the results of Bossuroy et al. (2016) where, in a multitasking environment, performance-based incentives had undesired impacts on non-incentivised outcomes (i.e. TB defaulting increased for patients of CHWs incentivised for case detection compared with patients of CHWs receiving a fixed salary).

4. Whether studies from which evidence was drawn were conducted in settings with similar social norms. Information about, for example, whether male CHWs in the Ayele et al. (1993) study faced cultural barriers working with women in the home etc. are examples of norms and attitudes that might have shaped intervention effects and so would be relevant to interpreting trial outcomes or exploring differential effects across settings.
5. Whether sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time. As explained above, the extent to which these proposed performance improvement interventions depend on the presence of other programmatic or contextual preconditions is not clear from the included studies.

Conducting process evaluations would have helped explain the success, failure, or unexpected consequences of a trial; illuminate causal mechanisms of a successful intervention as well as avenues for optimisation; evaluate fidelity and quality of implementation; and reveal the degree to which the treatment can be standardised (Linnan & Steckler, 2002; Oakley et al., 2006). The identification of core components is particularly relevant because so many trials tested compound interventions; generating information on the relationship between outcomes and individual intervention components, as suggested by MOST (Collins et al., 2011), would help those designing CHW programmes (Craig et al., 2008; Glasziou et al., 2008; Glenton et al., 2011a; Shepperd et al., 2009).

The review does not assess the long-term effects of interventions to improve CHW performance. Thirteen of the fourteen trials lasted less than 18 months; the one trial that lasted longer (25 months) found no effect on biological outcomes for patients. While providing additional inputs or supports to underperforming programmes may create short-term gains in patient health or CHW functioning, the novelty—and in turn the effects—of such inputs may wear off over time (Derenzi et al., 2012ABC). Moreover, trials may involve higher levels of support for the various inputs (e.g. supervision, refresher trainings) than may be available in non-experimental settings; further research is required on how high quality, sustainable operationalisation of these inputs is feasible at scale in low income settings.

Though the included programme designs were largely consistent with the taxonomy laid out in the background of this review, they do not represent the entire realm of approaches to improving CHW performance in low- and middle-income countries. Certain programme features (e.g. recruitment, incentivisation) were only assessed in one trial, belying the multiple ways such elements might be varied (e.g.

intensity, frequency, type etc.). Additionally, as mentioned in the introduction to this chapter, there may be entirely different approaches to program-level performance improvement not captured in the taxonomy or the included studies.

6.4.4 Quality of the Evidence

In the GRADE system, well-conducted randomised controlled trials are initially categorised as high quality and observational studies are initially categorised as low quality. The quality of the evidence can subsequently be rated down based on five criteria (risk of bias, inconsistency of effect, indirectness, imprecision, publication bias) or rated up based on three criteria (large effect size, dose-response gradient, plausible confounding) (Guyatt et al., 2008b).

While all studies included in this review were randomised controlled trials, due to weaknesses in design and implementation, the evidence for behavioural outcomes for patients, CHW quality of care, and utilisation of health services was graded moderate. The evidence for biological outcomes was graded high (no reason to downgrade Chang et al., 2011). No studies reported on CHW retention or adverse effects.

These assessments are made with the caveat that—as explained in Chapter 4—there are issues with the applicability, reproducibility, and clarity of GRADE in the field of public health (Langendam, 2012; Rehfuss & Akl, 2013). Where the results of a systematic review are highly heterogeneous and effect estimates rely on single studies or cannot be pooled, as in this review, applying GRADE criteria is challenging. Even the selection of the appropriate body of evidence can be difficult to do systematically and transparently; in order to make reviews policy-relevant heterogeneous measures are, as in this case, sometimes grouped under multifaceted “umbrella outcomes” (Rehfuss & Akl, 2013, p. 6). Finally, GRADE does not

integrate specifications that are particularly salient for assessing the effectiveness of complex interventions, for example, intervention delivery, economics, context and equity (Waters et al., 2011). Because Summary of Findings Tables are not well suited to showcase narrative measures of impact, one was not included.

6.4.5 Potential Biases in the Review Process

Firstly, CHW RCTs remain poorly indexed in electronic databases (Lewin et al., 2010). As noted in Chapter 4, a large number of terms are used in the literature to describe CHWs; while nearly 40 synonyms for CHW were used in the electronic search for this review, two of the local terms for CHWs used in the included studies (community-based distribution providers and *gérant/village* drug kit managers) were not among them (Grayson & Gomersall, 2003). Similarly, though no language restrictions were imposed on the search and several non-English terms for CHW were included in the electronic search strategy, few non-English studies were found. It is therefore possible that, despite the extensive searches undertaken, some relevant RCTs were not identified.

Secondly, as has been noted elsewhere (Lehmann & Sanders, 2007; Lewin et al., 2010; Love et al., 1997; Witmer, Seifer, Finocchio, Leslie, & Oneil, 1995) there is no widely accepted definition for this cadre of health worker. Moreover, as many studies do not adequately describe the characteristics of included health workers, it is difficult to apply any definition. As such, exclusions may be debated.

Thirdly, despite an extensive search for unpublished and ongoing trials (which included searching grey literature, handsearching conference abstracts, and contacting experts), the review remains vulnerable to publication bias. The studies located may be unrepresentative of all unpublished studies.

Fourthly, limitations of the search strategy aside, poor reporting and nonresponse meant the reviewer was unable to obtain relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some studies.

Finally, while randomised trials have a unique value in component selection experiments, the exclusion of all non-randomised studies without further consideration places a zero weighting on evidence from such studies. Updates of this systematic review less constrained by the financial and temporal limitations of a DPhil should consider including such evidence.

To ensure quality in conduct and reporting, the author followed the Methodological Expectations of Cochrane Intervention Reviews (MECIR) and PRISMA statement for reporting systematic reviews (see Appendix 6c).

6.4.6 Agreements and Disagreements With Other Studies or Reviews

In an effort to determine what types of healthcare provision can be safely task shifted, a number of reviews have been carried out on CHW effectiveness, as noted in the introduction to the chapter. Far fewer reviews focus on CHW performance optimisation.

Two systematic review (Braun, Catalani, Wimbush, & Israelski, 2013; Kok et al., 2014) and four narrative reviews have examined some aspect of performance (Bhattacharyya, 2001; Bhutta et al., 2010; Jaskiewicz & Tulenko, 2012; Kallander et al., 2013). Jaskiewicz and Tulenko (2012) focus on a subset of performance *outcomes*, analysing programme elements that influence CHW productivity but finding little primary evidence. Braun et al. (2013), Kallander et al. (2013) and Bhattacharyya (2001) focus on a subset of performance *interventions*. Braun et al. (2013) and Kallander et al. (2013) analysed mobile health interventions, both finding

randomised and observational evidence that mobile tools help CHWs improve the quality of care but noting that stronger evidence is needed. Bhattacharyya (2001) analysed the effects of incentives, finding observational evidence that monetary, non-monetary, and community-level factors can motivate CHWs. An ongoing systematic review (Daniels et al., 2014) will also examine incentives. Bhutta et al. (2010) conducted a review of CHW programmes with positive impact on Millennium Development Goals (MDGs); while they derived draft recommendations about recruitment, training and supervision these were based on post-hoc correlations.

The completed reviews concurred that: (1) while there is a burgeoning literature on CHW performance, much of it consists of observational studies or documentation of project rollout. There is scant high-quality empirical evidence regarding the exact degree to which one input or a combination of inputs has a larger or smaller influence on CHW performance. (2) Of high-quality studies that exist, few evaluated programmes at scale or followed an a-priori experimental design. (3) Most studies neglected to document the complete description and characteristics of CHWs deployed.

Between the original search and the update of this review, a systematic review was published with the aim of identifying intervention design factors that influence CHW performance in LMICs (Kok et al., 2015b). The review included any studies “that concerned CHWs working in promotional, preventive or curative primary health care in [LMICs]...[that] described a factor related to intervention design” published within a 5 year period (2007-2013) (Kok et al., 2015b, p. 1208). Papers were coded thematically for “factors assumed to influence CHW performance” (p. 1210). This method was poorly executed and not well suited to the aim of the paper. Not only did

the authors miss nearly 40+ trials⁶ that would have fit their inclusion criteria (including their arbitrary date limit), the purely observational analysis provides no basis for causal inference—simply because a factor is commonly mentioned, does not mean it is an active ingredient in CHW performance. As Collins et al. (2011) explains when describing the MOST Framework and as discussed in section 6.4.2, evaluating a package (e.g. an entire CHW program) does not indicate what components make a difference; components must be isolated in efficient experiments and later tested in combination for their interactions.

This review complements the growing evidence base by virtue of its holistic focus on all possible performance improvement mechanisms and relevant outcomes; pre-specified, systematic, and highly sensitive search strategy; inclusion of published and unpublished literature; inclusion of only the most rigorous trial evidence (i.e. data from randomised controlled trials); and specific focus on head-to-head comparisons of CHWs conducting the same tasks with different types of programme support—the ideal design for isolating programme design components.

6.5 Conclusion

The comprehensive mapping of existing CHW types undertaken during taxonomy construction helped inform the identification of potential intervention components (i.e. “aspect of an intervention that is of interest and can be separated out for study”) (Collins et al., 2011, p. 221). The MOST Framework suggests that, because the intervention components could be combined to form scores of different

⁶ (Andersson et al., 2009; Azad et al., 2010; Baqui et al., 2008; Bari et al., 2011; Bojang et al., 2011; Bolton et al., 2007; CDI Study Group, 2010; Chang et al., 2010a; Chinbuah et al., 2013; Chola et al., 2011; Colbourn et al., 2013a; Cooper et al., 2009; Darmstadt et al., 2010; Datiko & Lindtjorn, 2009; Dias et al., 2008; Ertl et al., 2011; Fottrell et al., 2013; Fylkesnes et al., 2013; Jafar et al., 2009; Kirkwood et al., 2013; Kumar et al., 2008; Kweku et al., 2009; le Roux et al., 2010; Lewycka et al., 2013; More et al., 2012; Neuner et al., 2008; Odendaal et al., 2009; Ohnmar et al., 2012; Pence et al., 2007; Rahman et al., 2009; Selke et al., 2010; Sloan et al., 2008; Swart et al., 2008; Tine et al., 2011; Tripathy et al., 2010; Tylleskar et al., 2011; Wang et al., 2007).

packages, the next step in intervention development is to examine each intervention component individually. Chapter 6 lays the groundwork for such future component selection experiments by summarising the existing evidence on interventions to improve the performance of community health workers for community-based primary health care outcomes in low- and middle-income countries.

This review has highlighted that rigorous assessments of interventions for improving the performance of CHWs are scarce and existing randomised controlled trials are beset by methodological deficiencies and incomplete reporting. Even with these caveats, however, the evidence from this review has critical implications for research and practice.

6.5.1 Implications for Practice

A broad review such as this one can be useful in answering the overarching question of whether evidence exists that interventions to improve CHW performance can work and to identify the “best” intervention (Higgins & Green, 2011). While there is insufficient evidence to assess which interventions to improve CHW performance are most efficacious or effective, there is moderate quality evidence that variations in recruitment, training, supervision, incentivisation, integration, and reporting can (i) improve certain behavioural outcomes for patients, (ii) improve utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services, and (iii) improve CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided (e.g. bedside manner, technical competence, achieving policy objectives). The evidence for biological outcomes was graded high, though the lack of effect on virologic failure and mortality

is likely not indicative of all possible biological outcomes. Evidence is insufficient to draw conclusions regarding the effects of such interventions on retention or to specify which performance improvement intervention strategies are likely to be most effective.

While nearly 40% of studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components, there is moderate quality evidence that the following practices improve behavioural outcomes for patients, utilisation of services, and/or CHW quality of care: (1) When recruiting CHWs, emphasising career possibilities rather than benefits to the community (Ashraf et al., 2015). (2) When supervising CHWs, providing escalating reminders for tasks that are overdue and following-up with underperforming CHWs (Derenzi et al., 2012ABC). (3) When incentivising CHWs, tailoring incentives to measured individual preferences—but only for CHWs performing a single repetitive task, not for CHWs who must perform multiple or more complex tasks (Andreoni et al., 2016, Bossuroy et al., 2016). (4) When equipping CHWs, using mobile phone-based procedural guidance applications (Gautham et al., 2015; Ginsburg et al., 2014).

That said, as discussed in Chapter 4 and section 6.4.3, CHW-led health delivery is complex, and so care should be taken when attempting to apply these strategies. Prior to implementation, practitioners should assess whether (i) there are important differences in health system arrangements that may alter the feasibility and acceptability of the given intervention, (ii) there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting, (iii) there are important differences in baseline health conditions between where the studies were done and the

implementation setting, (iv) studies from which evidence was drawn were conducted in settings with similar social norms, and (v) sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time (Lavis et al., 2009). The detailed information about each study available in Appendix 6d should help practitioners make these assessments.

Notwithstanding the operational insights noted above, this review does not answer all relevant questions pertaining to implementation. More evidence, from trials and non-randomised studies alike, is required to determine (i) what category of inputs has the greatest effect, (ii) what operationalisation of each category has the greatest effect (e.g. with regards to training: frequency, intensity, and type of assessment), (iii) how inputs in different categories interact, and (iv) what adverse effects may stem from such interventions. Given the recent resurgence in interest in the use of CHWs for a range of health issues, for example the recent high-level call by a global Technical Task Force to train one million CHWs in Africa, the only potential recommendation for policymakers, funders and researchers is that they proceed with implementation while using rigorous designs and careful measurement to evaluate programmes.

6.5.2 Implications for Research

Further rigorous research on the effectiveness of interventions to improve CHW performance is needed. Before this can take place, the selection and validity of outcome measures in this area must improve. The lack of consensus building efforts on metrics to CHW performance improvement is a hindrance to the accumulation of knowledge in this area.

Component selection experiments on a greater range of interventions to

improve performance ought to be conducted; this review found few trials on each of the hypothesised drivers of performance. In keeping with the MOST Framework (Collins et al., 2011), this chapter lays the groundwork for such future component selection experiments by summarising the existing evidence on interventions to improve the performance of community health workers.

The review found no trials testing variations in CHW characteristics (e.g. age, sex). While the ways a CHW's innate characteristics affect outcomes has been much debated in the literature (e.g. Bhattacharyya, 2001; López Quiñones, 1999; Ofosu-Amaah, 1983) it is worth considering the ethics of conducting such trials in the future. Many high-income countries have employment discrimination laws that prohibit selection on such criteria except in rare cases related to how an organisation works. That similar legislation is often poorly enforced in LMICs (Heymann, Stein, & Moreno, 2014) should not keep researchers from engaging with the equity issues at the heart of these debates. As such, while more factorial trials and component selection experiments on CHW programme features are warranted, this may not be true of trials on CHW characteristics.

In order to better understand which components of these interventions are most effective and how core components interact, greater attention must be paid to the quality of reporting in such trials (see: Montgomery et al., 2013a). Trialists also ought to describe more thoroughly the intervention under consideration (e.g. design, delivery, uptake context) as well as provide a complete description of the characteristics of CHWs deployed (e.g. see guidance in Chapter 5) All studies ought to include (1) biological patient data, (2) behavioural patient data, (3) quality of care indicators, (4) CHW retention data, (5) process evaluations, and (6) economic data. There is a dearth of data on cost effectiveness of CHW performance interventions

and, although it was not the aim of this review, an assessment of the cost effectiveness of various strategies ought to be undertaken.

Further research should also explore the sustainability of these findings and the transferability of these findings to non-experimentally controlled settings.

Chapter 7: Conclusion

7.1 Overview of the Thesis

Over the past twelve months, the Millennium Development Goals (MDGs) have expired and the world has entered a new cycle of global targets, the Sustainable Development Goals (SDGs). While significant health progress has been made over the life of the former—including a global reduction in the child mortality rate by nearly 50% since 1990 and the maternal mortality ratio by 45% in the same period—it is now clear that none of the three health goals have been met (United Nations, 2015). This reckoning has happened against the backdrop of a worsening global health workforce shortage and the deadliest Ebola epidemic in history (Campbell et al., 2013; Kruk, Myers, Varpilah, & Dahn, 2015).

As the world sets bold new targets focused on championing universal health coverage in the context of these new and on-going challenges, there is an emerging consensus among global health leaders that adequate preparedness against future epidemics and the on-going fight against disease will require building stronger health delivery systems, with particular emphasis on community-based primary health care (Dahn et al., 2015). In order to harness the potential of CHWs to extend health services to poor and marginalised populations and avoid the pitfalls of the post-Alma-Ata period, there is an urgent need to better understand the efficacy, types, and performance of community health workers (CHWs).

This dissertation has made a timely contribution in each of these areas. Following the logic of early stage intervention development, it has (i) rigorously and comprehensively reviewed existing evidence on CHW efficacy and effectiveness, (ii) established, in a faceted taxonomy, the definitional clarity required in theory building and for the optimisation of CHW programmes through component selection experiments; and (iii) identified existing empirical data on the performance of individual intervention components. In undertaking this process, the thesis produced an additional, methodological contribution in the form of an overview of reviews decision tool. This tool may reduce the risk of bias in overviews, and in so doing, facilitate early stage intervention development in the future. This chapter will conclude the thesis by reflecting on the process, summarising its results and implications, considering its strengths and weaknesses, and delineating directions for future research.

7.1.1 Reflections

Prior to undertaking the DPhil, the author of this dissertation had worked as a CHW program manager in Liberia. The author's role as a practitioner influenced this thesis in several ways. It was as a practitioner that the issue of underperforming national CHW programmes initially came to her attention. It was also as a practitioner that the notion of altering how CHWs were supported in their work (i.e. testing different intervention components/programme features) and improving the execution of routine tasks (i.e. implementation fidelity) occurred to her as a way of addressing this issue.

While her experience on the ground undoubtedly created the motivation for the thesis, it is possible it also affected the interpretation of certain findings. Several techniques were used to reduce risk of bias when answering questions of efficacy and

effectiveness (e.g. registered protocols, systematic and transparent searches, audited or double coding). It is possible that in the more interpretive undertaking of thematically synthesising codes for taxonomy construction, however, her pre-existing ideas of what CHW characteristics and programme features were important may have influenced the final facet listing. In order to avoid the privileging of her experience over what was demonstrated by the studies themselves an effort was made to be as transparent as possible in the coding and synthesis (e.g. illustrating each iteration of code names and groupings).

7.2 Summary of Findings by Thesis Aim

This DPhil was originally intended to report the results of a stepped-wedge trial of CHW effectiveness and performance in South-eastern Liberia. Due to the secular events (i.e. the ebola crisis), the trial has yet to be carried out. As a consequence, this thesis now engages with the issue of CHW programme optimisation from a slightly different angle, using evidence syntheses rather than a primary trial. While the trial would have allowed for learning in areas this thesis does not—for example, public-sector implementation and programme delivery costs in truly remote areas—the questions addressed herein inform future trial design and have produced findings relevant in their own right to questions of CHW efficacy, types and performance.

In line with the context outlined above, three related aims for the dissertation were set out in Chapter 1:

- (1) Summarise the effects of CHW delivered interventions for primary health outcomes in LMICs

- (2) Achieve definitional clarity around CHW characteristics and programme features
- (3) Identify the effect of CHW performance improvement interventions for primary health outcomes in LMICs

Each aim will be considered in turn.

7.2.1 Aim One: Summarise the Effects CHW delivered interventions for Primary Health Outcomes in LMICs

Identifying and reviewing existing information is the first step in a phased approach to intervention development that has its conceptual roots in the Multiphase Optimisation Strategy (MOST) (Collins et al., 2011) and the MRC Framework for Developing and Evaluating Complex Interventions (Craig, Dieppe, Macintyre, & Michie, 2008). While CHWs typically deliver medical interventions in combination (Paintain et al., 2014; Puett, Coates, Alderman, Sadruddin, & Sadler, 2012), first ascertaining whether and how they can do each safely and efficaciously is important for intervention design and optimisation. Two categories of findings emerged from this work: (a) one related to research synthesis methods and (b) one related to CHW efficacy and effectiveness.

(a) Research synthesis methods

Over the last five years, there has been a proliferation of reviews on CHW effectiveness. It was thus initially believed by the author of this DPhil that the best way to summarise the effects of CHW delivered interventions for primary health outcomes in LMICs was by using an overview of reviews.

In order to synthesise existing guidance concerning overview practice, a scoping review was conducted in Chapter 2. This found that there is emerging debate regarding overlapping systematic reviews; systematic review scope; quality of

included research; updating; and synthesising and reporting results. While three functions for overviews have been proposed—identify gaps, explore heterogeneity, summarise evidence—overviews cannot perform the first; are unlikely to achieve the second and third simultaneously; and can only perform the third under specific circumstances. Namely, when identified systematic reviews meet the following four conditions: (1) include primary trials that do not substantially overlap; (2) match overview scope; (3) are of high methodological quality; and (4) are up-to-date. These four conditions were developed into a pilot decision tool to assess the conditions under which employing an overview of systematic reviews is likely to lead to a high risk of bias. This tool will hopefully help facilitate the synthesis of evidence and thus early stage intervention development in the future.

(b) CHW efficacy and effectiveness

The application of the overview checklist in Chapter 3 demonstrated that the effects of lay health worker interventions for primary health outcomes in LMICs are best assessed in the context of a systematic review, rather than an overview. The protocol and results of this systematic review, the largest review of CHW efficacy and effectiveness to date, were reported in Chapter 4.

The systematic review identified 155 papers reporting 86 trials, of which nearly 40% were published in the past five years. The diversity of included studies limited meta-analysis to outcomes for nine of these ten comparison groups. For comparisons addressing mental health outcomes, included analyses showed moderate quality evidence that CHW delivered interventions can reduce the prevalence of depression, improve PTSD symptoms, and reduce functional impairment related to CMD. There was also low quality evidence that CHW-led interventions can lessen the severity of common mental disorders and very low quality evidence that CHW

delivered interventions may have a small to moderate effect on reducing adolescent depressive symptoms.

For comparisons addressing maternal and paediatric health outcomes, the included analyses showed high quality evidence that CHW delivered interventions reduce perinatal mortality and improve WAZ. They showed moderate quality evidence that CHW delivered interventions reduce child morbidity, reduce neonatal mortality, improve WHZ, and increase exclusive breastfeeding and the speed of its initiation. They also showed moderate quality evidence of no difference in infant mortality, stillbirths, maternal mortality, and maternal morbidity between care provided by CHWs and usual, usually facility-based, care. There is low-quality evidence that CHW delivered interventions may reduce child mortality, increase the uptake of clean delivery kits, improve infant cognitive development, and produce no difference in HAZ. Finally, there is very low quality evidence that CHW delivered interventions have a large effect on maternal knowledge of infant development, may reduce infant morbidity, and may increase facility-based delivery.

For comparisons assessing infectious disease outcomes, included analyses showed high quality evidence that CHW delivered interventions can improve TB completion rates versus facility-based care. They also showed moderate quality evidence that CHW DOTS is not superior to family member DOTS, and that facility-led maintenance of ARVs is no better than CHW-led maintenance in terms of loss to follow-up or death. There is low-quality evidence that CHW delivered interventions may improve TB cure rates versus facility-based care. Studies too diverse for meta-analysis also suggest that CHW-led interventions can improve retention in pre-ARV care and that CD4 count may be similar in home-based care provided by CHWs and facility-based care.

For comparisons assessing public and preventative health outcomes, the included analyses showed low quality evidence that CHW delivered interventions may reduce the presence of home hazards and increase the coverage of mass treatment or preventative therapy. Studies too diverse for meta-analysis suggest that home visitation to educate or deliver vaccines, whether by CHWs or other health professionals, is likely to be useful in improving the coverage of a range of vaccines for children and that CHW-led interventions with appropriate clinical support can lower systolic blood pressure.

Poor reporting of CHW characteristics, CHW programme features, and regional or political context unfortunately made it difficult to explore heterogeneity stemming from such differences and, consequently, to make inferences regarding how to implement and optimise CHW programmes in different settings. That said, careful grouping of studies by PICOS characteristics has nonetheless allowed for insight into specific areas of CHW efficacy. Importantly, while the magnitude of the effect of CHW led-interventions varied, the results indicate that for no outcomes did CHWs cause harm or perform worse than usual care. Moreover, there is mixed quality but consistent evidence that, given reduced training and remuneration costs, CHWs could potentially reduce the direct and indirect costs of health care.

7.2.2 Aim Two: Achieve Definitional Clarity Around CHW Characteristics and Programme Features

Definitional clarity regarding intervention components is a precursor to the next steps in intervention development: theory building and optimisation of programmes through component selection experiments (Collins et al., 2011). In answer to repeated calls for a comprehensive taxonomy that allows for the conceptualisation of CHW delivered interventions “in terms of their lay component

rather than in terms of the specific health issues [tasks] on which they focus” (Lewin et al., 2010, p. 42; van Ginneken et al., 2013), systematic review methods were used in Chapter 5 to develop a faceted taxonomy of CHW characteristics and programme features.

Two hundred and fifty-three records were included in taxonomy construction. Records were (i) existing classification schemes, (ii) primary randomised controlled trials of CHW delivered interventions and connected process evaluations/qualitative studies of CHW experience, or (iii) influential studies (>100 citations) in the CHW literature. Inductive thematic analysis was used to identify, describe and categorise the full range of CHW characteristics and programme features. A faceted taxonomy with two categories (CHW Characteristics and CHW Programme Features) and six dimensions (Integration, Recruitment, Training, Supervision, Incentives, and Equipment) emerged.

7.2.3 Aim Three: Identify the Effect of CHW Performance Improvement Interventions for Primary Health Outcomes in LMICs

The comprehensive mapping of existing CHW types undertaken during taxonomy construction helped inform the identification of potential intervention components (i.e. “aspect of an intervention that is of interest and can be separated out for study”) (Collins et al., 2011, p. 221). The MOST Framework suggests that, because the intervention components could be combined to form scores of different packages, the next step in intervention development is to examine each intervention component individually. Chapter 6 lays the groundwork for such future component selection experiments by summarising the existing evidence on interventions to improve the performance of community health workers for community-based primary health care outcomes in low- and middle-income countries.

The review identified 14 trials evaluating the effects of interventions for improving the performance of community health workers for community-based primary healthcare in low- and middle-income countries. The heterogeneity and poor reporting of included studies precluded meta-analysis.

A structured summary laid out moderate quality evidence of the efficacy of CHW performance interventions in (i) improving behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided (e.g. bedside manner, technical competence, achieving policy objectives). The evidence for biological outcomes was graded high, though the lack of effect on virologic failure and mortality may not be indicative of all possible biological outcomes. None of the studies included in this review reported on possible adverse effects of these interventions, either to patients, CHWs, or the health system. Evidence is insufficient to draw conclusions regarding the effects of such interventions on retention or to specify which performance improvement intervention strategies are likely to be most effective. Furthermore, nearly half of the studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components.

That said, there is moderate quality evidence that the following practices improve behavioural outcomes for patients, utilisation of services, and/or CHW quality of care: (1) When recruiting CHWs, emphasising career possibilities rather than benefits to the community (Ashraf et al., 2015). (2) When supervising CHWs,

providing escalating reminders for tasks that are overdue and following-up with underperforming CHWs (Derenzi et al., 2012ABC). (3) When incentivising CHWs, tailoring incentives to measured individual preferences—but only for CHWs performing a single repetitive task, not for CHWs who must perform multiple or more complex tasks (Andreoni et al., 2016, Bossuroy et al., 2016). (4) When equipping CHWs, using mobile phone-based procedural guidance applications (Gautham et al., 2015; Ginsburg et al., 2014).

While it is important to recognise that variations in recruitment, training, supervision, incentivisation, integration, and reporting can affect outcomes, the four strategies listed above ought to be interpreted with caution. As explained in Chapter 1, CHW programmes are complex interventions that unfold in complex systems; intervention effects may not simply be the sum of their components but emergent trajectories that result from dynamic processes—including interaction with the context in which they take place (Craig, Dieppe, Macintyre, & Michie, 2008; Hawe et al., 2009). As such, practitioners should assess prior to implementation whether (i) there are important differences in health system arrangements that may alter the feasibility and acceptability of the given intervention, (ii) there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting, (iii) there are important differences in baseline health conditions between where the studies were done and the implementation setting, (iv) studies from which evidence was drawn were conducted in settings with similar social norms, and (v) sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time.

7.3 Discussion

Mounting pressure to meet ambitious international health goals and avoid past mistakes points urgently to the need to better understand the efficacy, types, and performance of community health workers (CHWs). Following the logic of early stage intervention development, this thesis has made significant contributions to the existing literature on evidence based practice and community health worker programme design and management by: (i) producing a rigorous, up-to-date, and comprehensive systematic review of the current best evidence on CHW efficacy and effectiveness; (ii) answering repeated calls in the literature for a comprehensive taxonomy that allows for the conceptualisation of CHW delivered interventions “in terms of their lay component rather than in terms of the specific health issues [tasks] on which they focus” (Lewin et al., 2010, p. 42; van Ginneken et al., 2013); (iii) conducting the first systematic review of existing empirical data on the performance of individual intervention components; and (iv) advancing the science of research synthesis by conducting the first systematic synthesis of existing overview guidance and deriving a decision tool to minimise risk of bias in overview of reviews. Over the course of this work, several crosscutting issues relevant to evidence based practice and community health worker programme design and management have emerged. Matters related to the nature of evidence synthesis, methodological issues in evaluating complex interventions, and conceptualising and measuring CHW efficacy and performance will be discussed in turn.

7.3.1 The Nature of Evidence Synthesis

Overviews have gained popularity in recent years due to their ability to quickly summarise information about broad comparisons (Pieper et al., 2012; Thomson, 2014). The scoping review conducted in Chapter 2, however, demonstrates

that current overview guidance is heavily informed by consensus and personal experience rather than empirical evidence. The problems of overlap, scope, evaluating the quality and reporting of included research, updating included systematic reviews, and how to synthesise and report findings remain unresolved—raising the question of when overviews are an appropriate tool for evidence synthesis.

This question reflects the classic evidence-based practice dilemma of comprehensiveness and reliability versus speed. While Caird et al. (2015) acknowledge that the shortcomings in current overview practice may be “uncomfortable for reviewers,” they nonetheless conclude that overviews are a useful means of synthesising evidence across broad topic areas (p. 95). The decision tool presented in Chapter 2, by contrast, has attempted to differentiate between circumstances in which the efficiency of an overview does not compromise the validity of its outcome and those in which such expediency is likely to introduce serious bias in the form of overly precise estimates and the improper exclusion of relevant data.

As demonstrated in Chapter 3, however, the tool is not quick to apply. While the checks it outlines should normally be undertaken as part of a high quality overview of reviews, the time associated with arriving at the conclusion that the overview should not proceed and conducting a broad systematic review with multiple sub-groups in its place may prove to be too much in some contexts.

That said, how much “time” is available depends on how one conceives of the project of evidence synthesis (and indeed, evidence based practice itself): should synthesis be a time-bound activity undertaken in response to specific policy questions at the moment they are raised or should synthesis involve a longer-term investment – of time and finances—that is undertaken with the utmost rigour? The risk of the

former is potential misleading results, the risk of the latter is having no results ready at a given “policy moment.”

Davies (2015) distinguishes between *strategic* (i.e. over the medium to long term) and *operational* (i.e. day-to-day) use of evidence. Syntheses that prioritise rigour might take 12-18 months to complete and so, while useful for strategic policymaking, are only of use for operational decision-making if they have been completed before the need to use them arises. Where this is not the case, it may be better to undertake a rapid evidence assessment that can be completed in 1-3 months.

Yet consistent failure to invest in strategically oriented evidence synthesis of the type enabled by the overview tool has ramifications. The history of CHWs is instructive in this regard: a WHO report once declared: “there is no longer any question of whether CHWs can be key agents in improving health; the question is how their potential may be realised.” This question is highly pertinent and is the central query at issue in Chapter 6. The paper in question, however, is not a new agenda-setting piece—it is from nearly twenty years ago (Kahssay, Taylor, & Berman, 1998, p. 9). Methodologist, Julia Littel has noted that this revisiting of unresolved policy questions is not atypical (Littell, 2016). In not taking the long view and insisting on rigour in the first instance, we may be relegating ourselves to a constant re-living of one of Murphy’s laws: that there is never enough time to do it right the first time, but there is always enough time to do it over.

7.3.2 Methodological Issues in Evaluating Complex Interventions

An important finding of this dissertation is that a sizeable proportion of trials on CHW efficacy, effectiveness, and performance are poorly designed and reported. This undermines efforts to aggregate findings and to derive operational insight. The

nature of these shortcomings, as well as ways in which future trials might be improved, is discussed in the sections that follow.

Randomised trials and evidence based practice

As explained in Chapter 1, the privileging of RCTs over other methods has resulted in considerable debate, even within the medical community (Ashcroft, 2004; Bothwell, Greene, Podolsky, & Jones, 2016). It has been argued that there is no role for randomised trials in operations research (Zachariah et al., 2009) and objections to the use of randomised evaluations for social interventions have been well chronicled (Oakley et al., 2003).

The pragmatically oriented trials in Chapter 4, however, demonstrate the potential weakness of ecological evaluation and the value of randomisation in assessing intervention effects. In several trials (e.g. Bolton et al., 2007; Bolton et al., 2003; More et al., 2012), the intervention would have appeared an unqualified success had assessments been based on favourable trends in the intervention arm, rather than statistically significant differences between trial arms. While there are policy-relevant questions RCTs do not answer—e.g. sustainability, transferability, financing—by answering the question “can the intervention work?” a comprehensive systematic review of RCTs is an essential first step in furnishing policymakers with the necessary information for decision-making.

Moreover, Chapter 6 illustrated how randomisation can be used to isolate the effects of particular programme components—a critical step in intervention optimisation (Collins et al., 2011). While several CHW efficacy papers in Chapter 4 noted that certain CHW programme characteristics were important for programme success—e.g. “selection” (Agrasada et al., 2005; Haider et al., 2000) and “supervision” (Agrasada et al., 2005; Azad et al., 2010; Bari et al., 2011; le Roux et

al., 2013; Rahman et al., 2008; Rotheram-Borus et al., 2014)—none of these studies describe how these conclusions were reached. Moreover, these studies only identify components (e.g. “supervision”) associated with CHW performance, not how these components ought to be operationalised (e.g. “weekly supportive supervision by a nurse composed of supply audits, patient checks, and on-the-job training”). While the evidence found in Chapter 6 was insufficient to specify which performance improvement intervention strategies are likely to be most effective, the review isolated four specific strategies related to CHW recruitment, training, supervision, and incentives that moderate quality evidence suggests improve behavioural outcomes for patients, CHW quality of care, and/or utilisation of health services.

Though, in both cases, data from randomised control trials are not the only evidence relevant to the questions asked, this thesis underscores the utility of RCTs in answering efficacy and optimisation questions pertaining to community-based interventions.

Weaknesses in trial design

The body of evidence in the efficacy and performance reviews reflected, however, common pitfalls in trial design related to the evaluation of complex social interventions.

Efficacy vs. effectiveness

Firstly, over three quarters of included interventional trials were delivered in efficacy settings. In such settings, CHWs are more likely to have been carefully selected; managers more motivated; remuneration more available due to research funding; training and supervision more intensive; and recipient groups curated (Lewin et al., 2010; van Ginneken et al., 2013). In fact, several trials in the sample excluded residents of precisely the type of hard-to-reach geographies that CHWs were designed

to serve (World Health Organization, 1978): Jaffar et al. (2009) excluded those living on far-away islands, Kramer et al. (2014) those living in villages “considered by local leaders to be too difficult to reach in the rainy season” (p. 5323), and Ohnmar et al. (2012) those living in areas with security concerns or small populations. These exceptional inputs and exclusions, likely made the interventions easier to implement than they would be in non-experimental settings.

Economic data

Secondly, less than a quarter of included trials in either review reported economic data. Of these papers, less than two thirds were full economic evaluations. The evaluations that were conducted illustrated the difficulty of such assessments: (a) Most interventions were geographically limited and so could not benefit from potential economies of scale. (b) The CHWs in most trials delivered a single intervention. Such specialist CHWs, whether delivering a curative or preventative service, have been found to be expensive compared to those that cover multiple interventions (Makan, 1997). (c) Most evaluated programmes had been designed for the purpose of the trial. Given that CHWs and managerial staff typically become more efficient over time (Makan, 1997), it is possible that costs might fall and effectiveness could potentially increase—though it is also possible that effects may attenuate over time. (d) More than two thirds of the costing exercises were conducted from the provider’s perspective (i.e. only costs incurred by the provider were considered). Population level savings (e.g. relating to having a healthier workforce) that might be relevant to policymakers were not included. Given that cost of care is a major burden for people living in the rural areas CHWs typically serve (Jaffar et al., 2009), evaluations that take a broader societal perspective are warranted.

Sustainability and long-term effects

Thirdly, particularly in the CHW performance review, included trials did not assess the long-term effects of interventions. While providing additional inputs or supports to underperforming programmes may create short-term gains in patient health or CHW functioning, the novelty—and in turn the effects—of such inputs may wear off over time. On the other hand, any potential benefits of psychosocial interventions (Chatterjee et al., 2014) and behaviour change interventions (Colbourn et al., 2013) may require longer to take effect and so the short follow-up of such trials may also have resulted in an underestimation. Either way, it is important to assess the sustainability of these interventions at scale.

Selection and validity of outcome measures

Fourthly, selection and validity of outcome measures was a concern across studies included in the two systematic reviews. In Chapter 6, for example, there were a wide range of measurement instruments employed to assess “quality of care” and several measured activity, not quality. While CHWs must perform their assigned functions as a prerequisite to providing quality care, the two are not the same and should not be confused. Similarly, direct assessments of performance were rare.

Poor consistency in outcome selection and definition was likewise evident in Chapter 4, where the diversity of measures used for the same construct—e.g. virologic failure was defined slightly differently in Chang et al. (2010) (>400 copies/mL) and Jaffar et al. (2009) (> 500 copies/mL)—sometimes precluded meta-analysis. The lack of consensus building efforts on metrics to judge CHW efficacy, and particularly performance improvement, is a hindrance to the accumulation of knowledge in this area.

Hawthorne effect

Fifthly, a number of studies documented that “standard care” was poor, or non-existent, possibly reflecting the fact that many currently have “only ad hoc, sub-scale” CHW programmes (Dahn et al., 2015, p. 1). In such cases, the presence of any intervention is likely to produce a considerable Hawthorne effect. This makes it difficult to assess what processes can be safely task shifted to CHWs or test new performance improvement interventions. While the Hawthorne effect cannot be overestimated (e.g. Morrow et al. (1999) note that even with frequent home visits by investigators prior to the trial the outcome of interest did not change until the intervention had begun), future studies would do well to adjust for the attention effect.

Process evaluations

Finally, the use of process evaluations, would have helped explain the success, failure, or unexpected consequences of a trial; illuminate causal mechanisms of a successful intervention as well as avenues for optimisation; evaluate fidelity and quality of implementation; and reveal the degree to which the treatment can be standardised (Linnan & Steckler, 2002; Oakley et al., 2006). The identification of core components is particularly relevant given that so many trials in both Chapter 4 and 6 tested compound interventions; information on the relationship between outcomes and individual intervention components is critical for those designing CHW programmes (Craig et al., 2008; Glasziou et al., 2008; Glenton et al., 2011a; Shepperd et al., 2009).

Weaknesses in trial reporting

Across the majority of trials in both systematic reviews, information regarding CHW characteristics and programme features, control conditions, adverse events and health context was lacking. While information deficits are partially related to trial

design, as discussed in the section on process evaluations above, they are also related to sub-par reporting.

Implementation

Implementation fidelity indicates whether an intervention is delivered as intended by the developers—in terms of intervention design, delivery, uptake, and context (Borrelli et al., 2005; Montgomery et al., 2013b). Implementation fidelity is important to assess because, as explained in Chapter 6, non-significant results may either be the result of (i) an ineffective intervention, (ii) a failure to optimise the intervention, (iii) poor implementation, or some combination of these three (Basch et al., 1985; Borrelli, 2011; Collins et al., 2011; Hasson, 2010). In addition, measuring fidelity in smaller and/or pilot trials enables researchers to assess whether a programme will be feasible during larger trials or in different settings (Lancaster, Dodd, & Williamson, 2004).

Unfortunately, few included studies in the systematic reviews reported implementation information like dosage (Azad et al., 2010; Baker-Henningham et al., 2005; Baqui et al., 2008; Darmstadt et al., 2010; Kirkwood et al., 2013; Lewycka et al., 2013; Penfold et al., 2014) or frequency (Haider et al., 2000) in enough detail to permit the sub-group analyses necessary for any heterogeneity to be explored.

Moreover, as with most complex interventions, it is likely that components of the CHW delivered intervention interact with contextual factors and each other (Craig et al., 2008; Montgomery et al., 2013b). With regard to public health interventions, Lavis et al. (2009) have noted that differences in health system arrangements, environmental conditions, baseline health conditions, social norms, and resources levels may determine whether an intervention shown to be efficacious in one setting will be generalisable to another. Several trials in Chapter 4, for example, conducted

health facility strengthening in both arms alongside the CHW delivered intervention in one arm (e.g. Waiswa et al., 2015). The degree to which this continuity of care (Kerber et al., 2007; Liu et al., 2011) enhanced or enabled the concurrent CHW delivered intervention should be considered more carefully in future studies. Similarly, additional information on baseline programme characteristics would have provided greater assistance in understanding the threshold effect, if any, of the various CHW performance interventions examined in Chapter 6. Without better reporting of culture and context it is difficult to assess the efficacy of CHW efficacy and performance improvement interventions or their applicability elsewhere.

Adverse effects

Fewer than 10% of trials across the two systematic reviews reported on adverse effects and only 3% reported on the equity effects of their interventions. Given this, it is difficult to draw conclusions regarding the trade-off between benefits and harms (Lavis et al., 2009a; Welch et al., 2013). That said, many of the outcomes were adverse events (e.g. death) and CHW-led interventions, by nature of being community-based, tend to be targeted at the most vulnerable.

7.3.3 Conceptualising and Measuring CHW Efficacy and Performance

While enormous intellectual energy has been spent testing models of task shifting (Callaghan et al., 2010; Fulton et al., 2011; Joshi et al., 2014; Kredo et al., 2014; Mdege et al., 2013), less energy has been expended on issues related to systematically disaggregating the types of personnel involved. There is a distinct lack of standardised language to describe CHWs. As noted in Chapter 1, two recent systematic reviews of CHW delivered interventions (Bhutta et al., 2010; Gilmore & McAuliffe, 2013) identified nearly 70 unique terms (ranging from “barefoot doctor” to “lady health worker”) used worldwide for community health workers (see

Appendix 1a for a listing). Even with suitable search strategies, this variation hinders identification of relevant research, hampering synthesis and knowledge transfer (Glasziou et al., 2008).

This failure to identify research and disaggregate types of CHWs may mean that particular characteristics (literacy, CHW remuneration) that distinguish between interventions remain unidentified. If such unreported or “unseen” differences are “active ingredients” associated with effectiveness, the design of optimally effective interventions (Collins et al., 2011) and the uptake and replication of such interventions will be impeded.

The definitional clarity achieved by the taxonomy developed in Chapter 5 makes several important contributions toward helping identify such active ingredients and mitigating several of the trial design issues reviewed above.

Firstly, it provides a foundational step toward standardisation and accuracy of reporting CHW characteristics and programme features, as required by CONSORT and TREND. As explained in Chapter 6, given the range of vastly different health workforce configurations routinely classed as “CHW programmes,” adhering to guidance from existing extensions to CONSORT (Boutron et al., 2008; Des Jarlais et al., 2004; Hoffmann et al., 2014) may still frequently yield descriptions of interventions where salient features go unreported. The taxonomy clarifies and defines relevant CHW characteristics and programme features and provides standardised language with which to describe them. Moreover, it is presented in an easy-to-use form that trialists can complete and include as a descriptive appendix in new studies, thereby facilitating consistent and complete reporting.

Secondly, the taxonomy adds much-needed operational nuance to CHW delivered intervention components. For example, several trials in the CHW efficacy

review described in Chapter 4 identified “supervision,” “training,” and “selection” as critical to the success of the CHW-delivered intervention they were testing, but none defined the terms (Agrasada et al., 2005; Haider et al., 2000; Rahman et al., 2008; Rotheram-Borus et al., 2014). Rather than treat these as monolithic concepts, the taxonomy delineates the components of each. For example, supportive supervision is broken into four discrete tasks (verifying reports, restocking supplies, observing home visits, and collecting community feedback) that may all be separate active ingredients in the CHW programme features. This conceptual mapping enables (i) the relevant, precise reporting of CHW programme features necessary for mediation and moderation analysis (Preacher, Rucker, & Hayes, 2007) and (ii) the definitional clarity required to run component selection experiments, identified by the Multiphase Optimisation Strategy (MOST) as helpful in identifying the optimal combination (and dose) of intervention components (Collins et al., 2011).

Thirdly, the taxonomy will enable systematic reviews like those in Chapter 4 and Chapter 6 to use a standardised method for extracting information about CHW characteristics and program features. This will aid in (i) the identification of areas where research is strong and where gaps exist, (ii) the reliable synthesis of evidence, and (iii) further exploration of discrete, replicable characteristics (or combinations of characteristics) as potential mechanisms of action associated with intervention effectiveness (Michie et al., 2013).

Fourthly, this enhanced classification and characterisation will allow for more accurate replication of CHW programme models. Not only does this enable (i) the investigation of generalisability across behaviours, populations, and settings, but (ii) the operationalisation of the practice standards that permit the faithful propagation

and uptake of interventions found to be efficacious or effective (Abraham & Michie, 2008; Michie et al., 2013).

7.3.4 Considering the Results Against the Thesis Aims

In light of these deficiencies in data and the limitations of the methods employed, it is worth considering to what degree the work undertaken has allowed for the answering of the dissertation's research questions and, ultimately, what can be claimed versus what remains unknown.

The first objective of the dissertation was to assess the effects of CHW interventions for primary health outcomes in LMICs. Given the methods employed and the nature of the evidence identified, this question was ultimately only partially answered.

Because the research question concerned efficacy and effectiveness and the “complex” characteristics of the interventions evaluated, the choice was made to focus on impact evaluations most capable of minimising selection bias (RCTs). That said, it is clear that not all CHW-delivered interventions have been evaluated using such methods and, in particular, that a focus on RCTs is likely to systematically underrepresent public health interventions (Ogilvie et al., 2005; Petticrew & Roberts, 2003; Rychetnik et al., 2002). A better balance between rigour and comprehensiveness would likely have been struck by including all evaluations able to establish a credible counterfactual (e.g. non-randomised trials). This was not possible given the resources available for the DPhil, but would be a worthwhile future project. While Chapter 4 covers new topics never before systematically reviewed (CHW-delivered interventions for injury prevention, NCDs, healthy child feeding practices, child cognitive development, and mental health outcomes), because some CHW

interventions were not assessed, it remains a potentially incomplete, rather than comprehensive, picture of the evidence.

Moreover, despite the focus on the type of evaluation best suited to answer questions of efficacy, the pitfalls in trial design and reporting noted in the previous section limited the ability of the systematic review to make clear statements about the efficacy of several of the CHW-led interventions considered. When the body of evidence for each comparison was assessed using the GRADE system, only three of the thirty-seven comparisons scored “high quality.” Over three quarters of the trials were conducted in efficacy settings, few considered long term effects, and almost none reported implementation data. These three factors make it difficult to assess (i) if the interventions assessed will work in real-world settings, (ii) whether the observed effect will be maintained over time, and (iii) whether type III error occurred (Basch et al., 1985; Borrelli, 2011; Collins et al., 2011; Hasson, 2010). These issues are critical to well-formed claims concerning the effect of CHW-led interventions; as a result, even for interventions assessed, the data only provides a partial answer to the research question.

While this may seem to substantiate concerns that EBP will lead to therapeutic nihilism (Straus & McAlister, 2000), it is worth noting, as explained in Chapter 1, that EBP by definition combines best evidence, practitioner experience and patient/community circumstances. As a result, patients are able to benefit from practitioner judgement and contribute—as informed participants—to decision making, even when empirical data do not fully answer the question at hand (Gibbs & Gambrill, 2002; Sackett, 2000).

That said, it is likely that additional analyses and different types of data could expand or strengthen claims made in answer to the research question. Had more time

and resources been available, pooling across outcomes or performing meta-regression on the results may have allowed for further insights into the effects of different intervention types, control conditions, and scale or setting (e.g. efficacy vs. effectiveness). Power analysis could make it clear whether certain outcomes (e.g. maternal mortality) were underpowered to detect a meaningful clinical difference. Finally, qualitative studies could contribute usefully to a more holistic understanding of the processes underlying CHW-led interventions. A realist, qualitative or meta-ethnographic review exploring mechanisms underlying CHW-led interventions would, as noted in Chapter 4, also provide complementary information. It would, however, be a separate and considerable undertaking involving distinct criteria, synthesis, and interpretation (cf. Dixon-Woods, 2011; Thomas & Harden, 2008; Wong et al., 2013).

The second objective of the dissertation was to achieve definitional clarity around CHW characteristics and programme design features. While the taxonomy developed in Chapter 5 meets this objective and is offered for consideration by the field, the facets would likely be further refined if exposed to a consensus process.

CHWs are problematic for taxonomic representation because there are multiple organising principles at work (Rosenfeld and Morville, 2002). As was clear from empirical representations consulted during taxonomy construction, there is no single principle that differentiates CHWs in a way that prevents certain characteristics or program features from belonging to more than one superordinate concept. As a result, rather than use the more common tree-based taxonomy (which subdivide types on the basis of only one dimension at a time), it became clear that the taxonomy would need to take a form that allowed for the categorisation of content along multiple dimensions at the same time. By using facets—each facet being, in essence, a “base

taxonomy” comprising one of the fundamental dimensions along which content can be analysed (Lambe, 2014, p. 37), CHWs were more appropriately classified according to a number of unrelated aspects. While this has the advantage of more faithfully representing the number of CHW “types” it does not map well and may not be immediately recognisable as a taxonomy.

While the taxonomy produced answers the research question asked in Chapter 1, it is possible to dispute whether this was the correct question in the first place. An argument can be made, for example, that the definitional clarity on CHW tasks is also important. As explained in Chapter 5, effectiveness studies have often failed to separate the health impact of a specific task (e.g. a vaccine), from the delivery strategy used to reach the target population (Victora, 2004). As such, the author felt it was important to restrict the focus of the taxonomy to CHW characteristics and programme features in order to more clearly separate the *task* (e.g. malaria management) from the features of the *delivery agent* (e.g. CHWs).

It is likewise recognised, however, that taxonomies related to other cadres of healthcare workers (e.g. nurses, Iowa Intervention Project, 1992) typically list tasks. While the motivation for this project stemmed largely from the fact that existing reporting guidelines demand few salient details about the delivery agent (but many more about the task), cataloguing CHW tasks could nonetheless complement the work begun here. That said, while a taxonomy of CHW characteristics and programme features is useful to illustrate the heterogeneity of a class of health workers typically written about in monolithic terms, it is worth asking whether a taxonomy of tasks could serve a similar purpose: that CHWs perform an array of tasks is already well reflected in the literature. Given this, a systematic review of the evidence for each of

these CHW-delivered tasks, similar but more granular than that performed in Chapter 4, would likely be more worthwhile than the generation of a simple listing.

Thus, while consensus methods may be used in future to further refine the taxonomy and further sources (e.g. paper-based program documents) incorporated, it remains a long overdue, foundational first step in the comprehensive and standardised description of CHW characteristics and programme features.

The third and final aim of the dissertation was to identify CHW performance improvement interventions for primary health outcomes in LMICs. While history has demonstrated that broad contextual factors related to the health system or the political and economic climate can have an impact on CHW, in line with the MOST framework, the review in Chapter 6 was focused on modifiable components of the intervention itself, namely: CHW characteristics and programme features (Collins et al., 2011). Ultimately, the systematic review produced—while likewise conducted in line with quality and reporting guidelines MECIR and PRISMA—has similar issues to the review in Chapter 4: the methods employed favour rigour at the expense of comprehensiveness and the trial quality was not sufficient to make strong operational claims in answer to the research question. Examining non-randomised evidence in future may mitigate these two issues.

The review nonetheless raises important question regarding interpretation and optimisation more generally, as nearly 40% of studies located were compound interventions. How to best pick which components get assessed in the efficient experiments envisioned by the MOST framework remains an open question. The taxonomy produced in Chapter 5 reveals the huge number of possible component combinations—very conservatively assuming each facet is dichotomous would still yield $\sim 3 \times 10^{18}$ “types of CHWs—and underscores the impossibility of testing all

possible configurations. That these interventions typically have complex characteristics—in that their components may interact with each other or the context—amplifies this issue.

Thus, while the claims made about CHW performance in Chapter 6 provide an answer the original research question, they only do so with the caveat that other configurations may prove more efficient than those tested and that several factors (cf. Lavis et al. 2009, discussed above) may have a bearing on the success on the prescribed strategies when they are transported to different settings.

The implications of these findings will be considered in the next section.

7.4 Implications and Directions for Future Research

7.4.1 Implications for Practitioners And Policymakers

1. Practitioners and policymakers should consider expanding the remit of existing CHWs to include the areas where they have been found to be efficacious (specified in section 7.2.1), keeping in mind the interpretive caveats described earlier and using rigorous designs and careful measurement to evaluate programmes and document implementation along the way. As discussed in Chapter 4, CHW-led health delivery is a complex intervention, and the reviews conducted in this thesis do not answer all relevant questions pertaining to operationalisation of these programmes.
2. While there is insufficient evidence to assess which changes to CHW program features are most efficacious or effective in improving performance, there is moderate quality evidence that variations in recruitment, training, supervision, incentives integration, and reporting can improve behavioural outcomes for patients, CHW quality of care, and utilisation of health services. Specifically, practitioners should consider the four specific strategies listed in section 7.2.3,

but not before assessing the five factors indicated that may have a bearing on the success of such changes when they are transported to different settings (Lavis et al. 2009). Given the recent resurgence in interest in the use of CHWs for a range of health issues, for example the recent high-level call by a global Technical Task Force to train one million CHWs in Africa, it is recommended that policymakers and practitioners proceed with implementation while using rigorous designs and careful measurement to evaluate programmes.

3. Relatedly, when describing CHW programmes, practitioners should employ the taxonomy in Chapter 5 to help clearly delineate CHW characteristics and programme features in order to allow for more accurate replication of CHW programme models and, in turn, investigation of generalisability.

7.4.2 Implications for Researchers

1. Overviews are always susceptible to the interpretive, conceptual, and reporting biases of previous reviewers; however, when the four conditions in the decision tool described in Chapter 2 are not met, the overview outcomes risk being biased and misleading. In these cases, a broad systematic review with multiple comparisons would likely produce a more precise result than an overview. Given that the number of overviews published per year between 2004 and 2011 has risen fivefold, from six to thirty (Pieper et al., 2012), and that overviews of reviews are the most downloaded content in certain evidence synthesis journals (Thomson, 2014), it is critical that this be addressed.
2. Researchers should employ the enclosed taxonomy so as to produce consistent, precise, and complete reporting of relevant CHW-related information as required by the CONSORT statement and TREND. This definitional clarity is

required for optimisation of CHW programmes through component selection experiments and systematic review subgroup analysis and enables the reliable identification and synthesis of evidence required to identify research gaps.

3. An important finding of this dissertation is that a sizeable proportion of trials on CHW efficacy, effectiveness, and performance are poorly designed and reported. This undermines efforts to aggregate findings and to derive operational insight. Trialists ought to heed the issues discussed in section 7.3.2 when designing future trials.

7.4.3 Directions for Future Inquiry

This thesis was informed by and structured along the lines of frameworks for early stage intervention development (Collins et al., 2011; Craig et al., 2008). While it has made significant contributions in identifying evidence of CHW efficacy, establishing the definitional clarity required in theory building, and reviewing existing data on performance of individual intervention components, there remains more to do in creating and optimising evidence-based CHW delivered interventions.

1. As Berman et al. (1987) note in their seminal paper on CHWs, “CHW programmes represent a mode for the organisation of services rather than a type of intervention [task]” (p. 445). Effectiveness studies have often failed to separate the health impact of a specific task (e.g. a vaccine), from the delivery strategy used to reach the target population (Victora, 2004). For health issues in which evidence is currently of low quality or does not exist, new trials assessing the effectiveness of CHW delivery are warranted. These studies ought to incorporate process and have long-term follow-up so as to better assess how high quality, sustainable operationalisation of such interventions is feasible at scale in remote settings.

2. New studies of CHW effectiveness and performance should be accompanied by full economic studies in order to assess cost effectiveness. Given that cost of care is a major burden for people living in the rural areas CHWs typically serve (Jaffar et al., 2009) evaluations that take a broader societal perspective are warranted.
3. For health tasks where CHW delivery demonstrate benefits, the focus of new research ought to shift to optimising public health impact—i.e. understanding what CHW or programme characteristics enable or mediate effects (Collins et al., 2011). The studies identified in Chapter 6 do not represent the entire realm of approaches to improving CHW performance in LMICs (e.g. no studies examine the impact of different forms of CHW recruitment or incentives on programme outcomes). Future component selection experiments ought to fill these gaps and be used to refine existing theories of CHW performance (Kok et al., 2014; Naimoli et al., 2014).
4. As discussed above, randomised trials have a unique value in assessing intervention effects. That said, as has been pointed out elsewhere (Walt, 2005) these are nonetheless policy-relevant questions and ones RCTs do not answer (e.g. sustainability, transferability, financing). As such, further systematic reviews drawing on a range of study designs (e.g. process evaluations, ethnographic work) to answer such questions would complement this work.

7.5 Strengths and Weaknesses

7.5.1 Strengths

This dissertation has a number of strengths. The sequence and studies of the dissertation were based on MRC's framework for developing and evaluating complex interventions and the Multiphase Optimisation Strategy (Collins et al., 2011; Medical

Research Council, 2008). Both frameworks play an important role in advancing the relationship between evidence-based practice and community-based health delivery. This approach is important—iteratively incorporating existing evidence, theory, and component data into an intervention allows for the illumination of processes and mechanisms that might enable others to replicate the causal pathways responsible for their results.

Substantively, the thesis was focused on issues of urgent concern to a field attempting to meet ambitious international health goals and avoid repeating mistakes of the past: comprehensively identifying the tasks CHWs can perform safely and effectively and how their performance could be optimised. Guided by PRISMA and MECIR, this dissertation produced the most comprehensive and up-to-date systematic review of evidence of CHW efficacy and effectiveness thus far and was the first to apply these rigorous methods to systematically summarise evidence for CHW performance improvement (Chandler, Churchill, Higgins, Lasserson, & Tovey, 2013; Moher, Liberati, Tetzlaff, & Altman, 2009).

The dissertation was also the first to systematically summarise current methodological guidance for overviews. It exposed the degree to which overview guidance was heavily informed by consensus and personal experience and made a methodological contribution in the form of an overview of reviews decision tool that may help reduce the risk of bias in overviews in the future.

Furthermore, the dissertation answered repeated calls for a comprehensive taxonomy of CHW characteristics and programme features. Using systematic methods, the dissertation produced a taxonomy that meets high quality standards: it is jointly exhaustive, yet parsimonious and feasibly used to report the required information in published studies. Included components are observable and empirically

testable, conceptually unique, defined in a way that is accurate and easily understood, and grouped in a clear and meaningful manner. Finally, the taxonomy is hospitable, meaning that as innovations in CHW characteristics and programme features emerge they can be included in the taxonomy in the future.

7.5.2 Limitations

This thesis is nonetheless beset by limitations related both to its methods and substance. Firstly, complete double extraction and coding of the scoping review, two systematic reviews, and taxonomy construction would have helped minimise errors and reduce potential biases. Due to budgetary and time constraints as part of the DPhil work, this was not possible.

Secondly, the overview decision tool and taxonomy were constructed solely on the basis of existing literature. In the case of the taxonomy, this is compounded by the fact that RCTs were used as a source of primary data, despite not being the study design best suited to this task. Subjecting these tools to revision through a consensus process involving experts in field and practitioners may have (a) helped identify important issues not mentioned in the literature, (b) refined the categories and (c) ensured its relevance to and uptake by potential users of these tools.

Thirdly, randomised trials have a unique value in assessing intervention effects and the weaknesses of non-randomised studies have been described. Exclusion of all such studies without further consideration, however, places a zero weighting on evidence from non-randomised studies. Particularly in Chapter 6, where there were zero included studies for synthesis, it may have been instructive to include non-experimental studies. Updates of the two systematic reviews should consider including such evidence.

Fourthly, while international development databases (e.g. 3ie and J-PAL) and the grey literature were searched while undertaking both systematic reviews, it is likely that other non-health data (e.g. related to workplace interventions or human resource management) are applicable to the operational questions posed in Chapter 6. Referring to such sources in future may yield relevant insight.

Finally, while CHWs were once concerned with bringing about both access to care and a healthier distribution of power, as evidenced by this dissertation, their emphasis has shifted considerably to the former in recent decades. Yet, as several scholars have noted, the health of the people is just as—if not more—influenced by politics and power groups and by the distribution of land and wealth than it is by the treatment and prevention of disease (McKeown, 1976; Werner, 1981). By not engaging deeply with the dynamics of evidence translation and policymaking, this thesis does nothing to mitigate the potential far-reaching consequences of these forces.

7.6 Conclusion

Current interest in large-scale CHW programmes must be tempered with critical reflection on the disappointments that followed a similar wave of enthusiasm in the 1970s and 1980s (Crigler et al., 2013). CHW programmes are complex interventions that require immense effort to assess, optimise, scale and maintain.

Following the logic of early stage intervention development, this thesis had three objectives: (1) to assess the effects of CHW delivered interventions for primary health outcomes in LMICs, (2) to develop a comprehensive taxonomy of CHW characteristics and programme design features, and (3) to identify what interventions for improving the performance of community health workers, if any, may result in better primary healthcare outcomes in LMICs.

First, high quality evidence was found that CHW delivered interventions reduce perinatal mortality, improve child nutritional status, and improve tuberculosis completion rates versus facility-based care. There is also moderate quality evidence that CHW delivered interventions improve certain mental, infectious disease, paediatric, and maternal health outcomes. These findings build on and extend the findings of previous reviews. Practitioners and policymakers should consider expanding the remit of existing CHWs to include the areas.

Second, a faceted taxonomy with two categories (CHW Characteristics and CHW Programme Features) and six dimensions (Integration, Recruitment, Training, Supervision, Incentives, and Equipment) emerged. The definitional clarity provided by such a taxonomy allows for (a) consistent, precise, and complete reporting of relevant CHW-related information in trials and other studies, (b) the optimisation of CHW programmes through component selection experiments and systematic review subgroup analysis, (c) reliable evidence synthesis, and (d) more accurate replication of CHW programme models.

Third, while evidence was insufficient to specify which performance improvement intervention strategies are likely to be most effective, the review isolated four specific strategies related to CHW recruitment, training, supervision, and incentives that moderate quality evidence suggests improve behavioural outcomes for patients, CHW quality of care, and/or utilisation of health services. The complexity of these programs suggests that these strategies should, however, be interpreted with caution and considered in light of factors—discussed above—that may affect their efficacy if employed in different settings (Lavis et al. 2009).

While this thesis has made significant contributions to assessing and optimising CHW programmes, challenges in creating large-scale, sustainable public

sector CHW programmes remain. It is clear, however, that improvement in health facilities alone is not sufficient to avert a large proportion of deaths (Haines et al., 2007). Over one billion people around the world go their entire lives without seeing a health worker (Ghani, 2011; Moszynski, 2011) and it is typically the poorest and most at risk populations who remain unreached (Darmstadt et al., 2005). CHWs—with their unique ability to vastly extend access to high quality primary care—must cease to be regarded as a stop-gap measure in a second-rate health programme and begin to be understood as a central part of a first-class health system (Crigler et al., 2013). In bringing the tools of evidence based practice to bear on community health worker interventions, this dissertation—though only a beginning—has contributed to the theoretical, methodological, and empirical evidence base from which the field can continue to advance.

References

- Abbatt, F. (2005). Scaling up health and education workers: community health workers. London, UK: DFID Health Systems Resource Centre.
- Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychol*, 27(3), 379-387. doi: 10.1037/0278-6133.27.3.379
- Afenyadu, G., Agyepong, I., Barnish, G., & Adjei, S. (2005). Improving access to early treatment of malaria: a trial with primary school teachers as care providers. *Tropical Medicine & International Health*, 10(10), 1065-1072.
- Agrasada, G. V., Gustafsson, J., Kylberg, E., & Ewald, U. (2005). Postnatal peer counselling on exclusive breastfeeding of low-birthweight infants: a randomized, controlled trial. *Acta Paediatr*, 94(8), 1109-1115. doi: 10.1080/08035250510025752
- Ahmed, A., Salam, R. A., Lassi, Z. S., Das, J. K., Imama, N., & Bhutta, Z. A. (2014). Community based interventions for the prevention and control of tuberculosis. *Infectious Diseases of Poverty*, 3(27).
- Ahmed, A. M., Desta, A., Tekle, K., & Mweta, E. A. (1993). Pursuing better health care delivery at district level. *World Health Forum*, 14(14), 360-366.
- Akogun, O. B., Audu, Z., Weiss, M. G., Adelakun, A. O., Akoh, J. I., Akogun, M. K., . . . Kale, O. O. (2001). Community-directed treatment of onchocerciasis with ivermectin in Takum, Nigeria. *Trop Med Int Health*, 6(3), 232-243.
- Alam, K., Khan, J. A. M., & Walker, D. G. (2012a). Impact of dropout of female volunteer community health workers: An exploration in Dhaka urban slums. *BMC Health Services Research*, 12. doi: 10.1186/1472-6963-12-260
- Alam, K., Tasneem, S., & Oliveras, E. (2012b). Retention of female volunteer community health workers in Dhaka urban slums: a case-control study. *Health Policy and Planning*, 27(6), 477-486. doi: 10.1093/heapol/czr059
- Ali, B. S., Rahbar, M., Naeem, S., Gul, A., Mubeen, S., & Iqbal, A. (2003). The effectiveness of counselling on anxiety and depression by minimally trained counsellors: a randomised controlled trial. *American Journal of Psychotherapy*, 57(3).
- Altman, D. G., Deeks, J. J., & Sackett, D. L. (1998). Odds ratios should be avoided when events are common. *British Medical Journal*, 317(7168), 1318.
- Amouzou, A., Banda, B., Kachaka, W., Joos, O., Kanyuka, M., Hill, K., & Bryce, J. (2014a). Monitoring child mortality through community health worker reporting of births and deaths in Malawi: validation against a household mortality survey. *Plos One*, 9(2), e88939. doi: 10.1371/journal.pone.0088939
- Amouzou, A., Hazel, E., Shaw, B., Miller, N. P., Tafesse, M., Mekonnen, Y., . . . Black, R. E. (2016a). Effects of the integrated Community Case Management of Childhood Illness Strategy on Child Mortality in Ethiopia: A Cluster Randomized Trial. *American Journal of Tropical Medicine & Hygiene*, 94(3), 596-604. doi: 10.4269/ajtmh.15-0586
- Amouzou, A., Kanyuka, M., Hazel, E., Heidkamp, R., Marsh, A., Mleme, T., . . . Bryce, J. (2016b). Independent Evaluation of the integrated Community Case Management of Childhood Illness Strategy in Malawi Using a National Evaluation Platform Design. *American Journal of Tropical Medicine & Hygiene*, 94(3), 574-583. doi: 10.4269/ajtmh.15-0584

- Amouzou, A., Morris, S., Moulton, L. H., & Mukanga, D. (2014b). Assessing the impact of integrated community case management (iCCM) programs on child mortality: review of early results and lessons learned in sub-Saharan Africa. (Themed Issue: Current scientific evidence for integrated community case management (iCCM) in Africa: findings from the iCCM evidence symposium.). *Journal of Global Health*, 4(2).
- Amuron, B., Levin, J., Birunghi, J., Namara, G., Coutinho, A., Grosskurth, H., & Jaffar, S. (2011). Mortality in an antiretroviral therapy programme in Jinja, south-east Uganda: a prospective cohort study. *AIDS Res Ther*, 8, 39. doi: 10.1186/1742-6405-8-39
- Anderson, L. M., Oliver, S. R., Michie, S., Rehfuss, E., Noyes, J., & Shemilt, I. (2013a). Investigating complexity in systematic reviews of interventions by using a spectrum of methods. *Journal of Clinical Epidemiology*, 66(11), 1223-1229. doi: 10.1016/j.jclinepi.2013.06.014
- Anderson, L. M., Petticrew, M., Chandler, J., Grimshaw, J., Tugwell, P., O'Neill, J., . . . Shemilt, I. (2013b). Introducing a series of methodological articles on considering complexity in systematic reviews of interventions. *Journal of Clinical Epidemiology*, 66(11), 1205.
- Andersson, N., Cockcroft, A., Ansari, N. M., Omer, K., Baloch, M., Ho Foster, A., . . . Soberanis, J. L. (2009). Evidence-based discussion increases childhood vaccination uptake: a randomised cluster controlled trial of knowledge translation in Pakistan. *BMC Int Health Hum Rights*, 9 Suppl 1, S8. doi: 10.1186/1472-698X-9-S1-S8
- Andreoni, J., Callen, M., Khan, Y., Jaffar, K., & Sprenger, C. (2016). *Using Preference Estimates To Customize Incentives: An Application To Polio Vaccination Drives In Pakistan*. NBER Working Paper Series. Retrieved from <http://www.nber.org/papers/w22019>
- Anyangwe, S. C., & Mtonga, C. (2007). Inequities in the global health workforce: the greatest impediment to health in sub-Saharan Africa. *International Journal of Environmental Research and Public Health*, 4(2), 93-100.
- Aromataris, E., Fernandez, R., Godfrey, C., Holly, C., Khalil, H., & Tungpunkom, P. (2014). Methodology for JBI umbrella reviews. *The Joanna Briggs Institute Reviewers Manual*, 5-34.
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132-140.
- Arrowsmith, M. E., Aicken, C. R., Saxena, S., & Majeed, A. (2012). Strategies for improving the acceptability and acceptance of the copper intrauterine device. *Cochrane Database of Systematic Reviews*, 3, CD008896.
- Arshad, A., Salam, R. A., Lassi, Z. S., Das, J. K., Naqvi, I., & Bhutta, Z. A. (2014). Community based interventions for the prevention and control of tuberculosis. *Infectious Diseases of Poverty*, 3, 27.
- Ashcroft, R. E. (2004). Current epistemological problems in evidence based medicine. *Journal of medical ethics*, 30(2), 131-135. doi: 10.1136/jme.2003.007039
- Ashraf, N., Bandiera, O., & Jack, B. K. (2014). No margin, no mission? A field experiment on incentives for public service delivery. *Journal of Public Economics*, 120, 1-17.
- Ashraf, N., Bandiera, O., & Lee, S. S. (2015). *Do-Gooders And Go-Getters: Career Incentives, Selection, And Performance In Public Service Delivery*.

- Aung, T., Montagu, D., Khin, H. S. S., Win, Z., San, A. K., & McFarland, W. (2014). Impact of a social franchising program on uptake of oral rehydration solution plus zinc for childhood diarrhea in myanmar: A community-level randomized controlled trial. *Journal of Tropical Pediatrics*, 60(3), 189-197. doi: <http://dx.doi.org/10.1093/tropej/fmt108>
- Avula, R., Frongillo, E. A., Arabi, M., Sharma, S., & Schultink, W. (2011). Enhancements to nutrition program in Indian integrated child development services increased growth and energy intake of children. *J Nutr*, 141(4), 680-684. doi: 10.3945/jn.109.116954
- Ayele, F., Desta, A., & Larson, C. (1993). The Functional Status Of Community-Health Agents - A Trial Of Refresher Courses And Regular Supervision. *Health Policy and Planning*, 8(4), 379-384. doi: 10.1093/heapol/8.4.379
- Azad, K., Barnett, S., Banerjee, B., Shaha, S., Khan, K., Rego, A. R., . . . Costello, A. (2010). Effect of scaling up women's groups on birth outcomes in three rural districts in Bangladesh: a cluster-randomised controlled trial. *Lancet*, 375(9721), 1193-1202.
- Baawo, S., Ballard, M., Slagle, J., Mukherjee, S., Johnson, A., Munah, A., . . . Griffiths, T. (2014). *Reaching the unreached: Improving the performance of community health volunteers and the district health system to increase health service delivery in hard-to-reach communities in rural Liberia—the Konobo experience*. Paper presented at the 142nd APHA Annual Meeting and Exposition (November 15-November 19, 2014).
- Bailey, J. E., & Coombs, D. W. (1996). Effectiveness of an Indonesian model for rapid training of Guatemalan health workers in diarrhea case management. *Journal of Community Health*, 21(4), 269-276. <http://onlinelibrary.wiley.com/doi/10.1002/jc.10014>
- Baker, P. R., Costello, J. T., Dobbins, M., & E, B. W. (2014). The benefits and challenges of conducting an overview of systematic reviews in public health: a focus on physical activity. *Journal of Public Health*, 36(3), 517-521. doi: 10.1093/pubmed/fdu050
- Baker-Henningham, H., Powell, C., Walker, S., & Grantham-McGregor, S. (2005). The effect of early stimulation on maternal depression: a cluster randomised controlled trial. *Archives of disease in childhood*, 90(12), 1230-1234.
- Baqui, A. H., El-Arifeen, S., Darmstadt, G. L., Ahmed, S., Williams, E. K., Seraji, H. R., . . . Projahnmo Study, G. (2008). Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *Lancet*, 371(9628), 1936-1944.
- Baqui, A. H., Williams, E., El-Arifeen, S., Applegate, J. A., Mannan, I., Begum, N., . . . Projahnmo Study Group in, B. (2016). Effect of community-based newborn care on cause-specific neonatal mortality in Sylhet district, Bangladesh: findings of a cluster-randomized controlled trial. *J Perinatol*, 36(1), 71-76. doi: 10.1038/jp.2015.139
- Bari, A., Sadruddin, S., Khan, A., Khan, I., Khan, A., Lehri, I. A., . . . Qazi, S. A. (2011). Community case management of severe pneumonia with oral amoxicillin in children aged 2-59 months in Haripur district, Pakistan: a cluster randomised trial. *Lancet*, 378(9805), 1796-1803.
- Bari, S., Mannan, I., Rahman, M., Darmstadt, G., Seraji, M. H., Baqui, A., . . . Winch, P. (2006). Trends in Use of Referral Hospital Services for Care of Sick

- Newborns in a Community-based Intervention in Tangail District, Bangladesh. *J Health Popul Nutr*, 24(4), 519-529.
- Basch, C. E., Sliepeceich, E. M., Gold, R. S., Duncan, D. F., & Kolbe, L. J. (1985). Avoiding type III errors in health education program evaluations: a case study. *Health Education & Behavior*, 12(3), 315-331.
- Bass, J., Neugebauer, R., Clougherty, K. F., Verdeli, H., Wickramaratne, P., Ndogoni, L., . . . Bolton, P. (2006). Group interpersonal psychotherapy for depression in rural Uganda: 6-month outcomes: randomised controlled trial. *Br J Psychiatry*, 188, 567-573. doi: 10.1192/bjp.188.6.567
- Becker, L. (2006). *Umbrella reviews: what are they and do we need them?* Paper presented at the XVI Cochrane Colloquium, Dublin, Ireland.
- Becker, L., & Oxman, A. (2011). Chapter 22: Overviews of reviews. In J. Higgins & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. Version 5.1.0 (updated March 2011)). www.cochrane-handbook.org: The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org.
- Bee, P., Bower, P., Byford, S., Churchill, R., Calam, R., Stallard, P., . . . Abel, K. (2014). The clinical effectiveness, cost-effectiveness and acceptability of community-based interventions aimed at improving or maintaining quality of life in children of parents with serious mental illness: a systematic review. *NIHR HTA programme*.
- Begum, F., Bhuiyan, A., & Asiruddin, S. (2009). Community-based postnatal care study in Bangladesh. *International Journal of Gynecology & Obstetrics*, 107, S123.
- Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3), 489-520.
- Bennett, F. (1979). Primary health care and developing countries. *Social Science & Medicine*, 13, 505-514.
- Bensaude-Vincent, B., & Stengers, I. (1996). *A history of chemistry*: Harvard University Press.
- Berman, P. A., Gwatkin, D. R., & Burger, S. E. (1987). Community-Based Health-Workers - Head-Start or False Start Towards Health for All. *Social Science & Medicine*, 25(5), 443-459. doi: Doi 10.1016/0277-9536(87)90168-7
- Bertrand, J. T., Mangani, N., Mansilu, M., McBride, M. E., & Tharp, J. L. (1986). Strategies for family planning service delivery in Bas Zaire. *International Family Planning Perspectives*, 108-115.
- Betancourt, T. S., Newnham, E. A., Brennan, R. T., Verdeli, H., Borisova, I., Neugebauer, R., . . . Bolton, P. (2012). Moderators of treatment effectiveness for war-affected youth with depression in northern Uganda. *J Adolesc Health*, 51(6), 544-550. doi: 10.1016/j.jadohealth.2012.02.010
- Bhatt, A., Maru, R., & Prabhakar, A. S. (1981). Socio-economic profile and performance of CHVs in Gujarat: implications for selection and support. *Health Popul Perspect Issues*, 4(3), 181-198.
- Bhattacharyya, K. (2001). *Community Health Worker Incentives and Disincentives: How They Affect Motivation, Retention and Sustainability*. Viginia, USA: USAID Basic Support for Institutionalizing Child Survival Project (BASICS II), .
- Bhutta, Z., Lassi, Z. S., Pariyo, G., & Huicho, L. (2010). Global experience of community health workers for delivery of health related millennium development goals: a systematic review, country case studies, and

- recommendations for integration into national health systems. *Global Health Workforce Alliance*.
- Bhutta, Z., Memon, Z., Soofi, S., Salat, M., Cousens, S., & Martines, J. (2008a). Implementing community-based perinatal care: results from a pilot study in rural Pakistan. *Bulletin of the World Health Organization*, 86(6), 452-459. <http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/366/CN-00640366/frame.html>
- Bhutta, Z. A., Ali, S., Cousens, S., Ali, T. M., Haider, B. A., Rizvi, A., . . . Black, R. E. (2008b). Alma-Ata: Rebirth and Revision 6 Interventions to address maternal, newborn, and child survival: what difference can integrated primary health care strategies make? *Lancet*, 372(9642), 972-989. doi: [http://dx.doi.org/10.1016/S0140-6736\(08\)61407-5](http://dx.doi.org/10.1016/S0140-6736(08)61407-5)
- Bigirwa, P. (2009). Effectiveness of community health workers (CHWS) in the provision of basic preventive and curative maternal, newborn and child health (MNCH) interventions: A systematic review. *Health Policy and Development*, 7(3), 162-172.
- Birungi, N., Fadnes, L. T., Okullo, I., Kasangaki, A., Nankabirwa, V., Ndeezi, G., . . . Astrom, A. N. (2015). Effect of Breastfeeding Promotion on Early Childhood Caries and Breastfeeding Duration among 5 Year Old Children in Eastern Uganda: A Cluster Randomized Trial. *Plos One*, 10(5), e0125352. doi: [10.1371/journal.pone.0125352](https://doi.org/10.1371/journal.pone.0125352)
- Bishai, D., Sachatp, K., LeFevre, A., Thant, H. N., Zaw, M., Aung, T., . . . Social Franchising Research, T. (2015). Cost-effectiveness of using a social franchise network to increase uptake of oral rehydration salts and zinc for childhood diarrhea in rural Myanmar. *Cost Eff Resour Alloc*, 13, 3. doi: [10.1186/s12962-015-0030-3](https://doi.org/10.1186/s12962-015-0030-3)
- Björkman, M., & Svensson, J. (2009). Power To The People: Evidence From A Randomized Field Experiment On Community-Based Monitoring In Uganda. *Quarterly Journal of Economics*, 124(2), 735-769.
- Black, N. (2001). Evidence based policy: proceed with care. *British Medical Journal*, 323, 275-279.
- Blanas, D., & Ndiaye, Y. (2010). Community based malaria control in Saraya, Southeastern Senegal. *61st Annual Meeting of the American Society of Tropical Medicine and Hygiene*.
- Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay*, 20-24.
- Bojang, K. A., Akor, F., Conteh, L., Webb, E., Bittaye, O., Conway, D. J., . . . Greenwood, B. (2011). Two strategies for the delivery of IPTc in an area of seasonal malaria transmission in the Gambia: a randomised controlled trial. *PLoS Medicine / Public Library of Science*, 8(2), e1000409.
- Bolton, P., Bass, J., Betancourt, T., Speelman, L., Onyango, G., Clougherty, K., . . . Verdelli, H. (2007). Interventions for Depression Symptoms Among Adolescent Survivors of War and Displacement in Northern Uganda. *Journal of the American Medical Association*, 298(5).
- Bolton, P., Bass, J., Neugebauer, R., Verdelli, H., Clougherty, K. F., Wickramaratne, P., . . . Weissman, M. (2003). Group interpersonal psychotherapy for depression in rural Uganda: a randomized controlled trial. *Journal of the American Medical Association*, 289(23), 3117-3124. doi: [10.1001/jama.289.23.3117](https://doi.org/10.1001/jama.289.23.3117)

- Bonell, C., Hargreaves, J., Strange, V., Pronyk, P., & Porter, J. (2006). Should structural interventions be evaluated using RCTs? The case of HIV prevention. *Social Science & Medicine*, 63(5), 1135-1142. doi: 10.1016/j.socscimed.2006.03.026
- Boone, P., Mann, V., Eble, A., Mendiratta, T., Mukherjee, R., Figueiredo, R., . . . Elbourne, D. (2007). Community health and medical provision: impact on neonates (the CHAMPION trial). *BMC Pediatrics*, 7, 26. doi: 10.1186/1471-2431-7-26
- Borghi, J., Thapa, B., Osrin, D., Jan, S., Morrison, J., Tamang, S., . . . Costello, A. M. (2005). Economic assessment of a women's group intervention to improve birth outcomes in rural Nepal. *Lancet*, 366(9500), 1882-1884. doi: 10.1016/S0140-6736(05)67758-6
- Borrelli, B. (2011). The Assessment, Monitoring, and Enhancement of Treatment Fidelity In Public Health Clinical Trials. *J Public Health Dent*, 71(s1), S52-S63. doi: 10.1111/j.1752-7325.2011.00233.x
- Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., . . . Orwig, D. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol*, 73(5), 852-860. doi: 10.1037/0022-006X.73.5.852
- Bossuroy, T., Delavallade, C., & Pons, V. (2016). Fighting tuberculosis through community based counsellors: a randomized evaluation of performance based incentives in India *3ie Grantee Final Report*. New Delhi: International Initiative for Impact Evaluation (3ie).
- Bothwell, L. E., Greene, J. A., Podolsky, S. H., & Jones, D. S. (2016). Assessing the Gold Standard — Lessons from the History of RCTs. *New England Journal of Medicine*, 374(22), 2175-2181. doi: doi:10.1056/NEJMms1604593
- Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine*, 148(4), 295-309.
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research*, 8(1), 137-152.
- Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: developing taxonomy, themes, and theory. *Health Services Research*, 42(4), 1758-1772.
- Braun, R., Catalani, C., Wimbush, J., & Israelski, D. (2013). Community Health Workers and Mobile Technology: A Systematic Review of the Literature. *Plos One*, 8(6). doi: 10.1371/journal.pone.0065772
- Brieger, W. (2000). Implementation And Sustainability Of Community-Directed Treatment Of Onchocerciasis With Ivermectin: A Multicountry Study (pp. 62). Geneva.
- Brownson, R. C., Baker, E. A., Leet, T. L., & Gillespie, K. N. (2003). *Evidence-based public health*. New York: Oxford University Press.
- Brownson, R. C., Baker, E. A., Leet, T. L., Gillespie, K. N., & True, W. R. (2010). *Evidence-based public health*: Oxford University Press.
- Brownson, R. C., Fielding, J. E., & Maylahn, C. M. (2009). Evidence-based public health: a fundamental concept for public health practice. *Annual Review of Public Health*, 30, 175-201.
- Brownson, R. C., Gurney, J. G., & Land, G. H. (1999). Evidence-based decision making in public health. *Journal of Public Health Management and Practice*, 5(5), 86-97.

- Bruxvoort, K., Goodman, C., Kachur, S. P., & Schellenberg, D. (2014). How patients take malaria treatment: a systematic review of the literature on adherence to antimalarial drugs. *Plos One*, 9(1), e84555.
- Bryant, J. (1969). *Health and the developing world*. London, UK: Cornell University Press.
- Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., & Klassen, T. P. (2006). Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology*, 59(7), 697-703.
- Byrne, A., Hodge, A., Jimenez-Soto, E., & Morgan, A. (2014). What works? Strategies to increase reproductive, maternal and child health in difficult to access mountainous locations: a systematic literature review. *Plos One*, 9(2), e87683. doi: 10.1371/journal.pone.0087683
- Caird, J., Sutcliffe, K., Kwan, I., Dickson, K., & Thomas, J. (2015). Mediating policy-relevant evidence at speed: Are systematic reviews of systematic reviews a useful approach? *Evidence & Policy: A Journal of Research, Debate and Practice*, 11(1), 81-97.
- Caldwell, D. M., Welton, N. J., & Ades, A. E. (2010). Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *Journal of Clinical Epidemiology*, 63(8), 875-882. doi: 10.1016/j.jclinepi.2009.08.025
- Callaghan, M., Ford, N., & Schneider, H. (2010). A systematic review of task-shifting for HIV treatment and care in Africa. *Human Resources for Health*, 8, 8. doi: 10.1186/1478-4491-8-8
- Campbell, C., & Scott, K. (2011). Retreat from Alma Ata? The WHO's report on Task Shifting to community health workers for AIDS care in poor countries. *Global Public Health*, 6(2), 125-138. doi: 10.1080/17441690903334232
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24(4), 409.
- Campbell, J., Dussault, G., Buchan, J., Pozo-Martin, F., Arias, M. G., & Leone, M. (2013). A universal truth: No health without a workforce: Third global forum on human resources for health, Recife, Brazil. *Global Health Workforce Alliance and World Health Organization, Geneva*.
- Campbell, M., Fitzpatrick, R., Haines, A., & Kinmonth, A. L. (2000a). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal*, 321(7262), 694.
- Campbell, M., Grimshaw, J., & Steen, N. (2000b). Sample size calculations for cluster randomised trials. *Journal of health services research & policy*, 5(1), 12-16.
- Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., . . . Kinmonth, A. L. (2007). Designing and evaluating complex interventions to improve health care. *BMJ (Clinical research ed.)*, 334(7591), 455-459. doi: 10.1136/bmj.39108.379965.BE
- Carrón, J., Melian, M., & Leon, F. (1994). Final report of an operations research project: "Developing Tools of Low-Cost Use to Improve the Quality of Care of Rural CBD in Paraguay" Subcontract CI91. 96A January 1 1992-July 31 1994. Participating agencies: Centro Paraguayo de Estudios de Poblacion (CEPEP) the Population Council/INOPAL II.
- Carrón, J., Melián, M., & León, F. (1993). Desarrollando instrumentos de bajo costo para mejorar la calidad de los servicios rurales de planificación familiar: informe final (Enero 1992 - Diciembre 1993). Asunción, Paraguay: Centro Paraguayo de Estudios de Poblacion (CEPEP).

- Catterall, C. D. (1970). Taxonomy of Prescriptive Interventions. *Journal of School Psychology*, 8(1), 5-12. doi: Doi 10.1016/0022-4405(70)90047-6
- CDI Study Group. (2010). Community-directed interventions for priority health problems in Africa: results of a multicountry study. *Bulletin of the World Health Organization*, 88(7), 509-518. doi: 10.2471/BLT.09.069203
- Chalmers, I. (2003). Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. *The Annals of the American Academy of Political and Social Science*, 589(1), 22-40. doi: 10.1177/0002716203254762
- Chalmers, I. (2005). If evidence-informed policy works in practice, does it matter if it doesn't work in theory? *Evidence & Policy : A Journal of Research, Debate and Practice*, 1(2), 227-242.
- Chandler, J., Churchill, R., Higgins, J., Lasserson, T., & Tovey, D. (2013). Methodological standards for the conduct of new Cochrane Intervention Reviews (December).
- Chang, L. W., Kagaayi, J., Arem, H., Nakigozi, G., Ssempijja, V., Serwadda, D., . . . Reynolds, S. (2011). Impact of a mHealth Intervention for Peer Health Workers on AIDS Care in Rural Uganda: A Mixed Methods Evaluation of a Cluster-Randomized Trial. *AIDS and Behavior*, 15(8), 1776-1784. doi: 10.1007/s10461-011-9995-x
- Chang, L. W., Kagaayi, J., Nakigozi, G., Serwada, D., Quinn, T. C., Gray, R. H., . . . Holtgrave, D. (2013). Cost analyses of peer health worker and mHealth support interventions for improving AIDS care in Rakai, Uganda. *AIDS Care*, 25(5), 652-656. doi: 10.1080/09540121.2012.722600
- Chang, L. W., Kagaayi, J., Nakigozi, G., Ssempijja, V., Packer, A. H., Serwadda, D., . . . Reynolds, S. J. (2010a). Effect of peer health workers on AIDS care in Rakai, Uganda: a cluster-randomized trial. *Plos One*, 5(6), e10923.
- Chang, L. W., Nakigozi, G., Billioux, V. G., Gray, R. H., Serwadda, D., Quinn, T. C., . . . Reynolds, S. J. (2015). Effectiveness of peer support on care engagement and preventive care intervention utilization among pre-antiretroviral therapy, HIV-infected adults in Rakai, Uganda: a randomized trial. *AIDS and Behavior*, 19(10), 1742-1751. doi: 10.1007/s10461-015-1159-y
- Chang, S. M., Walker, S. P., Grantham-McGregor, S., & Powell, C. A. (2002). Early childhood stunting and later behaviour and school achievement. *J Child Psychol Psychiatry*, 43(6), 775-783.
- Chang, S. M., Walker, S. P., Grantham-McGregor, S., & Powell, C. A. (2010b). Early childhood stunting and later fine motor abilities. *Dev Med Child Neurol*, 52(9), 831-836. doi: 10.1111/j.1469-8749.2010.03640.x
- Chatterjee, S., Naik, S., John, S., Dabholkar, H., Balaji, M., Koschorke, M., . . . Thornicroft, G. (2014). Effectiveness of a community-based intervention for people with schizophrenia and their caregivers in India (COPSI): a randomised controlled trial.[Erratum appears in Lancet. 2014 Jun 14;383(9934):2046]. *Lancet*, 383(9926), 1385-1394.
- Chen, L., Evans, T., Anand, S., Boufford, J. I., Brown, H., Chowdhury, M., . . . Elzinga, G. (2004). Human resources for health: overcoming the crisis. *Lancet*, 364(9449), 1984-1990.
- Chevalier, C., Lapo, A., O'Brien, J., & Wierzba, T. F. (1993). *Why do village health workers drop out*. Paper presented at the World Health Forum.
- Child, D. (2006). *The essentials of factor analysis*: A&C Black.

- Chin-Quee, D., Mugeni, C., Nkunda, D., Uwizeye, M., Stockton, L., & Wesson, J. (2016). Balancing workload, motivation and job satisfaction in Rwanda: assessing the effect of adding family planning service provision to community health worker duties. *Reproductive Health*, 13(2).
- Chinbuah, M. A., Adjuik, M., Cobelens, F., Koram, K. A., Abbey, M., Gyapong, M., . . . Gyapong, J. O. (2013). Impact of treating young children with antimalarials with or without antibiotics on morbidity: a cluster-randomized controlled trial in Ghana. *Int Health*, 5(3), 228-235. doi: 10.1093/inthealth/ih021
- Chinbuah, M. A., Kager, P. A., Abbey, M., Gyapong, M., Awini, E., Nonvignon, J., . . . Gyapong, J. O. (2012). Impact of community management of fever (using antimalarials with or without antibiotics) on childhood mortality: a cluster-randomized controlled trial in Ghana. *American Journal of Tropical Medicine & Hygiene*, 87(5 Suppl), 11-20. doi: 10.4269/ajtmh.2012.12-0078
- Chola, L., Nkonki, L., Kankasa, C., Nankunda, J., Tumwine, J., Tylleskar, T., . . . Study Group, P.-E. (2011). Cost of individual peer counselling for the promotion of exclusive breastfeeding in Uganda. *Cost Eff Resour Alloc*, 9(1), 11. doi: 10.1186/1478-7547-9-11
- Chongsuvivatwong, V., Mo-Suwan, L., Tayakkanonta, K., Vitsupakorn, K., & McNeil, R. (1996). Impacts of training of village health volunteers in reduction of morbidity from acute respiratory infections in childhood in southern Thailand. *Southeast Asian Journal of Tropical Medicine & Public Health*, 27(2), 333-338.
- Christopher, J. B., May, A., Lewin, S., & Ross, D. A. (2011). Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa (Provisional abstract). *Human Resources for Health*, 9(1). <http://onlinelibrary.wiley.com/doi/10.1111/j.1469-7610.2011.02512.x>
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annu Rev Psychol*, 62, 73-101. doi: 10.1146/annurev.psych.093008.100427
- Clarke, K., Azad, K., Kuddus, A., Shaha, S., Nahar, T., Aumon, B. H., . . . Fottrell, E. (2014). Impact of a participatory intervention with women's groups on psychological distress among mothers in rural Bangladesh: secondary analysis of a cluster-randomised controlled trial. *Plos One*, 9(10), e110697. doi: 10.1371/journal.pone.0110697
- Clarke, M. (2005). Towards cost-effective tuberculosis control in the Western Cape of South Africa: intervention study involving lay health workers on agricultural farms. *Towards cost effective tuberculosis control in the Western Cape of South Africa: intervention study involving lay health workers on agricultural farms*.
- Clarke, M., Dick, J., & Bogg, L. (2006). Cost-effectiveness analysis of an alternative tuberculosis management strategy for permanent farm dwellers in South Africa amidst health service contraction. *Scandinavian Journal of Public Health*, 34(1), 83-91.
- Clarke, M., Dick, J., Zwarenstein, M., Lombard, C. J., & Diwan, V. K. (2005). Lay health worker intervention with choice of DOT superior to standard TB care for farm dwellers in South Africa: a cluster randomised control trial.[Erratum

- appears in *Int J Tuberc Lung Dis*. 2009 Jul;13(7):921]. *The International Journal of Tuberculosis and Lung Disease*, 9(6), 673-679.
- Cleary, M., & Hunt, G. E. (2010). The citation game. *Int J Ment Health Nurs*, 19(4), 221-222. doi: 10.1111/j.1447-0349.2010.00685.x
- Cochrane, A. L. (1972). *Effectiveness and efficiency: random reflections on health services* (Vol. 900574178). London: Nuffield Provincial Hospitals Trust.
- Cochrane Collaboration Qualitative & Implementation Methods Group. (2016). Cochrane Qualitative & Implementation Methods Group from <http://methods.cochrane.org/qi/>
- Cochrane Comparing Multiple Interventions Methods Group. (2012). *Review Type & Methodological Considerations: Background Paper for the First Part of the Paris CMIMG Discussion*. Paper presented at the Paris CMIMG Discussion, Paris, France.
- Cochrane Comparing Multiple Interventions Methods Group. (2013). *Undertaking, publishing and maintaining Cochrane Reviews that compare multiple interventions*.
- Cochrane Empty Reviews Project. (2013). The Cochrane Empty Reviews Project. Retrieved May 10, 2016, from <https://empty-reviews.org>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York, USA: Academic Press.
- Colbourn, T., Nambiar, B., Bondo, A., Makwenda, C., Tsetekani, E., Makonda-Ridley, A., . . . Costello, A. (2013a). Effects of quality improvement in health facilities and community mobilization through women's groups on maternal, neonatal and perinatal mortality in three districts of Malawi: MaiKhanda, a cluster randomized controlled effectiveness trial. *Int Health*, 5(3), 180-195. doi: 10.1093/inthealth/ih011
- Colbourn, T., Nambiar, B., Costello, A., & MaiKhanda. (2013b). Final Evaluation Report. The Impact of Quality Improvement at Health Facilities and Community Mobilisation by Women's Groups on Birth Outcomes: an Effectiveness Study in Three Districts Of Malawi. London: The Health Foundation.
- Colbourn, T., Pulkki-Brannstrom, A. M., Nambiar, B., Kim, S., Bondo, A., Banda, L., . . . Skordis-Worrall, J. (2015). Cost-effectiveness and affordability of community mobilisation through women's groups and quality improvement in health facilities (MaiKhanda trial) in Malawi. *Cost Eff Resour Alloc*, 13(1), 1. doi: 10.1186/s12962-014-0028-2
- Collins, L. M., Baker, T. B., Mermelstein, R. J., Piper, M. E., Jorenby, D. E., Smith, S. S., . . . Fiore, M. C. (2011). The multiphase optimization strategy for engineering effective tobacco use interventions. *Ann Behav Med*, 41(2), 208-226. doi: 10.1007/s12160-010-9253-x
- Conn, V. S., & Coon Sells, T. G. (2014). WJNR welcomes umbrella reviews. *Western Journal of Nursing Research*, 36(2), 147-151. doi: 10.1177/0193945913506968
- Cooper, H., & Koenka, A. C. (2012). The overview of reviews: unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *Am Psychol*, 67(6), 446-462.
- Cooper, P. J., Tomlinson, M., Swartz, L., Landman, M., Molteno, C., Stein, A., . . . Murray, L. (2009). Improving quality of mother-infant relationship and infant attachment in socioeconomically deprived community in South Africa: randomised controlled trial. *British Medical Journal*, 338, b974.

- Corluka, A., Walker, D. G., Lewin, S., Glenton, C., & Scheel, I. B. (2009). Are vaccination programmes delivered by lay health workers cost-effective? A systematic review. *Human Resources for Health*, 7, 81.
- Coutinho, S. B., de Lira, P. I. C., de Carvalho Lima, M., & Ashworth, A. (2005). Comparison of the effect of two systems for the promotion of exclusive breastfeeding. *Lancet*, 366(9491), 1094-1100. doi: 10.1016/s0140-6736(05)67421-1
- Craig, P., Dieppe, P., Macintyre, S., & Michie, S. (2008). Developing and Evaluating Complex Interventions: New Guidance. *British Medical Journal*, 337, a1655. doi: 10.1136/bmj.a1655
- Creanga, A. A., Bradley, H. M., Kidanu, A., Melkamu, Y., & Tsui, A. O. (2007). Does the delivery of integrated family planning and HIV/AIDS services influence community-based workers client loads in Ethiopia. *Health Policy and Planning*, 22(6), 404-414. doi: 10.1093/heapol/czm034
- Crigler, L., Glenton, C., Hodgins, S., LeBan, K., Lewin, S., Perry, H., . . . Jaskiewicz, W. (2013). Developing and strengthening community health worker programs at scale. A reference guide for program managers and policy makers. Draft. In USAID Maternal and Child Health Integrated Program (MCHIP) (Ed.).
- Crigler, L., & Hill, K. (2009). Rapid Assessment of Community Health Worker Programs in USAID Priority MCH Countries Draft Tool for Field Testing *USAID Health Care Improvement Project*.
- Crigler, L., Hill, K., Furth, R., & Bjerregaard, D. (2011). Community Health Worker Assessment and Improvement Matrix (CHW AIM): A Toolkit for Improving CHW Programs and Services *USAID Healthcare Improvement Project*.
- Crisp, N., Gawanas, B., & Sharp, I. (2008). Training the health workforce: scaling up, saving lives. *Lancet*, 371(9613), 689-691.
- Crispin, N., Wamae, A., Ndirangu, M., Wamalwa, D., Wangalwa, G., Watako, P., & Mbiti, E. (2012). Effects of selected socio-demographic characteristics of community health workers on performance of home visits during pregnancy: a cross-sectional study in Busia District, Kenya. *Glob J Health Sci*, 4(5), 78-90. doi: 10.5539/gjhs.v4n5p78
- Cueto, M. (2004). The origins of primary health care and selective primary health care. *American Journal of Public Health*, 94(11), 1864-1874.
- Curtale, F., Siwakoti, B., Lagrosa, C., LaRaja, M., & Guerra, R. (1995). Improving skills and utilization of community health volunteers in Nepal. *Social Science & Medicine*, 40(8), 1117-1125.
- Dahn, B., Woldemariam, A., Perry, H., Maeda, A., von Glahn, D., Panjabi, R., . . . Lu, C. (2015). Strengthening primary health care through community health workers: Investment case and financing recommendations.
- Daly, J., Willis, K., Small, R., Green, J., Welch, N., Kealy, M., & Hughes, E. (2007). A hierarchy of evidence for assessing qualitative health research. *Journal of Clinical Epidemiology*, 60(1), 43-49.
- Daniels, K., Odendaal, W. A., Nkonki, L., Hongoro, C., Colvin, C. J., & Lewin, S. (2014). Incentives for lay health workers to improve recruitment, retention in service and performance. *The Cochrane Library*.
- Darmstadt, G. L., Bhutta, Z. A., Cousens, S., Adam, T., Walker, N., de Bernis, L., & Lancet Neonatal Survival Steering Team. (2005). Evidence-based, cost-effective interventions: how many newborn babies can we save? *Lancet*, 365(9463), 977-988.

- Darmstadt, G. L., Choi, Y., Arifeen, S. E., Bari, S., Rahman, S. M., Mannan, I., . . . Bangladesh Projahnmo-2 Mirzapur Study, G. (2010). Evaluation of a cluster-randomized controlled trial of a package of community-based maternal and newborn interventions in Mirzapur, Bangladesh. *Plos One*, 5(3), e9696.
- Darmstadt, G. L., Lee, A. C., Cousens, S., Sibley, L., Bhutta, Z. A., Donnay, F., . . . Lawn, J. E. (2009). 60 Million non-facility births: who can deliver in community settings to reduce intrapartum-related deaths? *International Journal of Gynaecology & Obstetrics*, 107 Suppl 1, S89-112.
- Das, A., Friedman, J., Kandpal, E., Ramana, G., Gupta, R., das, . . . Ramesh, G. (2014). Strengthening malaria service delivery through supportive supervision and community mobilization in an endemic Indian setting: an evaluation of nested delivery models. *Malaria Journal*, 13(482).
- Das, J. K., Lassi, Z. S., Salam, R. A., & Bhutta, Z. A. (2013). Effect of community based interventions on childhood diarrhea and pneumonia: uptake of treatment modalities and impact on mortality. *BMC Public Health*, 13 Suppl 3, S29. doi: 10.1186/1471-2458-13-S3-S29
- Datiko, D. G., & Lindtjorn, B. (2009). Health extension workers improve tuberculosis case detection and treatment success in southern Ethiopia: a community randomised trial. *Plos One*, 4(5), e5443.
- Datiko, D. G., & Lindtjorn, B. (2010). Cost and cost-effectiveness of smear-positive tuberculosis treatment by Health Extension Workers in Southern Ethiopia: a community randomized trial. *Plos One*, 5(2), e9158.
- Datta, J., & Petticrew, M. (2013). Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Public Health*, 13, 568. doi: 10.1186/1471-2458-13-568
- Davies, P. (2015). Getting evidence into policy and practice. *Journal of Development Effectiveness*, 7(4), 393-401.
- Davies-Adetugbo, A., Fabiyi, A., Ojoofetimi, E., & Adetugbo, K. (1997). Breastfeeding training improves health worker performance in rural Nigeria. *East African Medical Journal*, 74(8), 510-513.
<http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/097/CN-00148097/frame.html>
- De La Torre, S. E. C., Ly, P., Nguon, C., Arantxa, R. F., Sintasath, D., Whittaker, M., & Singhasivanon, P. (2012). *Evaluation of community malaria worker performance in western cambodia: A quantitative and qualitative assessment*. Paper presented at the 61st Annual Meeting of the American Society of Tropical Medicine and Hygiene, Atlanta, Georgia.
- De Silva, M. J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C., & Patel, V. (2014). Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials*, 15(1), 1.
- De Zoysa, I., & Cole-King, S. (1983). *Remuneration of the community health worker: what are the options?* Paper presented at the World Health Forum.
- Decroo, T., Rasschaert, F., Telfer, B., Remartinez, D., Laga, M., & Ford, N. (2013). Community-based antiretroviral therapy programs can overcome barriers to retention of patients and decongest health services in sub-Saharan Africa: a systematic review. *International Health*, 5(3), 169-179.
- DeJong, G., Horn, S. D., Gassaway, J. A., Slavin, M. D., & Dijkers, M. P. (2004). Toward a taxonomy of rehabilitation interventions: using an inductive approach to examine the "black box" of rehabilitation. *Archives of Physical*

- Medicine and Rehabilitation*, 85(4), 678-686. doi: 10.1016/j.apmr.2003.06.033
- Delgado-Rodríguez, M. (2006). Systematic reviews of meta-analyses: applications and limitations. *Journal of Epidemiology and Community Health*, 60(2), 90-92.
- Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, 58(8), 635-641.
- Derenzi, B., Findlater, L., Payne, J., Birnbaum, B., Mangilima, J., Parikh, T., . . . Lesh, N. (2012). Improving Community Health Worker Performance Through Automated SMS. *ACM International Conference Proceeding Series*, 25-34 doi: 10.1145/2160673.2160677
- Des Jarlais, D. C., Lyles, C., & Crepaz, N. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *American Journal of Public Health*, 94(3), 361-366.
- Dias, A., Dewey, M. E., D'Souza, J., Dhume, R., Motghare, D. D., Shaji, K. S., . . . Patel, V. (2008). The effectiveness of a home care program for supporting caregivers of persons with dementia in developing countries: a randomised controlled trial from Goa, India. *Plos One*, 3(6), e2333. doi: 10.1371/journal.pone.0002333
- Dickersin, K. (2005). Publication Bias: Recognizing the Problem, Understanding its Origins and Scope, and Preventing Harm. In H. Rothstein, A. J. Surtton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, (pp. 11-33). West Sussex: John Wiley & Sons.
- Diodato, V. P., & Gellatly, P. (2013). *Dictionary of bibliometrics*: Routledge.
- Dixon-Woods, M. (2011). Using framework-based synthesis for conducting reviews of qualitative studies. *BMC Medicine*, 9(1), 1.
- Dobrow, M. J., Goel, V., & Upshur, R. E. G. (2004). Evidence-based health policy: context and utilisation. *Social Science & Medicine*, 58(1), 207-217. doi: 10.1016/s0277-9536(03)00166-7
- Druetz, T., Siekmans, K., Goossens, S., Ridde, V., & Haddad, S. (2013). The community case management of pneumonia in Africa: a review of the evidence. *Health Policy and Planning*. doi: 10.1093/heapol/czt104
- Drummond, M., & Jefferson, T. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal*, 313(7052), 275-283.
- Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, UK: Oxford University Press.
- Duffy, B., & Richards, W. (1978). Evaluation of SOFTLY Unpublished manuscript. *Australian Institute of Criminology, Canberra*.
- Effective Practice and Organisation of Care Group. (2014). Synthesising results when it does not make sense to do a meta- analysis. *EPOC Resources for review authors*. Retrieved July 12, 2014, from <http://epocoslo.cochrane.org/epoc-specific-resources-review-authors>
- Effective Practice and Organisation of Care Group. (2015). EPOC Resources for Review Authors. from <http://epoc.cochrane.org/epoc-specific-resources-review-authors>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634.

- Elder, J., Louis, T., Sutisnaputra, O., Sulaeiman, N., Ware, L., Shaw, W., . . . Graeff, J. (1992). The use of diarrhoeal management counselling cards for community health volunteer training in Indonesia: The HealthCom Project. *Journal of Tropical Medicine and Hygiene*, 95(5), 301-308.
<http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/082/CN-00217082/frame.html>
- Eldredge, J. D. (2000). Evidence-based librarianship: an overview. *Bulletin of the Medical Library Association*, 88(4), 289.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107-115.
- Eloff, I., Finestone, M., Makin, J. D., Boeving-Allen, A., Visser, M., Ebersohn, L., . . . Forsyth, B. W. C. (2014). A randomized clinical trial of an intervention to promote resilience in young children of HIV-positive mothers in South Africa. (Special Issue: Children born into families affected by HIV.). *AIDS*, 28(Suppl. 3), S347-S357.
- Emukah, E. C., Enyinnaya, U., Olaniran, N. S., Akpan, E. A., Hopkins, D. R., Miri, E. S., . . . Katabarwa, M. N. (2008). Factors affecting the attrition of community-directed distributors of ivermectin, in an onchocerciasis-control programme in the Imo and Abia states of south-eastern Nigeria. *Annals of Tropical Medicine and Parasitology*, 102(1), 45-51. doi: 10.1179/136485908x252241
- Engelbrechtsen, I. M., Jackson, D., Fadnes, L. T., Nankabirwa, V., Diallo, A. H., Doherty, T., . . . group, P. E.-s. (2014). Growth effects of exclusive breastfeeding promotion by peer counsellors in sub-Saharan Africa: the cluster-randomised PROMISE EBF trial. *BMC Public Health*, 14, 633. doi: 10.1186/1471-2458-14-633
- Ertl, V., Pfeiffer, A., Schauer, E., Elbert, T., & Neuner, F. (2011). Community-implemented trauma therapy for former child soldiers in Northern Uganda: a randomized controlled trial. *Journal of the American Medical Association*, 306(5), 503-512. doi: 10.1001/jama.2011.1060
- Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care*, 21(2), 240-245.
- Evidence-Based Medicine Working Group. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 268(17), 2420.
- Faraone, S. V. (2008). Interpreting Estimates of Treatment Effects: Implications for Managed Care. *Pharmacy and Therapeutics*, 33(12), 700-711.
- Field, A., & Hole, G. (2002). *How to design and report experiments*: Sage.
- Findley, S. E., Uwemedimo, O. T., Doctor, H. V., Green, C., Adamu, F., & Afenyadu, G. Y. (2013). Comparison of high- versus low-intensity community health worker intervention to promote newborn and child health in Northern Nigeria. *Int J Womens Health*, 5, 717-728. doi: 10.2147/ijwh.s49785
- Fisher, R. A. (1925). *Statistical methods for research workers*: Genesis Publishing Pvt Ltd.
- Fleury, S. (2011). Brazil's health-care reform: social movements and civil society. *Lancet*, 377(9779), 1724-1725.
- Flodgren, G., Eccles Martin, P., Shepperd, S., Scott, A., Parmelli, E., & Beyer Fiona, R. (2011). An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient

- outcomes. *Cochrane Database of Systematic Reviews*, (7).
<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009255/abstract>
 doi:10.1002/14651858.CD009255
- Florez-Arango, J., Iyengar, M., Dunn, K., & Zhang, J. (2011). Performance factors of mobile rich media job aids for community health workers. *Journal of the American Medical Informatics Association : JAMIA*, 18(2), 131-137.
<http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/329/CN-00787329/frame.html> doi:10.1136/jamia.2010.010025
- Flottorp, S. A., Oxman, A. D., Krause, J., Musila, N. R., Wensing, M., Godycki-Cwirko, M., . . . Eccles, M. P. (2013). A checklist for identifying determinants of practice: A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implementation Science*, 8. doi: Artn 35
 10.1186/1748-5908-8-35
- Foreit, J. R., & Foreit, K. G. (1984). Quarterly versus monthly supervision of CBD family planning programs: an experimental study in Northeast Brazil. *Studies in Family Planning*, 112-120.
- Fottrell, E., Azad, K., Kuddus, A., Younes, L., Shaha, S., Nahar, T., . . . Houweling, T. A. (2013). The effect of increased coverage of participatory women's groups on neonatal mortality in Bangladesh: A cluster randomized trial. *JAMA Pediatr*, 167(9), 816-825. doi: 10.1001/jamapediatrics.2013.2534
- Frankel, S. (1992). *The community health worker: effective programmes for developing countries*: Oxford University Press.
- Fraser, M. W., Richman, J. M., Galinsky, M. J., & Day, S. H. (2009). *Intervention Research: Developing Social Programs: Developing Social Programs*: Oxford University Press, USA.
- Freidson, E. (1974). *Professional dominance: The social structure of medical care*: Transaction Publishers.
- Fretheim, A., Flottorp, S., & Oxman, A. D. (2005). It is a capital mistake to theorize before one has data: a response to Eccle's criticism of the OFF theory of research utilization. *Journal of Clinical Epidemiology*, 58(2), 119.
- Fulton, B. D., Scheffler, R. M., Sparkes, S. P., Auh, E. Y., Vujicic, M., & Soucat, A. (2011). Health workforce skill mix and task shifting in low income countries: a review of recent evidence. *Human Resources for Health*, 9, 1.
- Fylkesnes, K., Sandoy, I. F., Jurgensen, M., Chipimo, P. J., Mwangala, S., & Michelo, C. (2013). Strong effects of home-based voluntary HIV counselling and testing on acceptance and equity: a cluster randomised trial in Zambia. *Social Science & Medicine*, 86, 9-16. doi: 10.1016/j.socscimed.2013.02.036
- Gagnier, J. J., Moher, D., Boon, H., Beyene, J., & Bombardier, C. (2012). Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. *BMC Med Res Methodol*, 12(1), 1.
- Galbraith, J. S., Herbst, J. H., Whittier, D. K., Jones, P. L., Smith, B. D., Uhl, G., & Fisher, H. H. (2011). Taxonomy for strengthening the identification of core elements for evidence-based behavioral interventions for HIV/AIDS prevention. *Health Educ Res*, 26(5), 872-885. doi: 10.1093/her/cyr030
- Gambrill, E. D. (2003). Evidence-based practice: Sea change or emperor's new clothes? *Journal of Social Work Education*, 39(1), 3-23.
- Gardner, J. M., Walker, S. P., Powell, C. A., & Grantham-McGregor, S. (2003). A randomized controlled trial of a home-visiting intervention on cognition and

- behavior in term low birth weight infants. *J Pediatr*, 143(5), 634-639. doi: 10.1067/S0022-3476(03)00455-4
- Gautham, M., Iyengar, M. S., & Johnson, C. W. (2015). Mobile phone-based clinical guidance for rural health providers in India. *Health Informatics J*, 21(4), 253-266. doi: 10.1177/1460458214523153
- Gawande, A. (2009). *The Checklist Manifesto: How To Get Things Right*. Metropolitan Books.
- Gellis, Z., & Reid, W. J. (2004). Strengthening evidence-based practice. *Brief Treatment and Crisis Intervention*, 4(2), 155.
- George, C. M., van Geen, A., Slavkovich, V., Singha, A., Levy, D., Islam, T., . . . Liu, X. (2012). A cluster-based randomized controlled trial promoting community participation in arsenic mitigation efforts in Bangladesh. *Environmental Health*, 11(1), 1.
- Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., . . . Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187), 998-1001. doi: 10.1126/science.1251178
- Ghani, R. (2011). 40 million children have no access to the most basic healthcare services. *British Medical Journal*, 343, d4667. doi: 10.1136/bmj.d4667
- Gibbs, L., & Gambrill, E. (2002). Evidence-Based Practice: Counterarguments to Objections. *Research on Social Work Practice*, 12(3), 452-476. doi: 10.1177/1049731502012003007
- Gilmore, B., & McAuliffe, E. (2013). Effectiveness of community health workers delivering preventive interventions for maternal and child health in low-and middle-income countries: a systematic review. *BMC Public Health*, 13(1), 1-14.
- Gilroy, K., & Winch, P. (2006). Management of sick children by community health workers. Intervention models and programme examples. Geneva, Switzerland: WHO, Department of Child and Adolescent Health and Development.
- Gilson, L., Walt, G., Heggenhougen, K., Owuor-Omondi, L., Perera, M., Ross, D., & Salazar, L. (1989). National community health worker programs: how can they be strengthened? *Journal of Public Health Policy*, 518-532.
- Ginsburg, O. M., Chowdhury, M., Wu, W., Chowdhury, M. T., Pal, B. C., Hasan, R., . . . Salim, R. (2014). An mHealth model to increase clinic attendance for breast symptoms in rural Bangladesh: can bridging the digital divide help close the cancer divide? *Oncologist*, 19(2), 177-185.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, 5(10), 3-8.
- Glasziou, P., & Longbottom, H. (1999). Evidence-based public health practice. *Australian and New Zealand Journal of Public Health*, 23(4), 436-440.
- Glasziou, P., Meats, E., Heneghan, C., & Shepperd, S. (2008). What is missing from descriptions of treatment in trials and reviews? *British Medical Journal*, 336(7659), 1472-1474.
- Glenton, C., Colvin, C. J., Carlsen, B., Swartz, A., Lewin, S., Noyes, J., & Rashidian, A. (2013a). Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: qualitative evidence synthesis. *Cochrane Database of Systematic Reviews*, 10, CD010414.
- Glenton, C., Khanna, R., Morgan, C., & Nilsen, E. S. (2013b). The effects, safety and acceptability of compact, pre-filled, autodisable injection devices when

- delivered by lay health workers. *Tropical Medicine & International Health*, 18(8), 1002-1016\.
- Glenton, C., Lewin, S., & Scheel, I. B. (2011a). Still too little qualitative research to shed light on results from reviews of effectiveness trials: A case study of a Cochrane review on the use of lay health workers. *Implementation Science*, 6. doi: 10.1186/1748-5908-6-53
- Glenton, C., Scheel, I. B., Lewin, S., & Swingler, G. H. (2011b). Can lay health workers increase the uptake of childhood immunisation? Systematic review and typology. *Tropical Medicine & International Health*, 16(9), 1044-1053. doi: <http://dx.doi.org/10.1111/j.1365-3156.2011.02813.x>
- Global Health Workforce Alliance. (2012). *Synthesis Paper: Developed out of the outcomes of four consultations on Community Health Workers and other Frontline Health Workers held in May/June 2012*. http://pdf.usaid.gov/pdf_docs/pnaed275.pdf
- Glouberman, S., & Zimmerman, B. (2002). Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like? *Commission On The Future Of Health Care in Canada*.
- Gogia, S., Ramji, S., Gupta, P., Gera, T., Shah, D., Mathew, J. L., . . . Panda, R. (2011). Community based newborn care: a systematic review and metaanalysis of evidence: UNICEF-PHFI series on newborn and child health, India. *Indian Pediatrics*, 48(7), 537-546.
- Gogia, S., & Sachdev, H. S. (2010). Home visits by community health workers to prevent neonatal deaths in developing countries: A systematic review. [French]. [Des visites a domicile par des professionnels communautaires de la sante permettent de reduire la mortalite infantile dans les pays en voie de developpement: Une revue systematique.]. *Bulletin of the World Health Organization*, 88(9), 658-666.
- Gökçay, G., Bulut, A., & Neyzi, O. (1993). Paraprofessional women as health care facilitators in mother and child health. *Trop Doct*, 23(2), 79-81.
- Goldenberg, M. J. (2005). Evidence-based ethics? On evidence-based practice and the "empirical turn" from normative bioethics. *BMC Med Ethics*, 6(1), E11. doi: 10.1186/1472-6939-6-11
- Gottfredson, G. D. (1987). Peer Group Interventions To Reduce The Risk Of Delinquent Behavior: A Selective Review And A New Evaluation. *Criminology*, 25(3), 671-714.
- Gough, D., Oliver, S., & Thomas, J. (2012). *An introduction to systematic reviews*: Sage.
- Grant, S. P., Mayo-Wilson, E., Melendez-Torres, G., & Montgomery, P. (2013). Reporting quality of social and psychological intervention trials: a systematic review of reporting guidelines and trial publications. *Plos One*, 8(5), e65442.
- Grantham-McGregor, S., Walker, S., Chang, S., & Powell, C. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children1. *American Journal of Clinical Nutrition*, 66, 247-253.
- Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., & Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study. *Lancet*, 338(8758), 1-5.
- Grantham-McGregor, S. M., Walker, S. P., Himes, J. H., & Powell, C. A. (1993). The effect of nutritional supplementation and stunting on morbidity in young

- children: The Jamaican study. *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 87, 109-113.
- Grayson, L., & Gomersall, A. (2003). A difficult business: finding the evidence for social science reviews (Working Paper 19). London: ESRC UK Centre for Evidence Based Policy and Practice.
- Green, C. P., Joyce, S., & Foreit, J. R. (2002). Using men as community-based distributors of condoms *Frontiers in Reproductive Health*. Washington, DC: Population Council.
- Gul, A., & Ali, B. (2004). The onset and duration of benefit from counselling by minimally trained counsellors on anxiety and depression in women. *JPMMA. The Journal of the Pakistan Medical Association*, 54(11), 549-552.
- Guyatt, G., & Rennie, D. (2007). *Users' guides to the medical literature: a manual for evidence-based clinical practice* (Vol. 706). Chicago: AMA press.
- Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., . . . Schünemann, H. J. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology*, 64(12), 1311-1316. doi: <http://dx.doi.org/10.1016/j.jclinepi.2011.06.004>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008a). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical research ed.)*.
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008b). Rating quality of evidence and strength of recommendations: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal*, 336(7650), 924.
- Guyatt, G. H., Thorlund, K., Oxman, A. D., Walter, S. D., Patrick, D., Furukawa, T. A., . . . Vist, G. (2013). GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles—continuous outcomes. *Journal of Clinical Epidemiology*, 66(2), 173-183.
- Habicht, J.-P., Victora, C., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology*, 28(1), 10-18.
- Hackett, K. M., & Sellen, D. (2015). Job satisfaction, self-efficacy, and performance of community health workers participating in a mobile health (mHealth) program to improve maternal, newborn and child health (MNCH) in Rural Tanzania. *Annals of Global Health*, 81 (1), 75.
- Hadley, M., & Maher, D. (2000). Community involvement in tuberculosis control: lessons from other health care programmes. *The International Journal of Tuberculosis and Lung Disease*, 4(5), 401-408.
- Haider, R., Ashworth, A., Kabir, I., & Huttly, S. R. A. (2000). Effect of community-based peer counsellors on exclusive breastfeeding practices in Dhaka, Bangladesh: a randomised controlled trial. *Lancet*, 356(9242), 1643-1647. doi: 10.1016/s0140-6736(00)03159-7
- Haines, A., Sanders, D., Lehmann, U., Rowe, A. K., Lawn, J. E., Jan, S., . . . Bhutta, Z. (2007). Achieving child survival goals: potential contribution of community health workers. *Lancet*, 369(9579), 2121-2131. doi: 10.1016/s0140-6736(07)60325-0
- Hall, J. (2011). Effective community-based interventions to improve exclusive breast feeding at four to six months in low- and low-middle-income countries: a

- systematic review of randomised controlled trials. *Midwifery*, 27(4), 497-502. doi: 10.1016/j.midw.2010.03.011
- Hall, J. J., & Taylor, R. (2003). Health for all beyond 2000: the demise, of the Alma-Ata declaration and primary health care in developing countries. *Medical Journal of Australia*, 178(1), 17-20.
- Halm, E. A., Lee, C., & Chassin, M. R. (2002). Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Annals of Internal Medicine*, 137(6), 511-520.
- Hammersley, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(1), 85-100. doi: 10.1332/1744264052703203
- Hanson, C., Manzi, F., Mkumbo, E., Shirima, K., Penfold, S., Hill, Z., . . . Schellenberg, J. (2015). Effectiveness of a Home-Based Counselling Strategy on Neonatal Care and Survival: A Cluster-Randomised Trial in Six Districts of Rural Southern Tanzania. *Plos Medicine*, 12(9), e1001881. doi: 10.1371/journal.pmed.1001881
- Hartling, L., Chisholm, A., Thomson, D., & Dryden, D. M. (2012a). A Descriptive Analysis of Overviews of Reviews Published between 2000 and 2011. *Plos One*, 7(11), e49667. doi: 10.1371/journal.pone.0049667
- Hartling, L., Chisholm, A., Thomson, D., & Dryden, D. M. (2012b). A Descriptive Analysis of Overviews of Reviews Published between 2000 and 2011. *Plos One*, 7(11).
- Hartling, L., Vandermeer, B., & Fernandes, R. M. (2014). Systematic reviews, overviews of reviews and comparative effectiveness reviews: a discussion of approaches to knowledge synthesis. *Evidence-Based Child Health a Cochrane Review Journal*, 9(2), 486-494.
- Harvey, S. A., Jennings, L., Chinyama, M., Masaninga, F., Mulholland, K., & Bell, D. R. (2008). Improving community health worker use of malaria rapid diagnostic tests in Zambia: package instructions, job aid and job aid-plus-training. *Malaria Journal*, 7. doi: 10.1186/1475-2875-7-160
- Harzing, A.-W. (2007). Publish or Perish. <http://www.harzing.com/pop.htm>.
- Harzing, A.-W. (2013). From publication to impact: Using Google Scholar and Publish or Perish to measure research impact. Retrieved March 3, 2016, from <http://lti.lse.ac.uk/events/networkED-seminar-series-11.php>
- Hasegawa, A., Yasuoka, J., Ly, P., Nguon, C., & Jimba, M. (2013). Integrating child health services into malaria control services of village malaria workers in remote Cambodia: service utilization and knowledge of malaria management of caregivers. *Malaria Journal*, 12. doi: 10.1186/1475-2875-12-292
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science*, 5(1), 1.
- Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how "out of control" can a randomised controlled trial be? *British Medical Journal*, 328(7455), 1561-1563. doi: 10.1136/bmj.328.7455.1561
- Hawe, P., Shiell, A., & Riley, T. (2009). Theorising Interventions as Events in Systems. *American Journal of Community Psychology*, 43(3-4), 267-276. doi: 10.1007/s10464-009-9229-9
- Haynes, R. B., Devereaux, P. J., & Guyatt, G. H. (2002). Clinical expertise in the era of evidence-based medicine and patient choice. *Evidence Based Medicine*, 7(2), 36-38.

- Haynes, R. B., Sackett, D. L., Gray, J. M., Cook, D. J., & Guyatt, G. H. (1996). Transferring evidence from research into practice: 1. The role of clinical care research evidence in clinical decisions. *ACP J Club*, 125(3), A14-16.
- Hedden, H. (2010). *The Accidental Taxonomist*. Medford, N.J.: Information Inc.
- Heggenhougen, K., Vaughan, P., Muhondwa, E. P., & Rutabanzibwa Ngaiza, J. (1987). *Community health workers: the Tanzanian experience*. Oxford University Press.
- Heymann, J., Stein, M. A., & Moreno, G. (2014). *Disability and equity at work*. Oxford University Press.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. doi: 10.1038/520429a
- Higgins, E. T. (1999). When do self-discrepancies have specific relations to emotions? The second-generation question of Tangney, Niedenthal, Covert, and Barlow (1998).
- Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration.
- Hill, A. B. (1951). The clinical trial. *British Medical Bulletin*, 7, 278-282.
- Hill, S., & Ryan, R. (2014). Cochrane review-interventions to improve safe and effective medicines use by consumers: an overview of systematic review. *Journal of Evidence-Based Medicine*, 7(3), 231-231.
- Hinshaw, S. P. (2002). Intervention research, theoretical mechanisms, and causal processes related to externalizing behavior patterns. *Development and Psychopathology*, 14(04), 789-818.
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Johnston, M. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *British Medical Journal*, 348.
- Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. Basic Books.
- Hopkins, H., Talisuna, A., Whitty, C. J., & Staedke, S. G. (2007). Impact of home-based management of malaria on health outcomes in Africa: a systematic review of the evidence. *DARE*, 6, 134. doi: 10.1186/1475-2875-6-134
- Houweling, T. A., Tripathy, P., Nair, N., Rath, S., Rath, S., Gope, R., . . . Prost, A. (2013). The equity impact of participatory women's groups to reduce neonatal mortality in India: secondary analysis of a cluster-randomised trial. *Int J Epidemiol*, 42(2), 520-532. doi: 10.1093/ije/dyt012
- Hsiao, W. C. (1984). Transformation of health care in China. *New England Journal of Medicine*, 310(14), 932-936.
- Hundley, V. A., Avan, B. I., Sullivan, C. J., & Graham, W. J. (2013). Should oral misoprostol be used to prevent postpartum haemorrhage in home-birth settings in low-resource countries? A systematic review of the evidence. *BJOG: An International Journal of Obstetrics & Gynaecology*, 120(3), 277-285; discussion 286-277.
- Ijumba, P., Doherty, T., Jackson, D., Tomlinson, M., Sanders, D., Swanevelder, S., & Persson, L. A. (2015). Effect of an integrated community-based package for maternal and newborn care on feeding patterns during the first 12 weeks of life: a cluster-randomized trial in a South African township. *Public Health Nutr*, 18(14), 2660-2668. doi: 10.1017/S1368980015000099
- Illich, I. (1976). *Limits to medicine*. London, UK: Marion Boyars.

- Ingersoll, G. L. (2000). Evidence-based nursing: What it is and what it isn't. *Nursing Outlook*, 48(4), 151-152.
- International Astronomical Union. (2008). Resolution B5: Definition of planet in the Solar System. Retrieved February 13, 2016, from http://www.iau.org/static/resolutions/Resolution_GA26-5-6.pdf
- Ioannidis, J. P. (2009). Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ*, 181(8), 488-493. doi: 10.1503/cmaj.081086
- Iowa Intervention Project. (1992). *Nursing Interventions Classification (NIC) Taxonomy of Nursing Intervention*. The University of Iowa.
- Iyengar, M. S., & Fjorez-Arango, J. F. (2013). Decreasing workload among community health workers using interactive, structured, rich-media guidelines on smartphones. *Technology and Health Care*, 21(2), 113-123. doi: 10.3233/thc-130713
- Jadad, A. R., Cook, D. J., & Browman, G. P. (1997). A guide to interpreting discordant systematic reviews. *CMAJ*, 156(10), 1411-1416.
- Jafar, T. H., Hatcher, J., Poulter, N., Islam, M., Hashmi, S., Qadri, Z., . . . Hypertension Research, G. (2009). Community-based interventions to promote blood pressure control in a developing country: a cluster randomized trial. *Annals of Internal Medicine*, 151(9), 593-601.
- Jafar, T. H., Islam, M., Bux, R., Poulter, N., Hatcher, J., Chaturvedi, N., . . . Hypertension Research, G. (2011). Cost-effectiveness of community-based strategies for blood pressure control in a low-income developing country: findings from a cluster-randomized, factorial-controlled trial. *Circulation*, 124(15), 1615-1625.
- Jafar, T. H., Islam, M., Hatcher, J., Hashmi, S., Bux, R., Khan, A., . . . Hypertension Research, G. (2010). Community based lifestyle intervention for blood pressure reduction in children and young adults in developing country: cluster randomised controlled trial. *British Medical Journal*, 340, c2641.
- Jafar, T. H., Jehan, I., Liang, F., Barbier, S., Islam, M., Bux, R., . . . Ebrahim, S. (2015). Control of Blood Pressure and Risk Attenuation: Post Trial Follow-Up of Randomized Groups. *Plos One*, 10(11), e0140550. doi: 10.1371/journal.pone.0140550
- Jaffar, S., Amuron, B., Foster, S., Birungi, J., Levin, J., Namara, G., . . . Jinja Trial, T. (2009). Rates of virological failure in patients treated in a home-based versus a facility-based HIV-care model in Jinja, southeast Uganda: a cluster-randomised equivalence trial. *Lancet*, 374(9707), 2080-2089. doi: 10.1016/S0140-6736(09)61674-3
- Jaskiewicz, W., & Tulenko, K. (2012). Increasing community health worker productivity and effectiveness: a review of the influence of the work environment. *Human Resources for Health*, 10(1), 38. doi: 10.1186/1478-4491-10-38
- Jenicek, M. (1997). Epidemiology, evidenced-based medicine, and evidence-based public health. *Journal of Epidemiology*, 7(4), 187-197.
- Jintaganont, P., Stoeckel, J., & Butaras, S. (1988). The impact of an oral rehydration therapy program in Southern Thailand. *American Journal of Public Health*, 78(10), 1302-1304.
- Jones, H. S., Liddell, H. G., & Scott, R. (1996). *A Greek-English lexicon: with a revised supplement 1996*. Clarendon Press.

- Jones, L., Othman, M., Dowswell, T., Alfievic, Z., Gates, S., Newburn, M., . . . Neilson James, P. (2012). Pain management for women in labour: an overview of systematic reviews. *Cochrane Database of Systematic Reviews*, (3). <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009234.pub2/abstract> doi:10.1002/14651858.CD009234.pub2
- Joos, O., Silva, R., Amouzou, A., Moulton, L. H., Perin, J., Bryce, J., & Mullany, L. C. (2016). Evaluation of a mHealth Data Quality Intervention to Improve Documentation of Pregnancy Outcomes by Health Surveillance Assistants in Malawi: A Cluster Randomized Trial. *Plos One*, 11(1), e0145238. doi: 10.1371/journal.pone.0145238
- Jordaan, E. R., Atkins, S., Van Niekerk, A., & Seedat, M. (2005). The development of an instrument measuring unintentional injuries in young children in low-income settings to serve as an evaluation tool for a childhood home injury prevention program. *Journal of Safety Research*, 36(3), 269-280.
- Jørgensen, A. W., Hilden, J., & Gøtzsche, P. C. (2006). Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *British Medical Journal*, 333(7572), 782.
- Joshi, R., Alim, M., Kengne, A. P., Jan, S., Maulik, P. K., Peiris, D., & Patel, A. A. (2014). Task shifting for non-communicable disease management in low and middle income countries--a systematic review. *Plos One*, 9(8), e103754.
- Jung, M. K., Kwang, W. K., Chul, H. O., Woo, H. J., Sung, H. K., & Dae, H. P. (2009). Assessment of village health worker training program in Tuguegarao, Philippine. *Journal of Preventive Medicine and Public Health*, 42(6), 377-385.
- Kadiyala, S. (2004). Scaling up HIV/AIDS interventions through expanded partnerships (STEPS) in Malawi. *Report produced by the International Food Policy Research Institute for the Social Development Family of the World Bank*.
- Kahssay, H. M., Taylor, M. E., & Berman, P. A. (1998). *Community Health Workers: The Way Forward*. Geneva, Switzerland: WHO.
- Kallander, K., Tibenderana, J., Kirkwood, B., Hill, Z., Strachan, D., Soremekun, S., . . . Meek, S. (2012). Inscale cluster randomized trial evaluating the effect of innovative motivation and supervision approaches on community health worker performance and retention in uganda and mozambique: Intervention design. *Trials*, 16(1).
- Kallander, K., Tibenderana, J. K., Akpogheneta, O. J., Strachan, D. L., Hill, Z., ten Asbroek, A. H. A., . . . Meek, S. R. (2013). Mobile Health (mHealth) Approaches and Lessons for Increased Performance and Retention of Community Health Workers in Low- and Middle-Income Countries: A Review. *Journal of Medical Internet Research*, 15(1). doi: 10.2196/jmir.2130
- Kalyango, J. N., Rutebemberwa, E., Alfven, T., Ssali, S., Peterson, S., & Karamagi, C. (2012). Performance of community health workers under integrated community case management of childhood illnesses in eastern Uganda. *Malaria Journal*, 11, 282. doi: 10.1186/1475-2875-11-282
- Kamal-Yanni, M. M., Potet, J., & Saunders, P. M. (2012). Scaling-up malaria treatment: a review of the performance of different providers. *Malaria Journal*, 11, 414.
- Kan, X. (1992). The current situation of village health worker programmes in China. *The community health worker: effective programmes for developing countries*, 88.

- Katz, J., West, K. R., Wu, L., Khattry, S. K., Pradhan, E. K., Christian, P., . . . Shrestha, S. R. (2002). Determinants of maternal vitamin A or beta-carotene supplementation coverage: Village-based female distributors in Nepal. *American Journal of Public Health*, 92(7), 1105-1107. doi: 10.2105/ajph.92.7.1105
- Kelly, M., Morgan, A., Ellis, S., Younger, T., Huntley, J., & Swann, C. (2010). Evidence based public health: a review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Social Science & Medicine*, 71(6), 1056-1062.
- Kemm, J. (2006). The limitations of 'evidence-based' public health. *Journal of Evaluation in Clinical Practice*, 12(3), 319-324.
- Kerber, K. J., de Graft-Johnson, J. E., Bhutta, Z. A., Okong, P., Starrs, A., & Lawn, J. E. (2007). Continuum of care for maternal, newborn, and child health: from slogan to service delivery. *Lancet*, 370(9595), 1358-1369.
- Khan, K. S., Ter Riet, G., Glanville, J., Sowden, A. J., & Kleijnen, J. (2001). *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*: NHS Centre for Reviews and Dissemination.
- Kidane, G., & Morrow, R. H. (2000). Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial. *Lancet*, 356(9229), 550-555.
- King, M. (1966). *Medical Care in Developing Countries. A Primer on the Medicine of Poverty and a Symposium from Makerere*. Paper presented at the Medical Care in Developing Countries. A Primer on the Medicine of Poverty and a Symposium from Makerere.
- Kirkwood, B. R., Manu, A., ten Asbroek, A. H. A., Soremekun, S., Weobong, B., Gyan, T., . . . Hill, Z. (2013). Effect of the Newhints home-visits intervention on neonatal mortality rate and care practices in Ghana: a cluster randomised controlled trial. *Lancet*, 381(9884), 2184-2192. doi: 10.1016/s0140-6736(13)60095-1
- Klein, R. (2000). From evidence-based medicine to evidence-based policy? *Journal of health services research & policy*, 5(2), 65.
- Kok, G., Gottlieb, N. H., Peters, G. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A., . . . Bartholomew, L. K. (2015a). A taxonomy of behaviour change methods: an Intervention Mapping approach. *Health Psychol Rev*, 1-16. doi: 10.1080/17437199.2015.1077155
- Kok, M. C., Dieleman, M., Taegtmeier, M., Broerse, J. E., Kane, S. S., Ormel, H., . . . de Koning, K. A. (2014). Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review. *Health Policy and Planning*, [21] p. doi: <http://dx.doi.org/10.1093/heapol/czu126>
- Kok, M. C., Dieleman, M., Taegtmeier, M., Broerse, J. E., Kane, S. S., Ormel, H., . . . de Koning, K. A. (2015b). Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review. *Health Policy and Planning*, 30(9), 1207-1227. doi: 10.1093/heapol/czu126
- Krathwohl, D. R. (1994). Reflections on the taxonomy: Its past, present, and future. *Yearbook (National Society for the Study of Education)*, 93(2), 181-202.
- Kredo, T., Adeniyi, F. B., Bateganya, M., & Pienaar, E. D. (2014). Task shifting from doctors to non-doctors for initiation and maintenance of antiretroviral therapy. *Cochrane Database of Systematic Reviews*, 7, CD007331.

- Kreitzer, A. E., & Madaus, G. F. (1994). Empirical investigations of the hierarchical structure of the taxonomy. *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education*, 64-81.
- Kruk, M. E., Myers, M., Varpilah, S. T., & Dahn, B. T. (2015). What is a resilient health system? Lessons from Ebola. *Lancet*, 385(9980), 1910-1912.
- Kumar, V., Kumar, A., Das, V., Srivastava, N. M., Baqui, A. H., Santosham, M., . . . Saksham Study, G. (2012). Community-driven impact of a newborn-focused behavioral intervention on maternal health in Shivgarh, India. *Int J Gynaecol Obstet*, 117(1), 48-55. doi: 10.1016/j.ijgo.2011.10.031
- Kumar, V., Mohanty, S., Kumar, A., Misra, R. P., & Santosham, M. (2008). Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. *Lancet*, 372(9644), 1151-1162. doi: [http://dx.doi.org/10.1016/S0140-6736\(08\)61483-X](http://dx.doi.org/10.1016/S0140-6736(08)61483-X)
- Kwasnik, B. H. (2000). The role of classification in knowledge representation and discovery. *Library Trends*, 48(1), 22-47.
- Kweku, M., Webster, J., Adjuik, M., Abudey, S., Greenwood, B., & Chandramohan, D. (2009). Options for the delivery of intermittent preventive treatment for malaria to children: a community randomised trial. *Plos One*, 4(9), e7256.
- Lamb, S. E., Becker, C., Gillespie, L. D., Smith, J. L., Finnegan, S., Potter, R., . . . Taxonomy, I. (2011). Reporting of complex interventions in clinical trials: development of a taxonomy to classify and describe fall-prevention interventions. *Trials*, 12, 125. doi: 10.1186/1745-6215-12-125
- Lambe, P. (2014). *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*: Elsevier.
- Lamming, L., Mason, D., Morton, K., Pears, S., Bijker, M., Sutton, S., & Hardeman, W. (2014). What do we Know About Brief Interventions to Promote Physical Activity? Systematic Review of Reviews. *European Health Psychologist*, 16(S), 808.
- Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *J Eval Clin Pract*, 10(2), 307-312. doi: 10.1111/j.2002.384.doc.x
- Langendam, M. (2012). *Cochrane MIF project: Enhancing the acceptance and implementation of Summary of Findings Tables in Cochrane Reviews*. Paper presented at the GRADE Working Group meeting, Auckland.
- Lassi, Z. S., & Bhutta, Z. A. (2015). Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes.[Update of Cochrane Database Syst Rev. 2010;(11):CD007754; PMID: 21069697]. *Cochrane Database of Systematic Reviews*, 3, CD007754.
- Lassi, Z. S., Haider, B. A., & Bhutta, Z. A. (2010). Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes. *Cochrane Database of Systematic Reviews*, 11.
- Lassi, Z. S., Haider, B. A., & Bhutta, Z. A. (2012). Community-based intervention packages for reducing maternal morbidity and mortality and improving neonatal outcomes. *Journal of Development Effectiveness*, 4(1), 151-187.
- Lavis, J. N., Oxman, A. D., Fretheim, A., & Lewin, S. (2009a). SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in

- balancing the pros and cons of policies. *Health Research Policy and Systems*, 7(Suppl 1), S16.
- Lavis, J. N., Oxman, A. D., Souza, N. M., Lewin, S., Gruen, R. L., & Fretheim, A. (2009b). SUPPORT Tools for evidence-informed health Policymaking (STP) 9: Assessing the applicability of the findings of a systematic review. *Health Research Policy and Systems*, 7(Suppl 1), S9.
- le Roux, I. M., le Roux, K., Comulada, W. S., Greco, E. M., Desmond, K. A., Mbewu, N., & Rotheram-Borus, M. J. (2010). Home visits by neighborhood Mentor Mothers provide timely recovery from childhood malnutrition in South Africa: results from a randomized controlled trial. *Nutrition Journal*, 9, 56.
- le Roux, I. M., le Roux, K., Mbeutu, K., Comulada, W. S., Desmond, K. A., & Rotheram-Borus, M. J. (2011). A randomized controlled trial of home visits by neighborhood mentor mothers to improve children's nutrition in South Africa. *Vulnerable Child Youth Stud*, 6(2), 91-102. doi: 10.1080/17450128.2011.564224
- le Roux, I. M., Rotheram-Borus, M. J., Stein, J., & Tomlinson, M. (2014). The impact of paraprofessional home visitors on infants' growth and health at 18 months. *Vulnerable Child Youth Stud*, 9(4), 291-304. doi: 10.1080/17450128.2014.940413
- le Roux, I. M., Tomlinson, M., Harwood, J. M., O'Connor, M. J., Worthman, C. M., Mbewu, N., . . . Rotheram-Borus, M. J. (2013). Outcomes of home visits for pregnant mothers and their infants: a cluster randomized controlled trial. *AIDS*, 27(9), 1461-1471.
- Lee, A. C., Lawn, J. E., Cousens, S., Kumar, V., Osrin, D., Bhutta, Z. A., . . . Darmstadt, G. L. (2009). Linking families and facilities for care at birth: what works to avert intrapartum-related deaths? *Int J Gynaecol Obstet*, 107 Suppl 1, S65-85, S86-68. doi: 10.1016/j.ijgo.2009.07.012
- Lee, A. C. C., Chandran, A., Herbert, H. K., Kozuki, N., Markell, P., Shah, R., . . . Baqui, A. H. (2014). Treatment of infections in young infants in low- and middle-income countries: a systematic review and meta-analysis of frontline health worker diagnosis and antibiotic access. *Plos Medicine*, 11(10).
- Lefeuvre, D., Dieng, M., Lamara, F., Raguin, G., & Michon, C. (2015). [Community health workers in HIV/AIDS care]. *Sante Publique (Vandoeuvre-Les-Nancey)*, 26(6), 879-888.
- LeFevre, A. E., Shillcutt, S. D., Waters, H. R., Haider, S., El Arifeen, S., Mannan, I., . . . Projahnmo Study, G. (2013). Economic evaluation of neonatal care packages in a cluster-randomized controlled trial in Sylhet, Bangladesh. *Bulletin of the World Health Organization*, 91(10), 736-745. doi: 10.2471/BLT.12.117127
- Lehmann, U., Friedman, I., & Sanders, D. (2004). Review of the utilisation and effectiveness of community-based health workers in Africa *Joint Learning Initiative on Human Resources for Health and Development (JLI), JLI Working Paper*: Global Health Trust.
- Lehmann, U., & Sanders, D. (2007). Community health workers: What do we know about them? The state of the evidence on programmes, activities, costs and impact on health outcomes of using community health workers *Evidence and Information for Policy, Department of Human Resources for Health* (Vol. 2, pp. 1-42). Geneva, Switzerland: World Health Organization.

- Leite, A. J., Puccini, R. F., Atalah, A. N., Alves Da Cunha, A. L., & Machado, M. T. (2005). Effectiveness of home-based peer counselling to promote breastfeeding in the northeast of Brazil: a randomized clinical trial. *Acta paediatrica*, 94(6), 741-746.
- Lewin, S., Munabi-Babigumira, S., Glenton, C., Daniels, K., Bosch-Capblanch, X., van Wyk, B. E., . . . Scheel, I. B. (2010). Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. *Cochrane Database of Systematic Reviews*(3). doi: 10.1002/14651858.CD004015.pub3
- Lewin, S., Zwarenstein, M., & Dick, J. (2005). Commentary: Lay health workers in primary and community health care - Response to Walt and Morrow. *International Journal of Epidemiology*, 34(6), 1252-1253.
- Lewis, S., & Clarke, M. (2001). Forest plots: trying to see the wood and the trees. *British Medical Journal*, 322(7300), 1479.
- Lewycka, S., Mwansambo, C., Rosato, M., Kazembe, P., Phiri, T., Mganga, A., . . . Costello, A. (2013). Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. *Lancet*, 381(9879), 1721-1735.
- Li, L., Tian, J., Tian, H., Sun, R., Liu, Y., & Yang, K. (2012). Quality and transparency of overviews of systematic reviews. *Journal of Evidence-Based Medicine*, 5(3), 166-173.
- Liberian Ministry of Health and Social Welfare. (2008). National Policy And Strategy On Community Health Services.
- Linnaeus, C. (1758). *Systema naturae*, vol. 1. *Systema naturae, Vol. 1*.
- Linnan, L., & Steckler, A. B. (2002). *Process Evaluation for Public Health Interventions and Research: An Overview*. San Francisco: Jossey-Bass.
- Littell, J. H. (2016). [Research Synthesis: Current Problems and Promise].
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*: Oxford University Press.
- Liu, A., Sullivan, S., Khan, M., Sachs, S., & Singh, P. (2011). Community health workers in global health: scale and scalability. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine*, 78(3), 419-435.
- Lopes, S. C., Cabral, A. J., & Sousa, B. d. (2014). Community health workers: to train or to restrain? A longitudinal survey to assess the impact of training community health workers in the Bolama Region, Guinea-Bissau. *Human Resources for Health*, 12, 8. doi: 10.1186/1478-4491-12-8
- López Quiñones, I. d. P. (1999). Community health workers: profile and training process in Colombia *HCT/AIEPI*, 20.1. Washington D.C.: OPS.
- Lorenz, E. N. (1995). *The Essence of Chaos*: University of Washington Press.
- Love, M. B., Gardner, K., & Legion, V. (1997). Community health workers: Who they are and what they do. *Health Education & Behavior*, 24(4), 510-522. doi: 10.1177/109019819702400409
- Lowe, D., Ryan, R., Santesso, N., & Hill, S. (2011). Development of a taxonomy of interventions to organise the evidence on consumers' medicines use. *Patient Educ Couns*, 85(2), e101-107. doi: 10.1016/j.pec.2010.09.024
- Lubega, M., Tumwesigye, N. M., Kadobera, D., Marrone, G., Wabwire-Mangen, F., Peterson, S., . . . Ekstrom, A. M. (2015). Effect of Community Support Agents on Retention of People Living With HIV in Pre-antiretroviral Care: A

- Randomized Controlled Trial in Eastern Uganda. *J Acquir Immune Defic Syndr*, 70(2), e36-43. doi: 10.1097/QAI.0000000000000723
- Lugada, E., Levin, J., Abang, B., Mermin, J., Mugalanzi, E., Namara, G., . . . Bunnell, R. (2010). Comparison of Home and Clinic-Based HIV Testing Among Household Members of Persons Taking Antiretroviral Therapy in Uganda: Results From a Randomized Trial. *Journal of Acquired Immune Deficiency Syndromes*, 55(2).
- Lwilla, F., Schellenberg, D., Masanja, H., Acosta, C., Galindo, C., Aponte, J., . . . Alonso, P. (2003). Evaluation of efficacy of community-based vs. institutional-based direct observed short-course treatment for the control of tuberculosis in Kilombero district, Tanzania. *Tropical Medicine & International Health*, 8(3), 204-210.
- MacCorquodale, D. W. (1982). Primary health care in the Dominican Republic: a study of health worker effectiveness. *J Trop Med Hyg*, 85(6), 251-254.
- Macinko, J., Guanais, F. C., & de Souza, M. d. F. M. (2006). Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002. *Journal of Epidemiology and Community Health*, 60(1), 13-19.
- Macintyre, S., & Petticrew, M. (2000). Good intentions and received wisdom are not enough. *Journal of Epidemiology and Community Health*, 54(11), 802-803.
- Maher, D., Van Gorkom, J., Gondrie, P., & Raviglione, M. (1999). Community contribution to tuberculosis care in countries with high tuberculosis prevalence: past, present and future. *The International Journal of Tuberculosis and Lung Disease*, 3(9), 762-768.
- Makan, B. (1997). *An Economic Analysis of Community Health Worker Programmes in the Western Cape Province*. Durban, South Africa: Health Systems Trust Retrieved from <http://www.hst.org.za/sites/default/files/ecoanal.pdf>.
- Malarcher, S., Meirik, O., Lebetkin, E., Shah, I., Spieler, J., & Stanback, J. (2011). Provision of DMPA by community health workers: what the evidence shows. *Contraception*, 83(6), 495-503.
- Manandhar, D. S., Osrin, D., Shrestha, B. P., Mesko, N., Morrison, J., Tumbahangphe, K. M., . . . de L Costello, A. M. (2004). Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial. *Lancet*, 364(9438), 970-979. doi: 10.1016/s0140-6736(04)17021-9
- Mangham-Jefferies, L., Pitt, C., Cousens, S., Mills, A., & Schellenberg, J. (2014). Cost-effectiveness of strategies to improve the utilization and provision of maternal and newborn health care in low-income and lower-middle-income countries: a systematic review. *BMC Pregnancy & Childbirth*, 14, 243.
- Mannan, H., Boostrom, C., Maclachlan, M., McAuliffe, E., Khasnabis, C., & Gupta, N. (2012). A systematic review of the effectiveness of alternative cadres in community based rehabilitation. *Human Resources for Health*, 10, 20. doi: 10.1186/1478-4491-10-20
- Manu, A. A. (2012). *Newhints Home Visits randomised controlled trial : impact on access to care for sick newborns and determinants, facilitators and barriers to this*. (Doctoral thesis), London School of Hygiene & Tropical Medicine. Retrieved from <http://researchonline.lshtm.ac.uk/768506/>
- Martínez-González, N. A., Berchtold, P., Ullman, K., Busato, A., & Egger, M. (2014). Integrated care programmes for adults with chronic conditions: a meta-review. *International Journal for Quality in Health Care*, mzu071.

- Mathes, A. (2004). Folksonomies-cooperative classification and communication through shared metadata. Retrieved February 14, 2016, from <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *The Annals of the American Academy of Political and Social Science*, 587(1), 16-30.
- McKeown, T. (1976). *The Role of Medicine: Dream, Mirage, or Nemesis?* London: Nuffield Provincial Hospitals Trust.
- McKercher, B. (2016). Towards a taxonomy of tourism products. *Tourism Management*, 54, 196-208. doi: 10.1016/j.tourman.2015.11.008
- Mdege, N. D., Chindove, S., & Ali, S. (2013). The effectiveness and cost implications of task-shifting in the delivery of antiretroviral therapy to HIV-infected patients: a systematic review. *Health Policy and Planning*, 28(3), 223-236.
- Medhanyie, A., Spigt, M., Dinant, G., & Blanco, R. (2012). Knowledge and performance of the Ethiopian health extension workers on antenatal and delivery care: a cross-sectional study. *Human Resources for Health*, 10, 44.
- Medical Research Council. (2000). A framework for development and evaluation of RCTs for complex interventions to improve health. London: Medical Research Council.
- Medical Research Council. (2008). Developing and evaluating complex interventions: new guidance. London: Medical Research Council.
- Meltzer, H. (1972). Review of reviews in industrial psychology, 1950–1969. *Personnel Psychology*, 25(2), 201-222.
- Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med*, 46(1), 81-95. doi: 10.1007/s12160-013-9486-6
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- Miller, J. S., Musominali, S., Baganizi, M., & Paccione, G. A. (2014). A process evaluation of performance-based incentives for village health workers in Kisoro district, Uganda. *Human Resources for Health*, 12, 19. doi: 10.1186/1478-4491-12-19
- Mills, E. J., Montori, V. M., & Guyatt, G. H. (2004). Evidence-based clinical practice. *Brief Treatment and Crisis Intervention*, 4(2), 187.
- Millward, L., Kelly, M., & Nutbeam, D. (2003). Public health interventions research: the evidence. London: Health Development Agency.
- Ministry of Health Indonesia. (2011). Pedomani Umum Pengelolaan Posyandu (Posyandu General Guide). Jakarta, Indonesia.
- Mohan, A., Nassir, H., & Niazi, A. (2003). Does routine home visiting improve the return rate and outcome of DOTS patients who delay treatment? *La Revue de Santé de la Méditerranée orientale*, 9(4), 702-708.
- MOHCW Zimbabwe. (2010). The Village Health Worker Strategic Direction. Harare, Zimbabwe.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D. F., & Group, Q. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*, 354(9193), 1896-1900.

- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269.
- Moher, D., Schulz, K. F., Altman, D. G., & Group, C. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet*, 357(9263), 1191-1194.
- Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for Developers of Health Research Reporting Guidelines. *PLoS Med*, 7(2), e1000217. doi: 10.1371/journal.pmed.1000217
- Montgomery, P., Grant, S., Hopewell, S., Macdonald, G., Moher, D., Michie, S., & Mayo-Wilson, E. (2013a). Protocol for CONSORT-SPI: an extension for social and psychological interventions. *Implementation Science*, 8(1), 99.
- Montgomery, P., Underhill, K., Gardner, F., Operario, D., & Mayo-Wilson, E. (2013b). The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology*, 66(8), 874-882.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., . . . Wight, D. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *British Medical Journal*, 350, h1258.
- More, N. S., Bapat, U., Das, S., Alcock, G., Patil, S., Porel, M., . . . Osrin, D. (2012). Community mobilization in Mumbai slums to improve perinatal care and outcomes: a cluster randomized controlled trial. *Plos Medicine*, 9(7), e1001257. doi: 10.1371/journal.pmed.1001257
- Morris, S. (2015). *An integrated toolkit to save newborn lives and brains in Kenya*. [Unpublished protocol].
- Morrow, A. L., Guerrero, M. L., Shults, J., Calva, J. J., Lutter, C., Bravo, J., . . . Butterfoss, F. D. (1999). Efficacy of home-based peer counselling to promote exclusive breastfeeding: a randomised controlled trial. *Lancet*, 353(9160), 1226-1231. doi: 10.1016/s0140-6736(98)08037-4
- Moseley, A. M., Elkins, M. R., Herbert, R. D., Maher, C. G., & Sherrington, C. (2009). Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *Journal of Clinical Epidemiology*, 62(10), 1021-1030.
- Moszynski, P. (2011). One billion people are affected by global shortage of healthcare workers. *British Medical Journal*, 342, d696. doi: 10.1136/bmj.d696
- Movsisyan, A., Melendez-Torres, G., & Montgomery, P. (2015). Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE. *Journal of Clinical Epidemiology*.
- Mtango, F. D., & Neuvians, D. (1986). Acute respiratory infections in children under five years. Control project in Bagamoyo District, Tanzania. *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 80(6), 851-858. doi: 10.1016/0035-9203(86)90241-5
- Muir Gray, J. A. (2000). Evidence-Based Public Health. In L. Trinder (Ed.), *Evidence-based practice: A critical appraisal* (pp. 89-110). Oxford: Blackwell Science.
- Mullen, E. J., Shlonsky, A., Bledsoe, S. E., & Bellamy, J. L. (2005). From concept to implementation: challenges facing evidence-based social work. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(1), 61-84. doi: 10.1332/1744264052703159

- Mullen, E. J., & Streiner, D. L. (2004). The Evidence For and Against Evidence-Based Practice. *Brief Treatment and Crisis Intervention*, 4(2), 111-121. doi: 10.1093/brief-treatment/mhh009
- Mumtaz, Z., Salway, S., Nykiforuk, C., Bhatti, A., Ataulhjan, A., & Ayyalasomayajula, B. (2013). The role of social geography on Lady Health Workers' mobility and effectiveness in Pakistan. *Social Science & Medicine*, 91, 48-57. doi: 10.1016/j.socscimed.2013.05.007
- Munos, M., Guiella, G., Roberton, T., Maiga, A., Tiendrebeogo, A., Tam, Y., . . . Baya, B. (2016). Independent Evaluation of the Rapid Scale-Up Program to Reduce Under-Five Mortality in Burkina Faso. *American Journal of Tropical Medicine & Hygiene*, 94(3), 584-595. doi: 10.4269/ajtmh.15-0585
- Munro, S., Lewin, S., Swart, T., & Volmink, J. (2007). A review of health behaviour theories: how useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? *BMC Public Health*, 7(104). doi: 10.1186/1471-2458-7-104
- Murray, L., Cooper, P., Arteche, A., Stein, A., & Tomlinson, M. (2015). Randomized controlled trial of a home-visiting intervention on infant cognitive development in peri-urban South Africa. *Dev Med Child Neurol*. doi: 10.1111/dmcn.12873
- Musa, B. M., Iliyasu, Z., Yusuf, S. M., & Uloko, A. E. (2014). Systematic review and metanalysis on community based interventions in tuberculosis care in developing countries. *Nigerian Journal of Medicine: Journal of the National Association of Resident Doctors of Nigeria*, 23(2), 103-117.
- Mutamba, B. B., Ginneken, N. v., Paintain, L. S., Wandiembe, S., & Schellenberg, D. (2013). Roles and effectiveness of lay community health workers in the prevention of mental, neurological and substance use disorders in low and middle income countries: a systematic review. *BMC Health Services Research*, 13(412).
- Mwai, G. W., Mburu, G., Torpey, K., Frost, P., Ford, N., & Seeley, J. (2013). Role and outcomes of community health workers in HIV care in sub-Saharan Africa: a systematic review. *Journal of the International Aids Society*, 16(1), 18586\.
- Naimoli, J., Frymus, D., Quain, E., & Roseman, E. (2012). *Community and Formal Health System Support for Enhanced Community Health Worker Performance: Final Report*. Paper presented at the US Government CHW Evidence Summit, Washington, D.C.
- Naimoli, J. F., Frymus, D. E., Wuliji, T., Franco, L. M., & Newsome, M. H. (2014). A Community Health Worker "logic model": towards a theory of enhanced performance in low- and middle-income countries. *Human Resources for Health*, 12, 56.
- Nakigozi, G., Makumbi, F. E., Bwanika, J. B., Atuyambe, L., Reynolds, S. J., Kigozi, G., . . . Kamya, M. R. (2015). Impact of Patient-Selected Care Buddies on Adherence to HIV Care, Disease Progression, and Conduct of Daily Life Among Pre-antiretroviral HIV-Infected Patients in Rakai, Uganda: A Randomized Controlled Trial. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 70(1), 75-82. doi: 10.1097/qai.0000000000000710
- Nankabirwa, V., Tylleskar, T., Nankunda, J., Engebretsen, I. M., Sommerfelt, H., Tumwine, J., & PROMISE EBF Research Consortium. (2011). Malaria Parasitaemia among Infants and Its Association with Breastfeeding Peer

- Counselling and Vitamin A Supplementation: A Secondary Analysis of a Cluster Randomized Trial. *Plos One*, 6(7). doi: 10.1371/
- Neuner, F., Onyut, P. L., Ertl, V., Odenwald, M., Schauer, E., & Elbert, T. (2008). Treatment of posttraumatic stress disorder by trained lay counselors in an African refugee settlement: a randomized controlled trial. *J Consult Clin Psychol*, 76(4), 686-694. doi: 10.1037/0022-006X.76.4.686
- Newell, J. N., Baral, S. C., Pande, S. B., Bam, D. S., & Malla, P. (2006). Family-member DOTS and community DOTS for tuberculosis control in Nepal: cluster-randomised controlled trial. *Lancet*, 367(9514), 903-909.
- Newell, K. W. (1975). *Health by the people*: WHO, Geneva, Switzerland.
- Nkonki, L. L., Daviaud, E., Jackson, D., Chola, L., Doherty, T., Chopra, M., & Robberstad, B. (2014). Costs of promoting exclusive breastfeeding at community level in three sites in South Africa. *Plos One*, 9(1), e79784. doi: 10.1371/journal.pone.0079784
- Noronha, J. (2010). Brazil's march towards universal coverage. *Bulletin of the World Health Organization*, 88, 646-647.
- Nsibande, D., Doherty, T., Ijumba, P., Tomlinson, M., Jackson, D., Sanders, D., & Lawn, J. (2013). Assessment of the uptake of neonatal and young infant referrals by community health workers to public health facilities in an urban informal settlement, KwaZulu-Natal, South Africa. *BMC Health Services Research*, 13, 47. doi: 10.1186/1472-6963-13-47
- Nullis-Kapp, C. (2005). Health worker shortage could derail development goals. *Bulletin of the World Health Organization*, 83(1), 5-6.
- O'Donnell, A., Anderson, P., Newbury-Birch, D., Schulte, B., Schmidt, C., Reimer, J., & Kaner, E. (2014). The impact of brief alcohol interventions in primary healthcare: a systematic review of reviews. *Alcohol and Alcoholism*, 49(1), 66-78.
- Oakley, A. (1998). Experimentation and social interventions: a forgotten but important history. *BMJ (Clinical research ed.)*, 317(7167), 1239-1242.
- Oakley, A. (2006). Resistances to 'new' technologies of evaluation : education research in the UK as a case study. *Evidence & Policy: A Journal of Research, Debate and Practice*, 2(1), 63-87.
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal*, 332(7538), 413-416.
- Oakley, A., Strange, V., Toroyan, T., Wiggins, M., Roberts, I., & Stephenson, J. (2003). Using Random Allocation to Evaluate Social Interventions: Three Recent U.K. Examples. *The Annals of the American Academy of Political and Social Science*, 589(1), 170-189. doi: 10.1177/0002716203254765
- Odendaal, W., van Niekerk, A., Jordaan, E., & Seedat, M. (2009). The impact of a home visitation programme on household hazards associated with unintentional childhood injuries: a randomised controlled trial. *Accid Anal Prev*, 41(1), 183-190. doi: 10.1016/j.aap.2008.10.009
- Ofosu-Amaah, V. (1983). *National experience in the use of community health workers*: WHO.
- Ogilvie, D., Egan, M., Hamilton, V., & Petticrew, M. (2005). Systematic reviews of health effects of social interventions: 2. Best available evidence: how low should you go? *J Epidemiol Community Health*, 59(10), 886-892. doi: 10.1136/jech.2005.034199

- Ohnmar, Tun, M., San, S., Than, W., & Chongsuvivatwong, V. (2012). Effects of malaria volunteer training on coverage and timeliness of diagnosis: a cluster randomized controlled trial in Myanmar. *Malaria Journal*, 11, 309.
- Okwundu, C. I., Nagpal, S., Musekiwa, A., & Sinclair, D. (2013). Home- or community-based programmes for treating malaria. *Cochrane Database of Systematic Reviews*, 5, CD009527.
- Oladapo, O. T., Fawole, B., Blum, J., & Abalos, E. (2012). Advance misoprostol distribution for preventing and treating postpartum haemorrhage. *Cochrane Database of Systematic Reviews*, 2, CD009336.
- Olang'o, C. O., Nyamongo, I. K., & Aagaard-Hansen, J. (2010). Staff attrition among community health workers in home-based care programmes for people living with HIV and AIDS in western Kenya. *Health Policy*, 97(2-3), 232-237. doi: 10.1016/j.healthpol.2010.05.004
- Olsen, O., Middleton, P., Ezzo, J., Gøtzsche, P. C., Hadhazy, V., Herxheimer, A., . . . McIntosh, H. (2001). Quality of Cochrane reviews: assessment of sample from 1998. *British Medical Journal*, 323(7317), 829-832.
- Omer, K., Mhatre, S., Ansari, N., Laucirica, J., & Andersson, N. (2008). Evidence-based training of frontline health workers for door-to-door health promotion: a pilot randomized controlled cluster trial with Lady Health Workers in Sindh Province, Pakistan. *Patient Educ Couns*, 72(2), 178-185.
<http://onlinelibrary.wiley.com/doi/10.1016/j.pec.2008.02.018>
- One Million CHWs Campaign. (2016). Operations Room. Retrieved 16 May, 2016, from <http://1millionhealthworkers.org/operations-room/>
- Open Working Group of the General Assembly on Sustainable Development Goals. (2014). Open Working Group proposal for Sustainable Development Goals (document A/68/970). New York: United Nations.
- Osika, J., Altman, D., Ekblad, L., Katz, I., Nguyen, H., Rosenfeld, J., . . . Tapera, S. (2011). Zimbabwe Health System Assessment 2010 *Health Systems 20/30 Project*. Bethesda, MD.
- Owais, A., Hanif, B., Siddiqui, A. R., Agha, A., & Zaidi, A. K. (2011). Does improving maternal knowledge of vaccines impact infant immunization rates? A community-based randomized-controlled trial in Karachi, Pakistan. *BMC Public Health*, 11, 239.
- Oxman, A. D., & Guyatt, G. H. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, 44(11), 1271-1278.
- Pachón, H., Schroeder, D. G., Marsh, D. R., Dearden, K. A., Ha, T. T., & Lang, T. T. (2002). Effect of an integrated child nutrition intervention on the complementary food intake of young children in rural north Viet Nam. *Food and Nutrition Bulletin*, 23(4 Suppl), 62-69.
- Paintain, L. S., Willey, B., Kendege, S., Sharkey, A., Kim, J., Buj, V., . . . Ngongo, N. (2014). Review Article: Community Health Workers and Stand-Alone or Integrated Case. Management of Malaria: A Systematic Literature Review. *American Journal of Tropical Medicine and Hygiene*, 91(3), 461-470. doi: 10.4269/ajtmh.14-0094
- Palazuelos, D., Ellis, K., Im, D. D., Peckarsky, M., Schwarz, D., Farmer, D. B., . . . Mitnick, C. D. (2013). 5-SPICE: the application of an original framework for community health worker program design, quality improvement and research agenda setting. *Global Health Action*, 6, 19658.

- Patel, R., Chang, T., Greysen, S. R., & Chopra, V. (2015). Social Media Use in Chronic Disease: A Systematic Review and Novel Taxonomy. *Am J Med*, 128(12), 1335-1350. doi: 10.1016/j.amjmed.2015.06.015
- Patouillard, E., Conteh, L., Webster, J., Kweku, M., Chandramohan, D., & Greenwood, B. (2011). Coverage, adherence and costs of intermittent preventive treatment of malaria in children employing different delivery strategies in Jasikan, Ghana. *Plos One*, 6(11), e24871.
- Pegurri, E., Fox-Rushby, J. A., & Damian, W. (2005). The effects and costs of expanding the coverage of immunisation services in developing countries: a systematic literature review. *Vaccine*, 23(13), 1624 - 1635.
- Pence, B. W., Nyarko, P., Phillips, J. F., & Debpuur, C. (2005). The effect of community nurses and health volunteers on child mortality: The Navrongo Community Health and Family Planning Project. *Population Council: Policy Research Division Working Papers*, 2000, 1-27.
- Pence, B. W., Nyarko, P., Phillips, J. F., & Debpuur, C. (2007). The effect of community nurses and health volunteers on child mortality: the Navrongo Community Health and Family Planning Project. *Scandinavian Journal of Public Health*, 35(6), 599-608. doi: 10.1080/14034940701349225
- Penfold, S., Manzi, F., Mkumbo, E., Temu, S., Jaribu, J., Shamba, D. D., . . . Armstrong Schellenberg, J. (2014). Effect of home-based counselling on newborn care practices in southern Tanzania one year after implementation: A cluster-randomised controlled trial. *BMC Pediatrics*, 14(1). doi: <http://dx.doi.org/10.1186/1471-2431-14-187>
- Pérez, L. M., & Martinez, J. (2008). Community health workers: social justice and policy advocates for community health and well-being. *American Journal of Public Health*, 98(1), 11-14.
- Perry, H., & Zulliger, R. (2012). How effective are community health workers? Baltimore: Johns Hopkins Bloomberg School of Public Health.
- Perry, H. B., Dhillon, R. S., Liu, A., Chitnis, K., Panjabi, R., Palazuelos, D., . . . Nyenswah, T. (2016). Community health worker programmes after the 2013–2016 Ebola outbreak. *Bulletin of the World Health Organization*, 94(7), 551-553. doi: 10.2471/blt.15.164020
- Perry, H. B., Zulliger, R., & Rogers, M. M. (2014). Community health workers in low-, middle-, and high-income countries: an overview of their history, recent evolution, and current effectiveness. *Annu Rev Public Health*, 35, 399-421. doi: 10.1146/annurev-publhealth-032013-182354
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991-996.
- Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2002). " Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *The Cochrane Library*.
- Petticrew, M. (2011). When are complex interventions 'complex'? When are simple interventions 'simple'? *Eur J Public Health*, 21(4), 397-398. doi: 10.1093/eurpub/ckr084
- Petticrew, M. (2015). Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. *Syst Rev*, 4, 36. doi: 10.1186/s13643-015-0027-1

- Petticrew, M., Anderson, L., Elder, R., Grimshaw, J., Hopkins, D., Hahn, R., . . . Welch, V. (2013). Complex interventions and their implications for systematic reviews: a pragmatic approach. *Journal of Clinical Epidemiology*, 66(11), 1209-1214. doi: 10.1016/j.jclinepi.2013.06.004
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, 57(7), 527-529.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford, UK: Blackwell Publishing.
- Petticrew, M., & Roberts, H. (2008). Systematic reviews--do they 'work' in informing decision-making around health inequalities? *Health Econ Policy Law*, 3(Pt 2), 197-211. doi: 10.1017/S1744133108004453
- Pieper, D., Antoine, S.-L., Mathes, T., Neugebauer, E. A. M., & Eikermann, M. (2014a). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, 67(4), 368-375. doi: <http://dx.doi.org/10.1016/j.jclinepi.2013.11.007>
- Pieper, D., Antoine, S. L., Morfeld, J. C., Mathes, T., & Eikermann, M. (2014b). Methodological approaches in conducting overviews: current state in HTA agencies. *Research Synthesis Methods*, 5(3), 187-199. doi: 10.1002/jrsm.1107
- Pieper, D., Antoine, S. L., Neugebauer, E. A. M., & Eikermann, M. (2014c). Up-to-dateness of reviews is often neglected in overviews: A systematic review. *Journal of Clinical Epidemiology*, 67(12), 1302-1308.
- Pieper, D., Buechter, R., Jerinic, P., & Eikermann, M. (2012). Overviews of reviews often have limited rigor: a systematic review. *Journal of Clinical Epidemiology*, 65(12), 1267-1273.
- Pieper, D., Mathes, T., & Eikermann, M. (2014d). Impact of choice of quality appraisal tool for systematic reviews in overviews. *Journal of Evidence-Based Medicine*, 7(2), 72-78.
- Pildal, J., Hrobjartsson, A., Jørgensen, K., Hilden, J., Altman, D., & Gøtzsche, P. (2007). Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *International Journal of Epidemiology*, 36(4), 847-857.
- Pineda, M. A., Bertrand, J. T., Santiso, R., & Guerra, S. (1983). Increasing the effectiveness of community workers through training of spouses: a family planning experiment in Guatemala. *Public Health Reports*, 98(3), 273.
- Piso, B., Semlitsch, T., Reinsperger, I., Breuer, J., Kaminski-Hartenthaler, A., Kien, C., . . . Siebenhofer, A. (2015). [Practical experience with overviews of reviews - valuable decision aid or academic exercise?]. *Zeitschrift für Evidenz Fortbildung und Qualität im Gesundheitswesen*, 109(4-5), 300-308.
- Pitchforth, J. O., & Mengersen, K. L. (2012). *Bayesian Meta-Analysis Case Studies in Bayesian Statistical Modelling and Analysis* (pp. 118-140): John Wiley & Sons, Ltd.
- Pitt, C., Tawiah, T., Soremekun, S., ten Asbroek, A. H. A., Manu, A., Tawiah-Agyemang, C., . . . Hanson, K. (2016). Cost and cost-effectiveness of newborn home visits: findings from the Newhints cluster-randomised controlled trial in rural Ghana. *The Lancet Global Health*, 4(1), e45-e56. doi: 10.1016/s2214-109x(15)00207-7
- Pongvongsa, T., Nonaka, D., Kobayashi, J., Mizoue, T., Phongmany, P., & Moji, K. (2011). Determinants Of Monthly Reporting By Village Health Volunteers In A Poor Rural District Of Lao Pdr. *Southeast Asian Journal of Tropical Medicine and Public Health*, 42(5), 1269-1281.

- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., . . . Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: ESRC Methods Programme.
- Porta, M., Greenland, S., Hernán, M., dos Santos Silva, I., & Last, J. M. (2014). *A dictionary of epidemiology*: Oxford University Press, USA.
- Powell, C., Baker-Henningham, H., Walker, S., Gernay, J., & Grantham-McGregor, S. (2004). Feasibility of integrating early stimulation into primary care for undernourished Jamaican children: cluster randomised controlled trial. *British Medical Journal*. doi: 10.1136/bmj.38132.503472.&C
- Powell, C., & Grantham-McGregor, S. (1989). Home visiting of varying frequency and child development. *Pediatrics*, 84(1), 157-164.
- Pradeep, J., Isaacs, A., Shanbag, D., Selvan, S., & Srinivasan, K. (2014). Enhanced care by community health workers in improving treatment adherence to antidepressant medication in rural women with major depression. *Indian J Med Res*, 139(2), 236-245.
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, 42(1), 185-227.
- Presidência da República Brasil, Lei No 10.507, De 10 De Julho De 2002 C.F.R. (2002).
- Puchalski Ritchie, L. M., van Lettow, M., Barnsley, J., Chan, A. K., Joshua, M., Martiniuk, A. L., . . . Zwarenstein, M. (2012). Evaluation of lay health workers' needs to effectively support anti-tuberculosis treatment adherence in Malawi. *Int J Tuberc Lung Dis*, 16(11), 1492-1497. doi: 10.5588/ijtld.12.0206
- Puett, C., Alderman, H., Sadler, K., & Coates, J. (2013). 'Sometimes they fail to keep their faith in us': community health worker perceptions of structural barriers to quality of care and community utilisation of services in Bangladesh. *Matern Child Nutr*. doi: 10.1111/mcn.12072
- Puett, C., Coates, J., Alderman, H., Sadruddin, S., & Sadler, K. (2012). Does greater workload lead to reduced quality of preventive and curative care among community health workers in Bangladesh? *Food and Nutrition Bulletin*, 33(4), 273-287.
- R Core Team. (2015). R Development Core Team (2010). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. from <http://www.R-project.org/>
- Rabbani, F., Mukhi, A. A., Perveen, S., Gul, X., Iqbal, S. P., Qazi, S. A., . . . Aftab, W. (2014). Improving community case management of diarrhoea and pneumonia in district Badin, Pakistan through a cluster randomised study--the NIGRAAN trial protocol. *Implementation Science*, 9, 186.
- Rahman, A., Fisher, J., Bower, P., Luchters, S., Thach, T., Yasamy, M. T., . . . Waheed, W. (2013). Interventions for common perinatal mental disorders in women in low- and middle-income countries: a systematic review and meta-analysis. *Bulletin of the World Health Organization*, 91(8), 593-601.
- Rahman, A., Iqbal, Z., Roberts, C., & Husain, N. (2009). Cluster randomized trial of a parent-based intervention to support early development of children in a low-income country. *Child: Care, Health & Development*, 35(1), 56-62.
- Rahman, A., Malik, A., Sikander, S., Roberts, C., & Creed, F. (2008). Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *Lancet*, 372(9642), 902-909.

- Rahman, A., Sikander, S., Malik, A., Ahmed, I., Tomenson, B., & Creed, F. (2012). Effective treatment of perinatal depression for women in debt and lacking financial empowerment in a low-income country. *Br J Psychiatry*, 201(6), 451-457. doi: 10.1192/bjp.bp.112.109207
- Rahman, S. M., Ali, N. A., Jennings, L., Seraji, M. H. R., Mannan, I., Shah, R., . . . Winch, P. J. (2010). Factors affecting recruitment and retention of community health workers in a newborn care intervention in Bangladesh. *Human Resources for Health*, 8. doi: 10.1186/1478-4491-8-12
- Ramadas, K., Sankaranarayanan, R., Jacob, B. J., Thomas, G., Somanathan, T., Mahé, C., . . . Nair, M. K. (2003). Interim results from a cluster randomized controlled oral cancer screening trial in Kerala, India. *Oral Oncology*, 39(6), 580-588. doi: 10.1016/s1368-8375(03)00041-1
- Ramadas, K., Sauvaget, C., Thomas, G., Fayette, J. M., Thara, S., & Sankaranarayanan, R. (2010). Effect of tobacco chewing, tobacco smoking and alcohol on all-cause and cancer mortality: a cohort study from Trivandrum, India. *Cancer Epidemiol*, 34(4), 405-412. doi: 10.1016/j.canep.2010.04.006
- Ramaiah, K. D., Vijay Kumar, K. N., Chandrakala, A. V., Augustin, D. J., Appavoo, N. C., & Das, P. K. (2001). Effectiveness of community and health services-organized drug delivery strategies for elimination of lymphatic filariasis in rural areas of Tamil Nadu, India. *Tropical Medicine & International Health*, 6(12), 1062-1069.
- Ranganathan, S. R. (1967). Prolegomena to library classification. *The Five Laws of Library Science*.
- Rao, J., & Scott, A. (1992). A simple method for the analysis of clustered binary data. *Biometrics*, 577-585.
- Razee, H., Whittaker, M., Jayasuriya, R., Yap, L., & Brentnall, L. (2012). Listening to the rural health workers in Papua New Guinea – The social factors that influence their motivation to work. *Social Science & Medicine*, 75(5), 828-835. doi: 10.1016/j.socscimed.2012.04.013
- Reader, T. W., Gillespie, A., & Roberts, J. (2014). Patient complaints in healthcare systems: a systematic review and coding taxonomy. *BMJ Qual Saf*, 23(8), 678-689. doi: 10.1136/bmjqs-2013-002437
- Rehfuess, E. A., & Akl, E. A. (2013). Current experience with applying the GRADE approach to public health interventions: an empirical study. *BMC Public Health*, 13(1), 9.
- Richard, J., & Pillemer, D. B. (1984). *Summing Up*: Harvard University Press.
- Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). *Qualitative research practice: A guide for social science students and researchers*: Sage.
- Robinson, K. A., Whitlock, E. P., O'Neil, M. E., Anderson, J. K., Hartling, L., Dryden, D. M., . . . Berkman, N. D. (2014). *Integration of Existing Systematic Reviews*. Rockville, MD, USA: Agency for Healthcare Research and Quality (US).
- Roemer, M. I. (1977). Primary care and physician extenders in affluent countries. *International Journal of Health Services*, 7(4), 545-555.
- Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29-48. doi: 10.1177/1356389007084674

- Rohde, J. (1993). Indonesias Posyandus: accomplishments and future challenges. In J. Rohde, M. Chatterjee, & D. Morley (Eds.), *Reaching Health for All* (pp. 135-157). New Delhi, India: Oxford University Press.
- Ronaghy, H. A., Najarzadeh, E., Schwartz, T. A., Russel, S. S., Solter, S., & Zeighami, B. (1976). The front line health worker: selection, training, and performance. *Am J Public Health*, 66(3), 273-277.
- Rose, H. (1994). *Love, Power, and Knowledge: Towards a Feminist Transformation of the Sciences*. Cambridge: Polity.
- Rosenfeld, L., & Morville, P. (2002). *Information architecture for the world wide web*: "O'Reilly Media, Inc."
- Rotheram-Borus, M. J., Tomlinson, M., le Roux, I. M., Harwood, J. M., Comulada, S., O'Connor, M. J., . . . Worthman, C. M. (2014). A cluster randomised controlled effectiveness trial evaluating perinatal home visiting among South African mothers/infants. *Plos One*, 9(10), e105934.
- Rothwell, P. M. (2005). Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*, 365(9454), 176-186.
- Rounsaville, B. J., Carroll, K. M., & Onken, L. S. (2001). A stage model of behavioral therapies research: Getting started and moving on from stage I. *Clinical Psychology: Science and Practice*, 8(2), 133-142.
- Ruebush, T. K., Weller, S. C., & Klein, R. E. (1994). Qualities of an ideal volunteer community malaria worker: A comparison of the opinions of community residents and national malaria service staff. *Social Science & Medicine*, 39(1), 123-131.
- Ruizendaal, E., Dierickx, S., Peeters Grietens, K., Schallig, H. D., Pagnoni, F., & Mens, P. F. (2014). Success or failure of critical steps in community case management of malaria with rapid diagnostic tests: a systematic review. *Malaria Journal*, 13, 229.
- Ryan, R. E., Kaufman, C. A., & Hill, S. J. (2009). Building blocks for meta-synthesis: data integration tables for summarising, mapping, and synthesising evidence on interventions for communicating with health consumers. *BMC Med Res Methodol*, 9, 16. doi: 10.1186/1471-2288-9-16
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*, 56(2), 119-127.
- Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *J Epidemiol Community Health*, 58(7), 538-545. doi: 10.1136/jech.2003.011585
- Rycroft-Malone, J. (2007). Theory and knowledge translation: setting some coordinates. *Nursing Research*, 56(4), S78-S85.
- Sackett, D. L. (2000). *Evidence-based medicine*: Wiley Online Library.
- Sackett, D. L., & Rosenberg, W. M. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312(13), 71-72.
- Sackett, D. L., & Wennberg, J. E. (1997). Choosing the best research design for each question. *British Medical Journal*, 315(7123), 1636.
- Sadrudin, S., Shehzad, S., Bari, A., Khan, A., Ibad ul, H., Khan, A., & Qazi, S. (2012). Household costs for treatment of severe pneumonia in Pakistan. *American Journal of Tropical Medicine & Hygiene*, 87(5 Suppl), 137-143.
- Saha, K., Rawat, R., Khaled, A., Haque, R., Afsana, K., Sanghvi, T., . . . Menon, P. (2013). A Comprehensive Training And Supervision Program Improved

- Frontline Health Workers' Knowledge Of Infant And Young Child Feeding In Rural Bangladesh. *Annals of Nutrition and Metabolism*, 63, 992-992.
- Saleem, A. F., Sadia, M., Naila, B.-A., & Zaidi, A. K. M. (2014). Impact of maternal education about complementary feeding on their infants' nutritional outcomes in low- and middle-income households: a community-based randomized interventional study in Karachi, Pakistan. *Journal of Health, Population and Nutrition*, 32(4), 623-633.
- Sanders, D. (1993). The potential and limits of health sector reforms in Zimbabwe. In J. Rohde, M. Chatterjee, & D. Morley (Eds.), *Reaching Health for All* (pp. 239-265). New Delhi, India: Oxford University Press.
- Sankaranarayanan, R., Mathew, B., Jacob, B. J., Thomas, G., Somanathan, T., Pisani, P., . . . Abraham, E. (2000). Early Findings from a Community-Based, Cluster-Randomized, Controlled Oral Cancer Screening Trial in Kerala, India. *Cancer*, 88(3), 664-673.
- Sankaranarayanan, R., Ramadas, K., Thara, S., Muwonge, R., Thomas, G., Anju, G., & Mathew, B. (2013). Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India. *Oral Oncology*, 49(4), 314-321. doi: 10.1016/j.oraloncology.2012.11.004
- Sankaranarayanan, R., Ramadas, K., Thomas, G., Muwonge, R., Thara, S., Mathew, B., . . . Trivandrum Oral Cancer Screening Study, G. (2005). Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet*, 365(9475), 1927-1933. doi: 10.1016/S0140-6736(05)66658-5
- Santaguida, P. L., Keshavarz, H., Carlesso, L. C., Lomotan, M., Gross, A., Macdermid, J. C., . . . Group, I. W. (2013). A description of the methodology used in an overview of reviews to evaluate evidence on the treatment, harms, diagnosis/classification, prognosis and outcomes used in the management of neck pain. *The Open Orthopaedics Journal*, 7, 461-472.
- Sarrami-Foroushani, P., Travaglia, J., Debono, D., Clay-Williams, R., & Braithwaite, J. (2015). Scoping Meta-Review: Introducing a New Methodology. *Clinical and Translational Science*, 8(1), 77-81.
- Satterfield, J. M., Spring, B., Brownson, R. C., Mullen, E. J., Newhouse, R. P., Walker, B. B., & Whitlock, E. P. (2009). Toward a Transdisciplinary Model of Evidence-Based Practice. *Milbank Quarterly*, 87(2), 368-390.
- Saunders, R. (1996). Economic Structural Adjustment Programme (ESAP)'s Fables II. *Southern Africa Report Archive*, 11(4).
- Schiffman, J., Darmstadt, G. L., Agarwal, S., & Baqui, A. H. (2010). Community-based intervention packages for improving perinatal health in developing countries: a review of the evidence. *Semin Perinatol*, 34(6), 462-476. doi: 10.1053/j.semperi.2010.09.008
- Schneider, H., Hlophe, H., & van Rensburg, D. (2008). Community health workers and the response to HIV/AIDS in South Africa: tensions and prospects. *Health Policy and Planning*, 23(3), 179-187. doi: 10.1093/heapol/czn006
- Schroeder, D., Pachón, H., Dearden, K., Ha, T., Lang, T., & Marsh, D. (2002). An integrated child nutrition intervention improved growth of younger, more malnourished children in northern Viet Nam. *Food and Nutrition Bulletin*, 23(4 Suppl), 50-58.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine*, 8(1), 1.

- Schünemann, H., Brožek, J., & Oxman, G. (2015). Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. 2013.
- Schünemann, H. J., Fretheim, A., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Research Policy and Systems*, 4(1), 25.
- Schwarzer, G. (2015). meta: General Package for Meta-Analysis (Version R package version 4.3-0). Retrieved from <http://CRAN.R-project.org/package=meta>
- Segall, M. (1977). Health and national liberation in the People's Republic of Mozambique. *International Journal of Health Services: Planning, Administration, Evaluation*, 7(2), 319.
- Segall, M. (2003). District health systems in a neoliberal world: a review of five key policy areas. *Int J Health Plann Manage*, 18(S1), S5-S26.
- Selke, H. M., Kimaiyo, S., Sidle, J. E., Vedanthan, R., Tierney, W. M., Shen, C., . . . Wools-Kaloustian, K. (2010). Task-shifting of antiretroviral delivery from health care workers to persons living with HIV/AIDS: clinical outcomes of a community-based program in Kenya. *Journal of Acquired Immune Deficiency Syndromes: JAIDS*, 55(4), 483-490.
- Shea, B., Moher, D., Graham, I., Pham, B., & Tugwell, P. (2002). A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Evaluation & the Health Professions*, 25(1), 116-129.
- Shea, B. J., Bouter, L. M., Peterson, J., Boers, M., Andersson, N., Ortiz, Z., . . . Grimshaw, J. M. (2007a). External validation of a measurement tool to assess systematic reviews (AMSTAR). *Plos One*, 2(12), e1350. doi: 10.1371/journal.pone.0001350
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., . . . Bouter, L. M. (2007b). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*, 7.
- Sheldon, B. (2001). The Validity of Evidence-Based Practice in Social Work : A Reply to Stephen Webb. *British Journal of Social Work*, 31(5), 801-809. doi: DOI 10.1093/bjsw/31.5.801
- Shelley, K. D. (2015). Application of a systems thinking lens to the design of effective and sustainable scale-up of national community health programs. *Annals of Global Health*, 81 (1), 171.
- Shepperd, S., Lewin, S., Straus, S., Clarke, M., Eccles, M. P., Fitzpatrick, R., . . . Sheikh, A. (2009). Can we systematically review studies that evaluate complex interventions? *Plos Medicine*, 6(8), e1000086.
- Shi, L. (1993). Health care in China: a rural-urban comparison after the socioeconomic reforms. *Bulletin of the World Health Organization*, 71(6), 723.
- Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. *British Medical Journal*, 336(7656), 1281-1283. doi: 10.1136/bmj.39569.510521.AD
- Shojania, K., Sampson, M. J., Ansari, M. T., Ji, J., Doucette, S., & Moher, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine*.
- Siabani, S., Driscoll, T., Davidson, P. M., & Leeder, S. R. (2015). Efficacy of a home-based educational strategy involving community health volunteers in improving self-care in patients with chronic heart failure in western Iran: A

- randomized controlled trial. *Eur J Cardiovasc Nurs*. doi: 10.1177/1474515115585651
- Sidel, V. W. (1972). The barefoot doctors of the People's Republic of China. *New England Journal of Medicine*, 286(24), 1292-1300.
- Sidibe, M., & Campbell, J. (2015). Reversing a global health workforce crisis. *Bulletin of the World Health Organization*, 93(1), 3. doi: 10.2471/BLT.14.151209
- Singh Jasvinder, A., Wells George, A., Christensen, R., Tanjong Ghogomu, E., Maxwell, L., MacDonald John, K., . . . Buchbinder, R. (2011). Adverse effects of biologics: a network meta-analysis and Cochrane overview. *Cochrane Database of Systematic Reviews*, (2). <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008794.pub2/abstract> doi:10.1002/14651858.CD008794.pub2
- Singh, P., & Sachs, J. D. (2013). 1 million community health workers in sub-Saharan Africa by 2015. *Lancet*, 382(9889), 363-365. doi: 10.1016/s0140-6736(12)62002-9
- Sloan, N. L., Ahmed, S., Mitra, S. N., Choudhury, N., & Chowdhury, M. (2008). Community-based Kangaroo Mother Care to prevent neonatal and infant mortality: A randomized, controlled cluster trial. *Pediatrics*, 121(5), e1047-e1059. doi: <http://dx.doi.org/10.1542/peds.2007-0076>
- Smith, J. M., Gubin, R., Holston, M. M., Fullerton, J., & Prata, N. (2013). Misoprostol for postpartum hemorrhage prevention at home birth: an integrative review of global implementation experience to date. *BMC Pregnancy & Childbirth*, 13, 44.
- Smith, V., Devane, D., Begley, C. M., & Clarke, M. (2011). Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol*, 11(1), 15.
- Smithells, R. (1974). Iatrogenic hazards and their effects. *Postgraduate Medical Journal*, 51, suppl 15: 39-52.
- Sokal, R. (1974). Classification: Purposes, principles, progress, prospects. *Science*, 185(4157), 1115-1123.
- Soofi, S., Ahmed, S., Fox, M. P., MacLeod, W. B., Thea, D. M., Qazi, S. A., & Bhutta, Z. A. (2012). Effectiveness of community case management of severe pneumonia with oral amoxicillin in children aged 2-59 months in Matiari district, rural Pakistan: a cluster-randomised controlled trial. *Lancet*, 379(9817), 729-737.
- Spaulding, A. B., Brickley, D. B., Kennedy, C., Almers, L., Packel, L., Mirjahangir, J., . . . Mbizvo, M. (2009). Linking family planning with HIV/AIDS interventions: a systematic review of the evidence. *AIDS*, 23 Suppl 1, S79-88.
- Sranacharoenpong, K., & Hanning, R. M. (2012). Diabetes Prevention Education Program for Community Health Care Workers in Thailand. *Journal of Community Health*, 37(3), 610-618. doi: 10.1007/s10900-011-9491-2
- Sripaipan, T., Schroeder, D. G., Marsh, D., Pachón, H., Dearden, K., Thu Ha, T., & Thi Lang, T. (2002). Effect of an integrated nutrition program on child morbidity due to respiratory infection and diarrhea in northern Viet Nam. *Food and Nutrition Bulletin*, 23(4).
- Standing, H., & Chowdhury, A. M. R. (2008). Producing effective knowledge agents in a pluralistic environment: What future for community health workers? *Social Science & Medicine*, 66(10), 2096-2107.

- Stanton, C. K., Newton, S., Mullany, L. C., Cofie, P., Tawiah Agyemang, C., Adiibokah, E., . . . Owusu-Agyei, S. (2013). Effect on Postpartum Hemorrhage of Prophylactic Oxytocin (10 IU) by Injection by Community Health Officers in Ghana: A Community-Based, Cluster-Randomized Trial. *Plos Medicine*, 10(10). doi: <http://dx.doi.org/10.1371/journal.pmed.1001524>
- Starfield, B. (1992). *Primary care: concept, evaluation, and policy*: Oxford University Press, USA.
- Stark, R. (1985). Lay workers in primary health care: victims in the process of social transformation. *Social Science & Medicine*, 20(3), 269-275.
- Stavri, Z., & Michie, S. (2012). Classification systems in behavioural science: current systems and lessons from the natural, medical and social sciences. *Health Psychology Review*, 6(1), 113-140. doi: 10.1080/17437199.2011.641101
- Steultjens, E. M., Dekker, J., Bouter, L. M., Leemrijse, C. J., & van den Ende, C. H. (2005). Evidence of the efficacy of occupational therapy in different conditions: an overview of systematic reviews. *Clinical Rehabilitation*, 19(3), 247-254. doi: 10.1191/0269215505cr870oa
- Stinson, W., Malianga, L., Marquez, L., & Madubuike, C. (1998). Quality supervision. *Quality Assurance Briefs*, 7(1), 4-6.
- Strachan, D. L., Kallander, K., ten Asbroek, A. H. A., Kirkwood, B., Meek, S. R., Benton, L., . . . Hill, Z. (2012). Interventions to improve Motivation and Retention of Community Health Workers Delivering Integrated Community Case Management (iCCM): Stakeholder Perceptions and Priorities. *American Journal of Tropical Medicine and Hygiene*, 87(5), 111-119. doi: 10.4269/ajtmh.2012.12-0030
- Straus, S. E., & McAlister, F. A. (2000). Evidence-based medicine: a commentary on common criticisms. *Canadian Medical Journal*, 163(7), 837-841.
- Straus, S. E., Richardson, W. S., Glasziou, P., & Haynes, R. B. (2005). *Evidence-based medicine: how to practice and teach EBM*. Oxford, UK: Elsevier.
- Subramanian, S., Sankaranarayanan, R., Bapat, B., Somanathan, T., Thomas, G., Mathew, B., . . . Ramadas, K. (2009). Cost-effectiveness of oral cancer screening: results from a cluster randomized controlled trial in India. *Bulletin of the World Health Organization*, 87(3), 200-206. doi: 10.2471/blt.08.053231
- Sunguya, B. F., Poudel, K. C., Mlunde, L. B., Shakya, P., Urassa, D. P., Jimba, M., & Yasuoka, J. (2013). Effectiveness of nutrition training of health workers toward improving caregivers' feeding practices for children aged six months to two years: a systematic review. *Nutrition Journal*, 12, 66.
- Suryakusuma, J. (2013, December 4). Children, the “Hunger Games” and posyandu. *The Jakarta Post*. Retrieved from <http://www.thejakartapost.com/news/2013/12/04/children-hunger-games-and-posyandu.html>
- Svitone, E. C., Garfield, R., Vasconcelos, M. I., & Craveiro, V. A. (2000). Primary health care lessons from the Northeast of Brazil: the Agentes de Saúde Program. *Revista Panamericana de Salud Publica*, 7(5), 293-301.
- Swart, L., van Niekerk, A., Seedat, M., & Jordaan, E. (2008). Paraprofessional home visitation program to prevent childhood unintentional injuries in low-income communities: a cluster randomized controlled trial. *Injury Prevention*, 14(3), 164-169.
- Swider, S. M. (2002). Outcome effectiveness of community health workers: an integrative literature review. *Public Health Nursing*, 19(1), 11-20.

- Taylor, C. E., Alter, J., Grover, P., Sangal, S., Andrews, S., & Takulia, H. (1976). *Doctors for the villages: study of rural internships in seven indian medical colleges*: Asia Publishing House.
- Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, 22(13), 2113-2126.
- Terris, M. (1977). Issues in primary care. False starts and lesser alternatives. *Bulletin of the New York Academy of Medicine*, 53(1), 129.
- The Cochrane Collaboration. (2016). Cochrane Glossary. Retrieved July 10, 2016, from <http://community-archive.cochrane.org/glossary/5#letters>
- The University of Arizona Rural Health Office and College of Public Health. (1998). Community Health Worker Evaluation Tool Kit.
- Theodoratou, E., Tzoulaki, I., Zgaga, L., & Ioannidis, J. P. (2014). Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. *British Medical Journal*, 348.
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*, 8(1), 1.
- Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., . . . Kavanagh, J. (2004). Integrating qualitative research with trials in systematic reviews. *British Medical Journal*, 328(7446), 1010-1012.
- Thompson, S. G., & Higgins, J. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21(11), 1559-1573.
- Thompson, S. G., & Higgins, J. (2005). Can meta-analysis help target interventions at individuals most likely to benefit? *Lancet*, 365(9456), 341-346.
- Thomson, D. (2014). Evidence synthesis in child health: Overviews of reviews. *Evidence-Based Child Health*, 9(1), 1-2. doi: 10.1002/ebch.1963
- Thomson, D., Foisy, M., Oleszczuk, M., Wingert, A., Chisholm, A., & Hartling, L. (2013). Overview of reviews in child health: evidence synthesis and the knowledge base for a specific population. *Evidence-Based Child Health: A Cochrane Review Journal*, 8(1), 3-10. doi: 10.1002/ebch.1897
- Thomson, D., Russell, K., Becker, L., Klassen, T., & Hartling, L. (2010). The evolution of a new publication type: Steps and challenges of producing overviews of reviews. *Research Synthesis Methods*, 1(3-4), 198-211. doi: 10.1002/jrsm.30
- Threlfall, A. G., Meah, S., Fischer, A. J., Cookson, R., Rutter, H., & Kelly, M. P. (2015). The appraisal of public health interventions: the use of theory. *Journal of Public Health*, 37(1), 166-171.
- Tian, M., Ajay, V., Dunzhu, D., Hameed, S., Li, X., Liu, Z., . . . Yan, L. (2015). A Cluster-Randomized Controlled Trial of a Simplified Multifaceted Management Program for Individuals at High Cardiovascular Risk (SimCard Trial) in Rural Tibet, China, and Haryana, India. *Circulation*. doi: 10.1161/CIRCULATIONAHA.115.015373
- Tine, R. C., Faye, B., Ndour, C. T., Ndiaye, J. L., Ndiaye, M., Bassene, C., . . . Gaye, O. (2011). Impact of combining intermittent preventive treatment with home management of malaria in children less than 10 years in a rural area of Senegal: a cluster randomized trial. *Malaria Journal*, 10, 358.
- Tine, R. C., Ndour, C. T., Faye, B., Cairns, M., Sylla, K., Ndiaye, M., . . . Gaye, O. (2014). Feasibility, safety and effectiveness of combining home based malaria management and seasonal malaria chemoprevention in children less than 10

- years in Senegal: a cluster-randomised trial. *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 108(1), 13-21.
- Toczek, A., Cox, H., Cros, P. d., Cooke, G., & Ford, N. (2013). Strategies for reducing treatment default in drug-resistant tuberculosis: systematic review and meta-analysis. *The International Journal of Tuberculosis and Lung Disease*, 17(3), 299-307.
- Tomlinson, M., Doherty, T., Ijumba, P., Jackson, D., Lawn, J., Persson, L. A., . . . Chopra, M. (2014). Goodstart: a cluster randomised effectiveness trial of an integrated, community-based package for maternal and newborn care, with prevention of mother-to-child transmission of HIV in a South African township. *Tropical Medicine & International Health*, 19(3), 256-266.
- Tomlinson, M., Rotheram-Borus, M. J., Harwood, J., le Roux, I. M., O'Connor, M., & Worthman, C. (2015). Community health workers can improve child growth of antenatally-depressed, South African mothers: a cluster randomized controlled trial. *BMC Psychiatry*, 15, 225. doi: 10.1186/s12888-015-0606-7
- Trickett, E. J., & Birman, D. (1989). Taking ecology seriously: A community development approach to individually based preventive interventions in schools. In B. E. Compas (Ed.), *Primary Prevention and Promotion in the Schools* (pp. 361-390). Thousand Oaks, CA, US: Sage Publications.
- Trickett, E. J., Kelly, J. G., & Vincent, T. A. (1985). The Spirit of Ecological Inquiry in Community Research. In E. Susskind & D. Klein (Eds.), *Community Research: Methods, Paradigms, And Applications* (pp. 283-333). New York: Praeger.
- Tripathy, P., Nair, N., Barnett, S., Mahapatra, R., Borghi, J., Rath, S., . . . Costello, A. (2010). Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomised controlled trial. *Lancet*, 375(9721), 1182-1192. doi: 10.1016/S0140-6736(09)62042-0
- Tripathy, P., Nair, N., Sinha, R., Rath, S., Gope, R. K., Rath, S., . . . Prost, A. (2016). Effect of participatory women's groups facilitated by Accredited Social Health Activists on birth outcomes in rural eastern India: a cluster-randomised controlled trial. *The Lancet Global Health*, 4(2), e119-e128. doi: 10.1016/s2214-109x(15)00287-9
- Tugwell, P., Kottnerus, J. A., & Idzerda, L. (2013). Complex interventions—how should systematic reviews of their impact differ from reviews of simple or complicated interventions? *Journal of Clinical Epidemiology*, 66(11), 1195-1196. doi: 10.1016/j.jclinepi.2013.09.003
- Turab, A., Pell, L., Bassani, D., Sajid, S., Shabina, A., Bhutta, Z., & Morris, S. (2014). The community-based delivery of an innovative neonatal kit to save newborn lives in rural Pakistan: design of a cluster randomized trial. *BMC Pregnancy & Childbirth*, 14(315).
- Turner, L., Shamseer, L., Altman Douglas, G., Weeks, L., Peters, J., Kober, T., . . . Moher, D. (2012). Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database of Systematic Reviews*, (11).
<http://onlinelibrary.wiley.com/doi/10.1002/14651858.MR000030.pub2/abstract>
 doi:10.1002/14651858.MR000030.pub2
- Tylleskar, T., Jackson, D., Meda, N., Engebretsen, I. M., Chopra, M., Diallo, A. H., . . . Group, P.-E. S. (2011). Exclusive breastfeeding promotion by peer

- counsellors in sub-Saharan Africa (PROMISE-EBF): a cluster-randomised trial. *Lancet*, 378(9789), 420-427. doi: 10.1016/S0140-6736(11)60738-1
- UNICEF/UNDP/World Bank/WHO TDR. (2000). Community-directed treatment of lymphatic filariasis in Africa: report of a multi-centre study in Ghana and Kenya (pp. 41-41).
- UNICEF/UNDP/World Bank/WHO TDR. (2008). Community-directed interventions for major health problems in Africa: a multi-country study: final report. Geneva: TDR.
- United Nations. (2015). The Millennium Development Goals Report 2015. Retrieved 12 June, 2016, from [http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20\(July%201\).pdf](http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf)
- van Ginneken, N., Tharyan, P., Lewin, S., Rao, G. N., Meera, S. M., Pian, J., . . . Patel, V. (2013). Non-specialist health worker interventions for the care of mental, neurological and substance-abuse disorders in low- and middle-income countries. *Cochrane Database of Systematic Reviews*, 11, CD009149.
- Vazir, S., Engle, P., Balakrishna, N., Griffiths, P. L., Johnson, S. L., Creed-Kanashiro, H., . . . Bentley, M. E. (2013). Cluster-randomized trial on complementary and responsive feeding education to caregivers found improved dietary intake, growth and development among rural Indian toddlers. *Matern Child Nutr*, 9(1), 99-117. doi: 10.1111/j.1740-8709.2012.00413.x.
- Vedanthan, R., Kamano, J. H., Naanyu, V., Delong, A. K., Were, M. C., Finkelstein, E. A., . . . Fuster, V. (2014). Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study): Study protocol for a randomized controlled trial. *Trials*, 15 (1) (143).
- Velema, J. P., Ebenso, B., & Fuzikawa, P. L. (2008). Evidence for the effectiveness of rehabilitation-in-the-community programmes. *Leprosy Review*, 79(1), 65-82.
- Vesel, L., ten Asbroek, A. H., Manu, A., Soremekun, S., Tawiah Agyemang, C., Okyere, E., . . . Kirkwood, B. R. (2013). Promoting skin-to-skin care for low birthweight babies: findings from the Ghana Newhints cluster-randomised trial. *Tropical Medicine & International Health*, 18(8), 952-961.
- Victora, C. G., Barreto, M. L., do Carmo Leal, M., Monteiro, C. A., Schmidt, M. I., Paim, J., . . . Travassos, C. (2011). Health conditions and health-policy innovations in Brazil: the way forward. *Lancet*, 377(9782), 2042-2053.
- Victora, C. G., Habicht, J.-P., & Bryce, J. (2004). Evidence-based public health: moving beyond randomized trials. *American Journal of Public Health*, 94(3), 400-405.
- Villar, J., Piaggio, G., Carroli, G., & Donner, A. (1997). Factors affecting the comparability of meta-analyses and largest trials results in perinatology. *Journal of Clinical Epidemiology*, 50(9), 997-1002.
- Volmink, J., & Garner, P. (2007). Directly observed therapy for treating tuberculosis. *Cochrane Database of Systematic Reviews*(4), CD003343.
- Vouking, M. Z., Tadenfok, C. N., Evina, C. D., Nsangou, M., Bonono, C. R., & Yondo, D. (2014). Revue systématique sur la contribution des femmes distributeurs communautaires dans la lutte contre l'onchocercose en Afrique sub-saharienne [Systematic review on the contribution of community women distributors in the fight against onchocerciasis in sub-Saharan Africa]. [Revue systematique sur la contribution des femmes distributeurs communautaires dans la lutte contre l'onchocercose en Afrique sub-saharienne.]. *Pan African Medical Journal*, 19, 117-123.

- Vouking, M. Z., Tamo, V. C., & Mbuagbaw, L. (2013). The impact of community health workers (CHWs) on Buruli ulcer in sub-Saharan Africa: a systematic review. *Pan African Medical Journal*, 15, 19.
- Wade, A., Osrin, D., Shrestha, B. P., Sen, A., Morrison, J., Tumbahangphe, K. M., . . . Costello, A. M. (2006). Behaviour change in perinatal care practices among rural women exposed to a women's group intervention in Nepal [ISRCTN31137309]. *BMC Pregnancy & Childbirth*, 6, 20. doi: 10.1186/1471-2393-6-20
- Waiswa, P., Pariyo, G., Kallander, K., Akuze, J., Namazzi, G., Ekirapa-Kiracho, E., . . . Uganda Newborn Study, T. (2015). Effect of the Uganda Newborn Study on care-seeking and care practices: a cluster-randomised controlled trial. *Global Health Action*, 8, 24584. doi: 10.3402/gha.v8.24584
- Walach, H., Falkenberg, T., Fonnebo, V., Lewith, G., & Jonas, W. B. (2006). Circular instead of hierarchical: methodological principles for the evaluation of complex interventions. *BMC Med Res Methodol*, 6, 29. doi: 10.1186/1471-2288-6-29
- Walker, D. G., & Jan, S. (2005). How do we determine whether community health workers are cost-effective? Some core methodological issues. *Journal of Community Health*, 30(3), 221-229.
- Walker, D. G., Wilson, R. F., Sharma, R., Bridges, J., Niessen, L., Bass, E. B., & Frick, K. (2012). Best Practices for Conducting Economic Evaluations in Health Care: A Systematic Review of Quality Assessment Tools *Methods Research Reports*: The Johns Hopkins University Evidence-based Practice Center.
- Walker, S. P., Chang, S. M., Powell, C. A., & Grantham-McGregor, S. M. (2004). Psychosocial intervention improves the development of term low-birth-weight infants. *J Nutr*, 134(6), 1417-1423.
- Walker, S. P., Chang, S. M., Powell, C. A., & Grantham-McGregor, S. M. (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *Lancet*, 366(9499), 1804-1807.
- Walker, S. P., Chang, S. M., Powell, C. A., Simonoff, E., & Grantham-McGregor, S. M. (2006). Effects of psychosocial stimulation and dietary supplementation in early childhood on psychosocial functioning in late adolescence: follow-up of randomised controlled trial. *British Medical Journal*, 333(7566), 472. doi: 10.1136/bmj.38897.555208.2F
- Walker, S. P., Chang, S. M., Vera-Hernandez, M., & Grantham-McGregor, S. (2011). Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics*, 127(5), 849-857. doi: 10.1542/peds.2010-2231
- Walker, S. P., Chang, S. M., Younger, N., & Grantham-McGregor, S. M. (2010). The effect of psychosocial stimulation on cognition and behaviour at 6 years in a cohort of term, low-birthweight Jamaican children. *Dev Med Child Neurol*, 52(7), e148-154. doi: 10.1111/j.1469-8749.2010.03637.x
- Walker, S. P., Grantham-McGregor, S. M., Himes, J. H., Powell, C. A., & Chang, S. M. (1996). Early childhood supplementation does not benefit the long-term growth of stunted children in Jamaica. *Community and International Nutrition*.
- Walker, S. P., Grantham-McGregor, S. M., Powell, C. A., & Chang, S. M. (2000). Effects of growth restriction in early childhood on growth, IQ, and cognition at age 11 to 12 years and the benefits of nutritional supplementation and

- psychosocial stimulation. *Journal of Pediatrics*, 137(1), 36-41. doi: 10.1067/mpd.2000.106227
- Walker, S. P., Powell, C. A., Grantham-McGregor, S. M., Himes, J. H., & Chang, S. M. (1991). Nutritional supplementation, psychosocial stimulation, and growth of stunted children: the Jamaican study. *American Journal of Clinical Nutrition*, 54(4), 642-648.
- Walt, G. (1988). CHWs: are national programmes in crisis? *Health Policy and Planning*, 3(1), 1-21.
- Walt, G. (1990). *Community health workers in national programmes: just another pair of hands?* Milton Keynes, UK: Open University Press.
- Walt, G. (2005). Commentary: Lay health workers in primary and community health care. *International Journal of Epidemiology*, 1251.
- Wang, L., Li, J., Chen, H., Li, F., Armstrong, G. L., Nelson, C., . . . Shapiro, C. N. (2007). Hepatitis B vaccination of newborn infants in rural China: evaluation of a village-based, out-of-cold-chain delivery strategy. *Bulletin of the World Health Organization*, 85(9), 688-694. doi: 10.2471/blt.06.037002
- Wang, X., Lindsley, K., & Li, T. (2012). Is there agreement in outcomes among Cochrane reviews to support 'Overviews' of reviews? A case study within the Cochrane Eyes and Vision Group (CEVG). The Cochrane Library. 2012 Supplement.
- Waters, E., Hall, B., Armstrong, R., Doyle, J., Pettman, T., & de Silva-Sanigorski, A. (2011). Essential components of public health evidence reviews: capturing intervention complexity, implementation, economics and equity. *Journal of Public Health*, 33(3), 462-465.
- Webb, S., & Webb, B. P. (1932). *Methods of social study*. Cambridge, UK: Cambridge University Press.
- Webb, S. A. (2001). Some considerations on the validity of evidence-based practice in social work. *British Journal of Social Work*, 57-79.
- Welch, V., Petticrew, M., Tugwell, P., Moher, D., O'Neill, J., Waters, E., & White, H. (2013). PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *Revista Panamericana de Salud Publica*, 34(1), 60-67.
- Werner, D. (1981). The Village Health Worker: Lackey Or Liberator? *World Health Forum*, 2(1), 46-68.
- Werner, D., & Bower, B. (2005). *Helping health workers learn: a book of methods, aids, and ideas for instructors at the village level*: Hesperian Foundation.
- Whitlock, E. P., Lin, J. S., Chou, R., Shekelle, P., & Robinson, K. A. (2008). Using Existing Systematic Reviews in Complex Systematic Reviews. *Annals of Internal Medicine*, 148(10), 776-782. doi: 10.7326/0003-4819-148-10-200805200-00010
- WHO Study Group. (1989). Strengthening the performance of community health workers in primary health care: report of a WHO Study Group. Geneva, Switzerland: World Health Organization.
- Wiley-Exley, E. (2007). Evaluations of community mental health care in low- and middle-income countries: a 10-year review of the literature. *Social Science & Medicine*, 64(6), 1231-1241. doi: 10.1016/j.socscimed.2006.11.009
- Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., Gargon, E., & Tugwell, P. (2012). Developing core outcome sets for clinical trials: issues to consider. *Trials*, 13, 132. doi: 10.1186/1745-6215-13-132

- Willis, J. R., Kumar, V., Mohanty, S., Singh, V., Kumar, A., Singh, J. V., . . . Saksham Study, G. (2012). Impact of community-based behaviour-change management on perceived neonatal morbidity: a cluster-randomized controlled trial in Shivgarh, Uttar Pradesh, India. *J Trop Pediatr*, 58(4), 286-291. doi: 10.1093/tropej/fmr097
- Willis, N., Hill, S., Kaufman, J., Lewin, S., Kis-Rigo, J., Freire, S. B. D. C., . . . Wiysonge, C. S. (2013). "Communicate to vaccinate": the development of a taxonomy of communication interventions to improve routine childhood vaccination. *Bmc International Health and Human Rights*, 13(1).
- Winch, P., Bagayoko, A., Diawara, A., Kane, M., Thiero, F., Gilroy, K., . . . Swedberg, E. (2003). Increases in correct administration of chloroquine in the home and referral of sick children to health facilities through a community-based intervention in Bougouni District, Mali. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 97(5), 481-490.
- Witmer, A., Seifer, S. D., Finocchio, L., Leslie, J., & Oneil, E. H. (1995). Community Health Workers: Integral Members Of The Health-Care Work-Force. *American Journal of Public Health*, 85(8), 1055-1058. doi: 10.2105/AJPH.85.8_Pt_1.1055
- Woelk, G. B. (1994). Primary health care in Zimbabwe: Can it survive?: An exploration of the political and historical developments affecting the implementation of PHC. *Social Science & Medicine*, 39(8), 1027-1035.
- Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: realist syntheses. *BMC Medicine*, 11(1), 1.
- Woodd, S. L., Grosskurth, H., Levin, J., Amuron, B., Namara, G., Birunghi, J., . . . Jaffar, S. (2014). Home-based versus clinic-based care for patients starting antiretroviral therapy with low CD4(+) cell counts: findings from a cluster-randomized trial. *AIDS*, 28(4), 569-576. doi: 10.1097/QAD.0000000000000056
- World Bank. (2014). Country and Lending Groups. Retrieved May 1, 2014, from <http://data.worldbank.org/about/country-and-lending-groups>
- World Health Organization. (1974). Training and Utilization of Feldshers in the USSR. Geneva, Switzerland.
- World Health Organization. (1978). *Declaration of Alma-Ata*. Paper presented at the International Conference on Primary Health Care, Alma-Ata, USSR.
- World Health Organization. (1987). Community health workers: pillars for health for all. Report of the Interregional conference, Yaoundé, Cameroon, 1-5 December 1986. Unpublished document SHS/CIH/87.2.
- World Health Organization. (2006). The World Health Report: 2006: Working Together For Health. Geneva, Switzerland.
- World Health Organization. (2007). Treat, Train, Retain: Global recommendations and guidelines. Geneva, Switzerland.
- World Health Organization. (2014). *The 2014 update, Global Health Workforce Statistics*. Retrieved from: <http://www.who.int/hrh/statistics/hwfstats/>
- World Health Organization, PEPFAR, & UNAIDS. (2008). Task Shifting: Rational Redistribution Of Tasks Among Health Workforce Teams. Global Recommendations And Guidelines. Geneva, Switzerland: World Health Organization.
- Wouters, E., Van Damme, W., van Rensburg, D., Masquillier, C., & Meulemans, H. (2012). Impact of community-based support services on antiretroviral treatment programme delivery and outcomes in resource-limited countries: a

- synthetic review. *BMC Health Services Research*, 12, 194. doi: 10.1186/1472-6963-12-194
- Wright, J., Walley, J., Philip, A., Pushpanathan, S., Dlamini, E., Newell, J., & Dlamini, S. (2004). Direct observation of treatment for tuberculosis: a randomized controlled trial of community health workers versus family members. *Tropical Medicine & International Health*, 9(5), 559-565.
- Yeboah-Antwi, K., Pilingana, P., Macleod, W. B., Semrau, K., Siazeele, K., Kalesha, P., . . . Hamer, D. H. (2010). Community case management of fever due to malaria and pneumonia in children under five in Zambia: a cluster randomized controlled trial. *PLoS Medicine / Public Library of Science*, 7(9), e1000340.
- Yeomans, P. D., Forman, E. M., Herbert, J. D., & Yuen, E. (2010). A randomized trial of a reconciliation workshop with and without PTSD psychoeducation in Burundian sample. *J Trauma Stress*, 23(3), 305-312. doi: 10.1002/jts.20531
- Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2014). Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *Lancet*, 384(9950), 1282-1293.
- Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2015). Parenting Skills and Emotional Availability: An RCT. *Pediatrics*, 135(5), e1247-1257. doi: 10.1542/peds.2014-2335
- Yuan, J. Q., Liu, Y. L., Yang, K. H., Chen, Y. L., & Yang, M. (2012). Survey of quality assessment methods of included studies in overviews of reviews. *Chinese Journal of Evidence-Based Medicine*, 12(2), 238-242.
- Yusuf, S., Wittes, J., Probstfield, J., & Tyroler, H. A. (1991). ANalysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *Journal of the American Medical Association*, 266(1), 93-98. doi: 10.1001/jama.1991.03470010097038
- Zachariah, R., Harries, A. D., Ishikawa, N., Rieder, H. L., Bissell, K., Laserson, K., . . . Reid, T. (2009). Operational research in low-income countries: what, why, and how? *The Lancet Infectious Diseases*, 9(11), 711-717.
- Zafar, S., Sikander, S., Hamdani, S. U., Atif, N., Akhtar, P., Nazir, H., . . . Rahman, A. (2016). The effectiveness of Technology-assisted Cascade Training and Supervision of community health workers in delivering the Thinking Healthy Program for perinatal depression in a post-conflict area of Pakistan - study protocol for a randomized controlled trial. *Trials*, 17 (1)(188).
- Zhu, N., Ling, Z., Shen, J., Lane, J., & Hu, S. (1989). Factors associated with the decline of the Cooperative Medical System and barefoot doctors in rural China. *Bulletin of the World Health Organization*, 67(4), 431.
- Zurovac, D., Sudoi, R. K., Akhwale, W. S., Ndiritu, M., Hamer, D. H., Rowe, A. K., & Snow, R. W. (2011). The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. *Lancet*, 378(9793), 795-803.
- Zwarenstein, M., Schoeman, J. H., Vundule, C., Lombard, C. J., & Tatley, M. (2000). A randomised controlled trial of lay health workers as direct observers for treatment of tuberculosis. *The International Journal of Tuberculosis and Lung Disease*, 4(6), 550-554.

Appendices to
Community Health Workers: Efficacy,
Taxonomy, and Performance

Madeleine Ballard

Trinity College, University of Oxford

Trinity Term, 2016

Thesis submitted in partial fulfilment of the requirements for
the degree of DPhil in Social Intervention in the Social
Sciences Division at the University of Oxford

Table of Contents

Appendix to Chapter 1	432
Appendix 1a: Alternative Names for Community Health Worker	432
Appendix to Chapter 2	434
Appendix 2a: Excluded studies	434
Appendices to Chapter 3	435
Appendix 3a: Low And Middle-Income Countries	435
Appendix 3b: Search Strategies	435
Appendix 3c: Proposed Data Extraction Form	441
Appendix 3d: Calculating the CCA	443
Appendix 3e: Application of AMSTAR	449
Appendices to Chapter 4	453
Appendix 4a: Search Strategies	453
Appendix 4b: Data Extraction Form	456
Appendix 4c: Characteristics of Studies	458
Characteristics of included studies (includes 'Risk of bias' tables)	458
Characteristics of excluded studies	579
Characteristics of ongoing studies	585
List of trials awaiting classification	602
Funnel plots	603
Quality of included economic evaluations	603
Appendix 4d: Data And Analyses	604
Comparison 1: CHW interventions for treatment of common mental disorders	605
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five	606
Comparison 3: CHW interventions to improve maternal and newborn health	606
Comparison 4: CHW interventions to promote cognitive development in infants	608
Comparison 5: CHW interventions to promote healthy feeding practices	609
Comparison 6: CHW interventions to improve TB treatment outcomes compared with other forms of adherence support	609
Comparison 7: CHW interventions for maintenance of antiretroviral therapy	610
Comparison 8: CHW interventions to promote injury prevention in children	610
Comparison 9: CHW interventions to promote immunization uptake and mass drug administration	611
Outcomes of trials not assigned to the above comparisons	629
Appendix 4e: PRISMA Checklist	632
Appendices to Chapter 5	634
Appendix 5a: Characteristics of Studies	634
Appendix 5b: Taxonomy	639
Appendix 5c: Citations for illustrative examples	639
Appendices to Chapter 6	644
Appendix 6a: Search Strategy	644
Appendix 6b: Data Extraction Form	652
Appendix 6c: PRISMA Checklist	656
Appendix 6d: Characteristics of studies	659
Characteristics of included studies (includes 'Risk of bias' tables)	659
Characteristics of excluded studies	680
Characteristics of studies awaiting classification	681

Characteristics of ongoing studies	681
Quality of included economic evaluations	685
Appendix 6e: Differences Between Protocol and Review	686

Appendix to Chapter 1

Appendix 1a: Alternative Names for Community Health Worker

1	Accompagnateurs
2	Activista
3	Agente comunitario de salud
4	Agente comunitário de saúde
5	Allied health personnel
6	Anganwadi
7	Animatrice
8	Auxiliary health worker
9	Barangay health worker
10	Barefoot doctor
11	Basic health worker
12	Behvarz
13	Brigadista
14	Colaborador voluntario
15	Community assistants
16	Community drug distributor
17	Community health representative
18	Community Health Advocates
19	Community health agent
20	Community Health Aides
21	Community health promoter
22	Community health volunteer
23	Community health worker
24	Community mobilizer
25	Community nutrition worker
26	Community resource person
27	Community Reproductive Health Worker
28	Community support worker
29	Community volunteers
30	Community-based Workers
31	Female community health volunteer
32	Female multipurpose health worker
33	Health and nutrition worker
34	Health promoter
35	Home health aides
36	Kader
37	Lady health worker
38	Lay Health Visitor
39	Lay Health Worker
40	Link worker

41	Maternal & Child Health Promotion Workers
42	Maternal and child health worker
43	Mental Health Workers
44	Monitora
45	Mother coordinator
46	Nutrition Volunteers
47	Nutrition Worker
48	Outreach educator
49	Paramedical worker
50	Peer volunteer
51	Postnatal Support Worker
52	Promotora
53	Raodat
54	Rural health motivator
55	Rural Health Worker
56	Saksham Sahaya
57	Sevika
58	Shastho karmis
59	Shastho shebika
60	Village drug-kit manager
61	Village Health Guide
62	Village health helper
63	Village Health Promoters
64	Village health worker
65	Village Malaria Worker
66	Visitadora
67	Voluntary Malaria Workers
68	Voluntary workers
69	Women Group Leaders

Appendix to Chapter 2

Appendix 2a: Excluded studies

Reference	Reason for Exclusion
Becker 2006	Primary focus is not to describe overview characteristics or provide methodological guidance
Delgado-Rodriguez 2006	Primary focus is not to describe overview characteristics or provide methodological guidance
Pieper 2014	Primary focus is not to describe overview characteristics or provide methodological guidance
Robinson 2014	Overviews are not the subject of the paper (whether and how to incorporate existing systematic reviews into new systematic reviews)
Sarami-Foroushani 2015	Overviews are not the subject of the paper (scoping meta-review)
Thomson 2014	Primary focus is not to describe overview characteristics or provide methodological guidance
Whitlock 2008	Overviews are not the subject of the paper (whether and how to incorporate existing systematic reviews into new systematic reviews)

Appendices to Chapter 3

Appendix 3a: Low And Middle-Income Countries

Based on World Bank Country and Lending Groups:

Upper-middle-income economies

Albania, Algeria, American Samoa, Angola, Azerbaijan, Belarus, Belize, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Fiji, Gabon, Grenada, Iran, Islamic Rep., Iraq, Jamaica, Jordan, Kazakhstan, Lebanon, Libya, Macedonia, Malaysia, Maldives, Marshall Islands, Mauritius, Mexico, Mongolia, Montenegro, Namibia, Palau, Panama, Paraguay, Peru, Romania, Serbia, South Africa, St. Lucia, St. Vincent and the Grenadines, Suriname, Thailand, Tonga, Tunisia, Turkey, Turkmenistan, Tuvalu.

Lower-middle-income economies

Armenia, Bangladesh, Bhutan, Bolivia, Cabo Verde, Cameroon, Congo, Rep., Côte d'Ivoire, Djibouti, Egypt, El Salvador, Georgia, Ghana, Guatemala, Guyana, Honduras, India, Indonesia, Kenya, Kiribati, Kosovo, Kyrgyz Republic, Lao PDR, Lesotho, Mauritania, Micronesia, Moldova, Morocco, Myanmar, Nicaragua, Nigeria, Pakistan, Papua New Guinea, Philippines, Samoa, São Tomé and Príncipe, Senegal, Solomon Islands, Sri Lanka, Sudan, Swaziland, Syrian Arab Republic, Tajikistan, Timor-Leste, Ukraine, Uzbekistan, Vanuatu, Vietnam, West Bank and Gaza, Yemen, Zambia

Low-income economies

Afghanistan, Benin, Burkina Faso, Burundi, Cambodia, Central African Republic, Chad, Comoros, Dem. Rep Congo, Eritrea, Ethiopia, The Gambia, Guinea, Guinea-Bissau, Haiti, Dem Rep. Korea, Liberia, Madagascar, Malawi, Mali, Mozambique, Nepal, Niger, Rwanda, Sierra Leone, Somalia, South Sudan, Tanzania, Togo, Uganda, Zimbabwe

Appendix 3b: Search Strategies

3ie Systematic Reviews and Policy Briefs

“community health”

In: systematic reviews

The search was conducted on February 24, 2015 and produced 13 hits.

BiblioMap: EPPI-Centre database of health promotion research

1. Community Health Agent*
2. Community Health Aide*
3. Community health promoter*
4. Community mobili*er*

5. Community drug distributor*
6. Community health worker*
7. Village health worker*
8. Rural Health Worker*
9. Lay Health Worker*
10. Lady health worker*
11. nutrition worker*
12. frontline health worker*
13. Barangay health worker*
14. basic health worker*
15. Auxiliary health worker*
16. health extension worker*
17. community health volunteer*
18. village health volunteer*
19. accompanier*
20. accompagnateur*
21. activista*
22. animatrice*
23. brigadista*
24. kader*
25. promotora*
26. monitora*
27. sevika*
28. fhw*
29. lhw*
30. chw*
31. vhw*
32. chv*
33. shastho shebika*
34. shasto karmis
35. anganwadi*
36. barefoot doctor*
37. agente comunitario de salud*
38. agente comunitario de saude
39. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR
22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31
OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38

The search was conducted on February 25, 2015 and produced 14 hits.

Campbell Library of Systematic Reviews

“community health”

The search was conducted on February 24, 2015 and produced 0 hits.

CINAHL, Ebsco

1. TI ((community OR village) N2 (agent* OR aid* OR promot* OR mobili?er* OR distribut*r* OR worker))
2. TI ((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) N2 (worker* OR volunteer*))
3. TI ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR monitora* OR promotora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente comunitario de saude"))
4. S1 OR S2 OR S3

The search was conducted on February 24, 2015 and produced 38 hits.

Cochrane Database of Systematic Reviews, Wiley

1. ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)):ti,ab
3. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente comunitario de saude"):ti,ab
4. #1 or #2 or #3
(In: Cochrane reviews, other reviews, technology assessments, and economic evaluations).

The search was conducted on February 24, 2015 and produced 66 hits.

Database of Abstracts of Reviews of Effectiveness (DARE)

((community or village) near2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab) OR (((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near2 (worker* or volunteer*)):ti,ab) OR (((accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or shastho shebika or shasto karmis or anganwadi* or barefoot doctor or agente comunitario de salud or agente

comunitario de saude)):ti,ab)

The search was conducted on February 24, 2015 and produced 26 hits.

Database of Promoting Health Effectiveness Reviews (DoPHER)

1. Community Health Agent*
2. Community Health Aide*
3. Community health promoter*
4. Community mobili*er*
5. Community drug distributor*
6. Community health worker*
7. Village health worker*
8. Rural Health Worker*
9. Lay Health Worker*
10. Lady health worker*
11. nutrition worker*
12. frontline health worker*
13. Barangay health worker*
14. basic health worker*
15. Auxiliary health worker*
16. health extension worker*
17. community health volunteer*
18. village health volunteer*
19. accompanier*
20. accompagnateur*
21. activista*
22. animatrice*
23. brigadista*
24. kader*
25. promotora*
26. monitora*
27. sevika*
28. fhw*
29. lhw*
30. chw*
31. vhw*
32. chv*
33. shastho shebika*
34. shasto karmis
35. anganwadi*
36. barefoot doctor*
37. agente comunitario de salud*
38. agente comunitario de saude
39. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR
22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31
OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38

The search was conducted on February 25, 2015 and produced 19 hits.

Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index - Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities)

1. TI=((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*))
2. TI=((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) NEAR/2 (worker* OR volunteer*))
3. TI=(accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente comunitario de saude")
4. #3 OR #2 OR #1
5. TI=(meta-analysis or review or search or systematic)
6. #5 AND #4

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, CCR-EXPANDED, IC Timespan=All years

The search was conducted on February 24, 2015 and produced 66 hits.

Embase, Ovid (1974-2015 WK 8)

Searched using the terms in the main body of the text. The search was conducted on April 15, 2015 and produced 1133 hits.

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) Evidence Library

The reviewer scanned a complete list of all EPPI-Centre reviews, and is in alphabetical order within each year found at:

<http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=62>. The search was conducted on February 24, 2015 and produced 1 hit.

Global Health, Ovid (1973 to 2015 Week 07)

Searched using the terms in the main body of the text. The search was conducted on February 24, 2015 and produced 289 hits.

Health Technology Assessment

((community or village) near2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab) OR (((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near2 (worker* or volunteer*)):ti,ab) OR (((accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or shastho shebika or shasto karmis or anganwadi* or barefoot doctor or agente comunitario de salud or agente comunitario de saude)):ti,ab)

The search was conducted on February 24, 2015 and produced 4 hits.

MEDLINE In-Process & Other Non-Indexed Citations, Ovid

Searched using the terms in the main body of the text. The search was conducted on February 24, 2015 and produced 837 hits.

MEDLINE, Ovid

Searched using the terms in the main body of the text. The search was conducted on February 24, 2015 and produced 761 hits.

NHS Economic Evaluation Database (EED)

((community or village) near2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab) OR (((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near2 (worker* or volunteer*)):ti,ab) OR (((accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or shastho shebika or shasto karmis or anganwadi* or barefoot doctor or agente comunitario de salud or agente comunitario de saude)):ti,ab)

The search was conducted on February 24, 2015 and produced 16 hits.

POPLINE

All fields: review or metaanalysis or systematic or search

Keywords: Community workers or community-based distribution workers or community-based distribution or community health services

The search was conducted on February 25, 2015 and produced 88 hits.

Trip Database

“community health workers” or “lay health workers” in systematic reviews

The search was conducted on February 24, 2015 and produced 41 hits.

Appendix 3c: Proposed Data Extraction Form

Review ID (First author, year)

Variable	Value	Type	Options
INCLUSION CRITERIA			
Clearly stated set of objectives with pre-defined eligibility criteria for studies		Dichotomous	• Yes • No
Explicit, reproducible methodology		Dichotomous	• Yes • No
Systematic search		Dichotomous	• Yes • No
Systematic presentation, and synthesis, of the characteristics and findings of the included studies.		Dichotomous	• Yes • No
CHWs are primarily based in the community (as opposed to a health facility)		Dichotomous	• Yes • No
CHWs belong to the formal health system (i.e. are managed by the government or an implementing NGO) in a low-income country (as defined by World Bank, 2014).		Dichotomous	• Yes • No
CHWs perform interventions related to healthcare delivery.		Dichotomous	• Yes • No
CHWs have received organized training but have no formal or paraprofessional certification or tertiary education degree.		Dichotomous	• Yes • No

CHARACTERISTICS		
Stated aim of the review		Narrative
Last search/data as assessed up-to-date		Date Day/Month/Year
>2 years since last search (to Feb 2014)		Dichotomous • Yes • No
Number and type of included trials		Continuous
Population		Continuous
Intervention / comparison		Narrative
Outcomes		Narrative
O1: Assumed risk with Comparator		
O1: Corresponding risk with intervention		
O1: Relative effect (95%CI)		
O1: Number of participants (Studies)		
O1: Quality of the evidence (GRADE)		
O1: Comments		
O2: Assumed risk with Comparator		
O2: Corresponding risk with intervention		
O2: Relative effect (95%CI)		
O2: Number of participants (Studies)		
O2: Quality of the evidence (GRADE)		
O2: Comments		
REVIEW QUALITY (AMSTAR)		
Prespecified question and inclusion criteria		Nominal • Yes • No • Can't answer • Not applicable
Duplicate study selection and data extraction		Nominal • Yes • No • Can't answer • Not applicable
Comprehensive lit search		Nominal • Yes • No • Can't answer • Not applicable
Grey lit included		Nominal • Yes • No • Can't answer • Not applicable

Lists included and excluded studies		Nominal	• Yes • No • Can't answer • Not applicable
Describes characteristics of included studies		Nominal	• Yes • No • Can't answer • Not applicable
Study quality assessed		Nominal	• Yes • No • Can't answer • Not applicable
Study quality used appropriately in formulating conclusions		Nominal	
Likelihood of publication bias considered/tested		Nominal	• Yes • No • Can't answer • Not applicable
Studies combined using appropriate methods		Nominal	• Yes • No • Can't answer • Not applicable
Potential for conflict of interest addressed		Nominal	• Yes • No • Can't answer • Not applicable

Appendix 3d: Calculating the CCA

$CCA = (N-r)/(rc-r)$	
N= total publications (includes double counting)	
r = index publications	
c = # of reviews	

Calculation of Corrected Coverage Area:

	N	r	c	CCA
Child health	17	7	10	15.9
HIV	20	4	7	66.7
Mental, Neurological and Substance Abuse	26	16	10	6.9
MNH	88	30	18	11.4
NCD	3	2	2	50.0
Nutrition	4	3	3	16.7
TB	28	9	6	42.2

Citation Matrix on which the Calculation of the Corrected Coverage Area was based:

Primary Study ID	Systematic Review ID	Category
Chinbuah 2012	Amouzou 2014	Child Health
Chinbuah 2012	Pantain 2014	Child Health
Kalyango 2012	Pantain 2014	Child Health
Kalyango 2012	Druetz 2013	Child Health
Kidane 2000	Bhutta 2010	Child Health
Kidane 2000	Bigirwa 2012	Child Health
Kidane 2000	Hopkins 2007	Child Health
Kidane 2000	Lewin 2010	Child Health
Kidane 2000	Okwundu 2013	Child Health
Kidane 2000	Pantain 2014	Child Health
Lemma 2009	Kamal-Yanni 2012	Child Health
Miller	Amouzou 2014	Child Health
Mtango 1986	Druetz 2013	Child Health
Mtango 1986	Lewin 2010	Child Health
Soofi 2012	Das 2009	Child Health
Chang 2010	Lefevre 2015	HIV
Chang 2010	Mdege 2013	HIV
Chang 2010	Mwai 2013	HIV
Chang 2010	Wouters 2012	HIV
Jaffar 2009	Callaghan 2010	HIV
Jaffar 2009	Decroo 2013	HIV
Jaffar 2009	Kredo 2014	HIV
Jaffar 2009	Lefevre 2015	HIV
Jaffar 2009	Mdege 2013	HIV
Jaffar 2009	Mwai 2013	HIV
Jaffar 2009	Wouters 2012	HIV
Jaffar 2009	Decroo 2013	HIV
Jaffar 2009	Decroo 2013	HIV
Nachegea 2010	Wouters 2012	HIV

Selke 2010	Decroo 2013	HIV
Selke 2010	Lefevvre 2015	HIV
Selke 2010	Mdege 2013	HIV
Selke 2010	Mwai 2013	HIV
Selke 2010	Wouters 2012	HIV
Selke 2010	Decroo 2013	HIV
Ali 2003	Mutamba 2013	Mental Health
Ali 2003	Van Ginneken 2013	Mental Health
Ali 2003	Wiley-Exley 2007	Mental Health
Baker-Henningham 2005	Mutamba 2013	Mental Health
Baker-Henningham 2005	Van Ginneken 2013	Mental Health
Bolton 2003	Van Ginneken 2013	Mental Health
Bolton 2007	Van Ginneken 2013	Mental Health
Cooper 2009	Mutamba 2013	Mental Health
Dias 2008	Mutamba 2013	Mental Health
Dias 2008	Van Ginneken 2013	Mental Health
Ertl 2011 2011	Van Ginneken 2013	Mental Health
Gardner 2003	Lewin 2010	Mental Health
Gardner 2003	Mutamba 2013	Mental Health
Grantham McGregor 1991	Mutamba 2013	Mental Health
Hamadani 2006	Mutamba 2013	Mental Health
Hughes 2009	Rahman 2013	Mental Health
Neuner 2008	Mutamba 2013	Mental Health
Powell 1989	Mutamba 2013	Mental Health
Rahman 2008	Bee 2014	Mental Health
Rahman 2008	Bhutta 2010	Mental Health
Rahman 2008	Bigirwa 2014	Mental Health
Rahman 2008	Gilmore 2013	Mental Health
Rahman 2008	Joshi 2014	Mental Health
Rahman 2008	Van Ginneken 2013	Mental Health
Rahman 2009	Rahman 2013	Mental Health
Walker 2004	Mutamba 2013	Mental Health

Agrasada 2005	Bhutta 2010	MNCH
Agrasada 2005	Gilmore 2013	MNCH
Agrasada 2005	Lewin 2010	MNCH
Ali 2005	Byrne 2014	MNCH
Andersson 2009	Glenton 2011	MNCH
Azad 2010	Bhutta 2010	MNCH
Azad 2010	Lassi 2010	MNCH
Baqui 2008	Bhutta 2010	MNCH
Baqui 2008	Bigirwa 2009	MNCH
Baqui 2008	Darmstadt 2009	MNCH
Baqui 2008	Gogia 2010	MNCH
Baqui 2008	Gogia 2011	MNCH
Baqui 2008	Lee 2009	MNCH
Baqui 2008	Lewin 2010	MNCH
Baqui 2008	Schiffman 2010	MNCH
Baqui 2008	Bhutta 2008	MNCH
Baqui com care (a) 2008	Lassi 2010	MNCH
Baqui home care (a) 2008	Lassi 2010	MNCH
Bari 2006	Bhutta 2010	MNCH
Bari 2006	Bigirwa 2010	MNCH
Bari 2006	Gilmore 2013	MNCH
Bari 2006	Lassi 2010	MNCH
Bari 2006	Lewin 2010	MNCH
Borghi 2005	Mangham-Jeffries 2014	MNCH
Colbourn 2013 (a&b)	Lassi 2015	MNCH
Colbourn 2013 (a&b)	Mangham-Jeffries 2014	MNCH
Cooper 2009	Gilmore 2013	MNCH
Cooper 2009	Rahman 2013	MNCH
Coutinho 2005	Lewin 2010	MNCH
Darmstadt 2010	Bhutta 2010	MNCH
Darmstadt 2010	Gogia 2011	MNCH
Darmstadt 2010	Lassi 2015	MNCH

Darmstadt 2010	Schiffman 2012	MNCH
Ferreira 2010	Arrowsmith 2012	MNCH
Fottrell 2013	Lassi 2015	MNCH
Fottrell 2013	Mangham-Jeffries 2014	MNCH
Gökçay 1993	Bhutta 2010	MNCH
Gökçay 1993	Glenton 2011	MNCH
Gökçay 1993	Lewin 2010	MNCH
Haider 2000	Bhutta 2010	MNCH
Haider 2000	Bigirwa 2011	MNCH
Haider 2000	Gilmore 2013	MNCH
Haider 2000	Hall 2011	MNCH
Haider 2000	Lewin 2010	MNCH
Kirkwood 2013	Lassi 2015	MNCH
Kumar 2008	Bhutta 2008	MNCH
Kumar 2008	Bigirwa 2015	MNCH
Kumar 2008	Darmstadt 2009	MNCH
Kumar 2008	Gilmore 2013	MNCH
Kumar 2008	Gogia 2010	MNCH
Kumar 2008	Lee 2009	MNCH
Kumar 2008	Lewin 2010	MNCH
Kumar 2008	Schiffman 2013	MNCH
Kumar 2008	Gogia 2011	MNCH
Kumar 2008/ Darmstadt 2006	Gilmore 2013	MNCH
Kumar 2008/ Darmstadt 2006	Lassi 2010	MNCH
LeFevre 2013	Mangham-Jeffries 2014	MNCH
Leite 2005	Bhutta 2010	MNCH
Leite 2005	Bigirwa 2013	MNCH
Leite 2005	Lewin 2010	MNCH
Lewycka 2013	Lassi 2015	MNCH
Lewycka 2013	Mangham-Jeffries 2014	MNCH
Luby 2006	Lewin 2010	MNCH

Manandhar 2004	Bhutta 2008	MNCH
Manandhar 2004	Gogia 2011	MNCH
Manandhar 2004	Lee 2009	MNCH
Manandhar 2004	Lewin 2010	MNCH
Manandhar 2004	Schiffman 2015	MNCH
Manandhar 2004/ Osrin 2003/ Wade 2006	Lassi 2010	MNCH
More 2012	Lassi 2015	MNCH
Morrow 1999	Bhutta 2010	MNCH
Morrow 1999	Gilmore 2013	MNCH
Morrow 1999	Lewin 2010	MNCH
Pence 2005	Bhutta 2010	MNCH
Pence 2005	Christopher 2011	MNCH
Pence 2005	Lewin 2010	MNCH
Sloan 2008	Bhutta 2010	MNCH
Sloan 2008	Gilmore 2013	MNCH
Sloan 2008	Lewin 2010	MNCH
Swart 2008	Lewin 2010	MNCH
Tripathy 2010	Bhutta 2010	MNCH
Tripathy 2010	Mutamba 2013	MNCH
Tripathy 2010	Mangham-Jeffries 2014	MNCH
Tripathy 2010	Rahman 2013	MNCH
Tripathy 2010	Gogia 2011	MNCH
Tripathy 2010	Lassi 2011	MNCH
Tylleskar 2011	Gilmore 2013	MNCH
Wang 2007	Lewin 2010	MNCH
Jafar 2009	Bhutta 2010	NCDs
Sankaranarayanan 2005	Bhutta 2010	NCDs
Sankaranarayanan 2005	Joshi 2014	NCDs
Pachon 2002	Sunguya 2013	Nutrition
Sripaipan 2002	Bhutta 2010	Nutrition
Sripaipan 2002	Lewin 2010	Nutrition
Vazir 2012	Sunguya 2013	Nutrition

CDI study group	Ahmed 2014	TB
Chongsuvivatwong 1996	Lewin 2010	TB
Clarke 2005	Bhutta 2010	TB
Clarke 2005	Lewin 2010	TB
Clarke 2005	Musa 2014	TB
Clarke 2005	Ahmed 2014	TB
Clarke 2005	Arshad 2017	TB
Datiko 2009	Bhutta 2010	TB
Datiko 2009	Musa 2014	TB
Lwilla 2003	Bhutta 2010	TB
Lwilla 2003	Lewin 2010	TB
Lwilla 2003	Musa 2014	TB
Lwilla 2003	Volmink 2007	TB
Lwilla 2003	Ahmed 2014	TB
Mohan 2003	Bhutta 2010	TB
Mohan 2003	Lewin 2010	TB
Newell 2006	Ahmed 2014	TB
Newell 2006	Bhutta 2010	TB
Newell 2006	Volmink 2007	TB
Wright 2004	Ahmed 2014	TB
Wright 2004	Arshad 2016	TB
Wright 2004	Bhutta 2010	TB
Wright 2004	Musa 2014	TB
Wright 2004	Volmink 2007	TB
Zwarenstein 2000	Ahmed 2014	TB
Zwarenstein 2000	Bhutta 2010	TB
Zwarenstein 2000	Lewin 2010	TB
Zwarenstein 2000	Volmink 2007	TB

Appendix 3e: Application of AMSTAR

Review ID	Prespecified question and inclusion criteria	Duplicate study selection and data extraction	Comprehensive lit search	Grey lit included	Lists included and excluded studies	Describes characteristics of included studies	Study quality assessed	Study quality used appropriately in formulating conclusions	Studies combined using appropriate methods	Likelihood of publication bias considered/tested	Potential for conflict of interest addressed	Total score
Ahmed 2014	Y	Y	Y	Y	N	Y	Y	N	Y	N	N	7
Amouzou 2014	N	N	Y	Y	N	Y	Y	Y	Y	N	N	6
Arrowsmith 2012	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10
Arshad 2014	N	Y	Y	N	N	Y	Y	N	N	N	N	4
Bee 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10
Bhutta 2008	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	8
Bhutta 2010	N	N	Y	Y	N	Y	N	N	N	N	N	3
Bigirwa 2009	Y	N	N	N	N	Y	N	N	Y	N	N	3
Bruxvoort 2014	N	N	N	N	N	Y	N	N	N	N	N	1
Byrne 2014	Y	Y	N	N	N	Y	Y	N	Y	N	N	5
Callaghan 2010	N	N	Y	Y	N	Y	Y	Y	Y	Y	N	7
Christopher 2011	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	8
Corluka 2009	N	Y	Y	N	N	Y	Y	Y	Y	N	Y	7
Darmstadt 2009	C	C	Y	Y	N	Y	N	N	N	Y	N	4
Das 2013	N	N	Y	Y	N	N	Y	Y	Y	N	N	5
Decroo 2009	N	Y	Y	Y	N	Y	N	N	N	N	N	4
Druetz 2013	N	N	Y	Y	N	Y	Y	N	Y	N	N	5

Gilmore 2013	Y	N	N	N	Y	Y	Y	Y	Y	N	N	6
Glenton 2011	N	Y	Y	Y	N	Y	Y	Y	Y	N	N	6
Glenton 2013	C	N	Y	Y	Y	Y	Y	Y	NA	N	N	6
Gogia 2010	Y	Y	Y	N	Y	Y	N	N	Y	Y	N	7
Gogia 2011	Y	C	Y	N	N	Y	Y	Y	Y	Y	N	7
Hall 2011	N	N	N	N	Y	Y	Y	Y	Y	N	N	5
Hopkins 2007	N	C	N	Y	Y	Y	N	N	Y	N	N	4
Hundley 2013	C	Y	Y	N	N	Y	Y	N	Y	N	N	5
Joshi 2014	Y	Y	Y	N	N	Y	Y	N	Y	N	N	6
Kamal-Yanni 2012	N	N	N	N	N	Y	N	N	N	N	N	1
Kredo 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10
Lassi 2010/2012 /2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10
Lee 2009	N	N	N	N	N	Y	Y	Y	Y	Y	N	5
Lee 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9
Lefevre 2015	N	C	N	Y	N	Y	N	N	Y	N	N	3
Lewin 2010	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	10
Malar-cher 2011	C	N	Y	Y	N	y	Y	Y	N	N	N	5
Mangham-Jefferies 2014	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	8
Mannan 2012	N	C	N	N	N	Y	N	N	Y	N	N	2
Mdege 2013	N	Y	Y	Y	N	Y	Y	Y	Y	N	N	7
Musa 2014	N	Y	N	N	N	Y	Y	Y	Y	N	N	5
Mutamba 2013	Y	N	N	N	N	Y	Y	Y	Y	N	N	5
Mwai 2013	N	N	N	N	N	Y	Y	N	Y	N	N	3

Okwundu 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	10
Oladapo 2012	Y	Y	Y	Y	NA	NA	NA	NA	NA	NA	NA	4
Pantain 2014	N	C	Y	Y	N	Y	Y	N	Y	N	N	5
Pegurri 2005	C	C	Y	N	N	Y	Y	Y	N	N	N	4
Rahman 2013	N	Y	Y	Y	N	Y	N	N	Y	Y	N	6
Ruizendaal 2014	N	Y	N	N	Y	Y	Y	Y	Y	N	N	6
Schiffman 2010	N	N	N	N	N	Y	N	N	N	N	N	1
Smith 2013	Y	Y	N	Y	Y	Y	N	N	Y	N	N	6
Spaulding 2009	Y	Y	Y	Y	N	Y	Y	N	N	Y	N	7
Sunguya 2013	Y	Y	Y	N	N	Y	Y	Y	Y	Y	N	8
Toczek 2012	Y	N	Y	N	N	Y	N	N	Y	N	N	4
van Ginneken 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11
Velema 2008	N	C	Y	Y	N	Y	N	N	NA	N	N	3
Volmink 2007	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	9
Vouking 2013	N	Y	Y	Y	N	Y	N	N	N	N	N	4
Vouking 2014	N	Y	Y	Y	N	N	N	N	N	N	N	3
Wiley-Exley 2007	N	N	N	N	N	Y	N	N	N	N	N	1
Wouters 2012	N	Y	Y	N	N	Y	N	N	Y	N	N	4

Appendices to Chapter 4

Appendix 4a: Search Strategies

AMED, Ovid

Searched using the terms in the main body of the text. The search was conducted on April 15, 2015 and produced 246 hits.

BiblioMap, EPPI-Centre

Keyword search: Person(s) providing the intervention: community worker OR health promotion practitioner

The search was conducted on April 15, 2015 and produced 581 hits.

CINAHL, EBSCO

"(MM "Community Health Workers") OR ((community OR village) N2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*)) OR ((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) n2 (worker* or volunteer*)) OR (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or colaborador voluntario or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or drug-kit manager or anganwadi* or "barefoot doctor" or distributeurs communautaires or relais communautaires or agentes polivalentes elementares) OR ((saude or salud) N2 (agente or voluntario or tecnicos or promotores)) OR ((health* or medical*) N3 (auxiliar* or attendant* or allied or village or volunteer))

Limiter: Randomised Controlled Trial

The search was conducted on April 15, 2015 and produced 153 hits.

Cochrane Central Register of Controlled Trials (CENTRAL)

1. MeSH descriptor: [Community Health Workers] explode all trees
2. (community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)
3. (village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) near/2 (worker* or volunteer*)
4. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or colaborador voluntario or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or drug-kit manager or anganwadi* or "barefoot doctor" or distributeurs communautaires or relais communautaires or agentes polivalentes elementares)
5. (saude or salud) near/2 (agente or voluntario or tecnicos or promotores)
6. (health* or medical*) near/3 (auxiliar* or attendant* or allied or village or volunteer)
7. #1 or #2 or #3 or #4 or #5 or #6

The search was conducted on April 15, 2015 and produced 3261 hits.

Databases within ISI Web of Science

(TS=((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*)) OR TS=((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) near/2 (worker* or volunteer*)) OR TS=(accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or colaborador voluntario or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or drug-kit manager or anganwadi* or "barefoot doctor" or distributeurs communautaires or relais communautaires or agentes polivalentes elementares) OR TS=((saude or salud) NEAR/2 (agente or voluntario or tecnicos or promotores)) OR TS=((health* or medical*) NEAR/3 (auxiliar* or attendant* or allied or village or volunteer))) AND TS=(RCT OR random* or factorial* or crossover* or placebo* or trial or groups) AND WC=(Health Care Sciences & Services OR Health Policy & Services OR Nursing)
Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH
Timespan=All years

The search was conducted on April 15, 2015 and produced 1869 hits.

EMBASE, Ovid

Searched using the terms in the main body of the text. The search was conducted on April 15, 2015 and produced 751 hits.

Global Health, Ovid

1. ((community or village) adj2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)).tw.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) adj2 (worker)).tw.
3. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "colaborador voluntario" or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or "drug-kit manager" or anganwadi* or "barefoot doctor" or "distributeurs communautaires" or "relais communautaires" or "agentes polivalentes elementares").tw.
4. ((saude or salud) adj2 (agente or voluntario or tecnicos or promotores)).tw.
5. ((health* or medical*) adj3 (auxiliar* or attendant* or allied or village or volunteer)).tw.
6. 1 or 2 or 3 or 4 or 5
7. randomised controlled trial.pt.
8. controlled clinical trial.pt.
9. randomised.ab.
10. placebo.ab.
11. randomly.ab.
12. trial.ab.
13. groups.ab.
14. 7 or 8 or 9 or 10 or 11 or 12 or 13
15. 7 and 14

The search was conducted on April 15, 2015 and produced 1670 hits.

Impact Evaluation Database (3ie)

Keyword: community health worker

Evaluation method: Randomised Control Trials (RCT)

The search was conducted on July 12, 2016 and produced 62 hits.

IPA (Innovations for Poverty Action) Database

Keyword: community

Program area: Health

The search was conducted on July 12, 2016 and produced 6 hits.

JOLIS

Keyword: "Community Health Agent" or "Community Health Aides" or "Community health promoter" or "Community mobilizer" or "Community drug distributor" or "community health worker" or "Village health worker" or "Rural Health Worker" or "Lay Health Worker"

The search was conducted on April 15, 2015 and produced 23 hits.

J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database

Keyword: Community worker

Sector: Health

The search was conducted on July 12, 2016 and produced 23 hits.

LILACS

tw:(mh:"community health workers") AND (type_of_study:("clinical_trials"))

The search was conducted on April 15, 2015 and produced 1 hit.

MEDLINE In-Process & Other Non-Indexed Citations, Ovid

1. ((community or village) adj2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)).tw.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or extension or health or community) adj2 (worker)).tw.
3. (acompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or monitora* or promotora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "colaborador voluntario" or feldsher or "shastho shebika" or "shastho karmis" or Sevika or ASHA or "drug-kit manager" or anganwadi* or "barefoot doctor" or "distributeurs communautaires" or "relais communautaires" or "agentes polivalentes elementares").tw.
4. ((saude or salud) adj2 (agente or voluntario or tecnicos or promotores)).tw.
5. ((health* or medical*) adj3 (auxiliar* or attendant* or allied or village or volunteer)).tw.
6. 1 or 2 or 3 or 4 or 5

The search was conducted on April 15, 2015 and produced 2012 hits.

MEDLINE, Ovid

Searched using the terms in the main body of the text. The search was conducted on April 15, 2015 and produced 7619 hits.

PAHO

Index term: "COMMUNITY HEALTH AIDES/"

The search was conducted on April 15, 2015 and produced 99 hits.

POPLINE, K4Health

All fields: random or factorial or crossover or cross over or cross-over or assign or allocate or allocation or RCT

Keywords: community workers or community-based distribution workers or community-based distribution or community health services

The search was conducted on April 15, 2015 and produced 39 hits.

PsycINFO, Ovid

Searched using the terms in the main body of the text. The search was conducted on

April 15, 2015 and produced 1188 hits.

TRoPHI, EPPI-Centre

Keyword search: Person(s) providing the intervention: community worker OR health promotion practitioner

The search was conducted on April 15, 2015 and produced 560 hits.

WHO Global Health Library (AIM (AFRO), IMEMR (EMRO), IMSEAR (SEARO), WPRIM (WPRO), WHOLIS (KMS))

MESH search term: Community Health Workers

The search was conducted on April 15, 2015 and produced 327 hits.

World Bank Independent Evaluation Group

Advanced Search>>[all items]>>Systematic Reviews and Impact Evaluations>>Health, Nutrition and Population

The search was conducted on July 12, 2016 and produced 7 hits.

Appendix 4b: Data Extraction Form

BASIC INFORMATION		Training description	
Record name		Refresher or ongoing training	
Year		Certification process	
Trial registration number		Key competency	
Study category		Ratio of supervisors to CHWs	
INCLUSION CRITERIA		Supervisor type	
RCT?		Supervision Frequency	
CHWs in LMIC		Supervision description	
CHWs perform interventions related to		RISK OF BIAS	

healthcare delivery			
CHWs are primarily based in the community		Adequate sequence generation	
CHWs live in the catchment they serve		Allocation concealment	
Organized training but no formal degree		Blinding	
Reason for exclusion		Incomplete Outcome data	
STUDY CHARACTERISTICS		Selective reporting	
Unit of Randomisation		Recruitment Bias	
Study design		Baseline imbalance	
Country		Loss of clusters	
Funding Source		Unit of Analysis Issues	
Healthcare Setting		Comparability with individually randomised trials	
Start date		D Adequate sequence generation	
End date		D: Allocation concealment	
Duration		D: Blinding	
Study Objective		D: Incomplete outcome reporting	
Experimental arm		D: Free of selective reporting	
Control arm		D: Recruitment bias	
Number of clusters		D: baseline imbalance	
Participants		D: Loss of clusters	
N Intervention		D: Unit of analysis	
N Control		D: Comparability	
Implementation			
Other Notes		OUTCOMES	
CHW CHARACTERISTICS		List primary & secondary outcomes	
Local term for CHW		Outcome name (for continuous outcomes)	
% Female		O1 Definition	
% Male		O1 Intervention Mean	

CHW Age		O1 Intervention SD	
Literacy		O1 Control Mean	
Education (mean yrs.)		O1 Control SD	
Incentive		O1 Follow-up	
Incentive description		O1 ICC	
Ratio CHW to Population		Outcome name (for binary outcomes)	
Setting		O2 definition	
Recruiter		O2 N Positive intervention	
Recruitment advertising		O2 N Negative intervention	
Recruitment criteria		O2 N Positive control	
Recruitment interview		O2 N Negative control	
Recruitment process or test		O2 Follow-up	
Training duration (days)		O2 ICC	

Appendix 4c: Characteristics of Studies

Characteristics of included studies (includes 'Risk of bias' tables)

Agrasada 2005

Trial reg.	None
Country	Philippines
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Mother/infant dyad
Population	LOCAL TERM FOR CHW: Counsellor RECRUITMENT CRITERIA: A postnatal care seminar was held at the hospital after which 14 out of the 37 women who attended were invited to undergo counselling training HEALTHCARE SETTING: Household PATIENT POPULATION: First-time mothers were eligible if they were 18 y or older, intended to breastfeed, and had vaginally delivered a LBW singleton with an Apgar score of 8 or higher at 5 min. The infant had to be born between 37 and 42 weeks of gestation. INTERVENTION SAMPLE SIZE: 68 CONTROL SAMPLE SIZE: (A) 69, (B)
Intervention	OBJECTIVE: Test the efficacy of postnatal peer counselling among first-time mothers that aimed to increase exclusive breastfeeding of term LBW infants. INTERVENTION DESCRIPTION: Patients received home-based counselling visits by counsellors trained in breastfeeding counselling (n=68). Eight home visits were scheduled at infant age days 3–5, 7–10 and 21, and at 1.5 months, and then monthly up to 5.5 mo. KEY CHW COMPETENCY: The breastfeeding counsellors informed mothers of the benefits of exclusively breastfeeding infants up to 6 months, and assisted mothers in preventing and managing breastfeeding problems. The childcare counsellors assisted mothers on infant care and increasing mother–infant interaction using activities such

	as infant massage and smile therapy. TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: The breastfeeding counselling and childcare counselling courses ran parallel but unknown to each other. Both courses consisted of 40 h of interactive didactics, role-playing and practical training. A certified lactation counsellor trained the breastfeeding counsellors and a maternal child healthcare specialist trained the childcare counsellors REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: A certified lactation counsellor monitored the breastfeeding counsellors and a maternal child healthcare specialist monitored the childcare counsellors INCENTIVE: Monetary INCENTIVE DESCRIPTION: The peer counsellors were not paid a salary, but received local transport costs during training and home visits. RATIO CHW TO POPULATION: 1:8 START DATE: 01/2001 END DATE: 08/2002 CONTROL DESCRIPTION: (A) Control group of mothers who did not receive counselling. (B) Patients received home-based counselling visits by counsellors trained in general childcare. This arm was included as a control as, in line with the protocol, CHWs were not trained to provide the intervention in question.
Outcomes	OUTCOMES: Primary: Exclusive breastfeeding prevalence Secondary: were duration of breastfeeding, Mean WAZ, and diarrhoea morbidity. DURATION OF FOLLOW UP: 6 months
Funding	Swedish International Development Cooperation Agency, InDevelop and Uppsala University

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "The allocation sequence, based on a table of random numbers..."
Allocation concealment (selection bias)	Low Risk	Quote: "Once informed consent was obtained, the research assistant drew the topmost, opaque, sequentially numbered envelope from the pack. The mother's and infant's names were first written on the envelope, then copied into a tracking logbook. The research assistant then opened the envelope to determine the group assignment of the mother. "
Blinding (performance and detection bias)	Low Risk	Quote: "Both the physician and the trained interviewer [collecting data] were unaware of which group the mother belonged to."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "All analyses were by intention to treat."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned

Ali 2003 (incl. Gul 2004)

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Counsellors RECRUITMENT CRITERIA: Out of 73 women who came for the interview, 21 were selected on the basis of their communication skills, motivation, attitude, ability to read and write Urdu, and freedom to move in the community. HEALTHCARE SETTING: Household

	PATIENT POPULATION: Women with anxiety or depression based on screening w/indigenous instrument INTERVENTION SAMPLE SIZE: 216 CONTROL SAMPLE SIZE: 150
Intervention	OBJECTIVE: To assess if eight counselling sessions conducted by women minimally trained as community counsellors could reduce the mean level of anxiety and/or depression in women of their own community INTERVENTION DESCRIPTION: Counselling subjects once a week for eight weeks. KEY CHW COMPETENCY: Provide counselling to women with depression TRAINING DURATION (DAYS): 11 TRAINING DESCRIPTION: Training encompassed basic information regarding anxiety/depression, stress/anger management, and communication/counselling skills. Communication covered active listening, proving, and feedback, whereas counselling dealt with supportive, problem-solving, and cognitive-behavioural techniques. The model used was participatory and facilitating. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:10 START DATE: 01/2001 END DATE: 07/2001 CONTROL DESCRIPTION: No treatment
Outcomes	OUTCOMES: Anxiety and depression ("The Aga Khan University Anxiety and Depression Scale (AKUADS)) DURATION OF FOLLOW UP: 2 months
Funding	Aga Khan Research Council

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Using computer-generated random numbers, 216 cases were randomised to the intervention and 150 to the control group"
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned
Blinding (performance and detection bias)	Low Risk	Quote: "The data collectors were not from the same community and were blind to the intervention and control groups"
Incomplete outcome data (attrition bias)	High Risk	Quote: "The dropout rate was 44% in the intervention group and 39% in the controls...There was no significant difference found between the dropout rates of the two groups (p value=0.481) and in the socio-demographic characteristics of the dropouts and non-dropouts within the two groups"
Selective reporting (reporting bias)	Unclear Risk	No protocol identified

Andersson 2009

Trial reg.	ISRCTN12421731
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Enumeration areas of 4-5 villages CLUSTERS: 18 Intervention, 14 Control
Population	LOCAL TERM FOR CHW: Field team

	RECRUITMENT CRITERIA: Recruited by study team. HEALTHCARE SETTING: Community PATIENT POPULATION: Children <60 months INTERVENTION SAMPLE SIZE: 536 CONTROL SAMPLE SIZE: 422
Intervention	OBJECTIVE: Test the effect on uptake of informed discussion of vaccination costs and benefits, without relying on improved health services. INTERVENTION DESCRIPTION: The intervention comprised three structured discussions separately with male and female groups in each cluster. The first discussion shared findings about vaccine uptake from the baseline study; the second focussed on the costs and benefits of childhood vaccination; the third focussed on local action plans. KEY CHW COMPETENCY: The activities of the field teams included: meeting community leaders to explain the purpose of the intervention and seek permission to work in the community; identifying suitable members for the discussion groups; scheduling and facilitating the three phases of discussion (sometimes requiring several meetings for a phase); and assisting the groups to list local barriers to vaccination and develop action plans. TRAINING DURATION (DAYS): 6-9 TRAINING DESCRIPTION: Classroom sessions and field practice. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Local supervisors supported and monitored the work of the field teams and documented the outcome of the three phase discussions, using structured checklists and reporting formats. They visited the teams in the field, provided feedback, and assisted them to remedy any problems encountered with the intervention implementation. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 08/2006 END DATE: 03/2007 CONTROL DESCRIPTION: Both intervention and control clusters received a district-wide health promotion programme emphasizing household hygiene.
Outcomes	OUTCOMES: Primary: uptake of measles and full DPT vaccination. Secondary: "cascada" of intermediate outcomes leading to vaccination uptake: conscious knowledge, attitudes about vaccination, subjective norms, intention to change, agency/self efficacy, and discussion within the household DURATION OF FOLLOW UP: 1 year
Funding	International Development Research Centre

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "a computer generated random number sequence assigned 18 intervention and 14 control clusters."
Allocation concealment (selection bias)	Low Risk	Quote: "The sequence was concealed and the intervention assigned centrally."
Blinding (performance and detection bias)	Low Risk	Quote: "Interviewers in the household surveys were blind of intervention status of different clusters."
Incomplete Outcome data	Low Risk	Quote: "All analyses followed the intention to treat principle, considering all children in designated intervention clusters as exposed irrespective of parental participation in the structured discussions"
Selective reporting (reporting bias)	Low Risk	Comment: The primary outcome noted in the protocol (immunization) is reported in the study
Recruitment Bias	Low Risk	Comment: All children in the enumeration areas were included in the analysis

Baseline imbalance	Low Risk	Comment: Groups significantly different on "willing to travel to vaccinate." Study adjusted for baseline differences using a logit model generalised estimating equation
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of Analysis Issues	Low Risk	Quote: "Cluster was the unit of randomisation, intervention and principal analysis."

Aung 2014 (incl. Bishai 2015)

Trial reg.	ISRCTN73606238
Country	Myanmar
Methods	STUDY DESIGN: C-RCT, matched-pairs UNIT OF RANDOMISATION: Village tracts (i.e. clusters of villages) CLUSTERS: Intervention: 52, Control: 52
Population	LOCAL TERM FOR CHW: Community Health Workers RECRUITMENT CRITERIA: Minimum qualifications were middle school education, residence in the intervention village tract and previous health experience such as an auxiliary midwife or a Red Cross aide. HEALTHCARE SETTING: Household PATIENT POPULATION: Children <5 years old INTERVENTION SAMPLE SIZE: N.S. (total = 3605) CONTROL SAMPLE SIZE: N.S. (total = 3605)
Intervention	OBJECTIVE: Assess ORS plus zinc uptake in rural Myanmar. INTERVENTION DESCRIPTION: Creation of a social franchising network of local community health workers by Population Services International, an international non-profit organization, to provide the village tracts with access to ORS plus zinc, to provide health information, to make referrals to district facilities and to prescribe and sell ORS plus zinc. Social franchising is the application of commercial methods to the delivery of subsidized services with a social benefit goal, such as improving rural health. PSI established a supply network to maintain a stock of ORS packaged with 10 days of zinc supplementation to increase access and for health workers to sell at a low, subsidized cost. The wholesale price was 100 kyats (US 13 cents) per pack of two ORS units plus 10 zinc tablets, estimated as the at-cost price of purchase and delivery. The recommended price for health workers to sell was 200 kyats—lower than the 300–500 kyats for other ORS brands without zinc available in stores in the study area. KEY CHW COMPETENCY: Physical exam of children w/diarrhoea (i.e. whether to prescribe ORS plus zinc alone or to prescribe and refer to district health facilities). Referrals are made for children with at least one sign of dehydration or severe diarrhoea. TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: Curriculum, based on national and WHO standards, on the prevention and treatment or referral of diarrhoea, malaria, pneumonia and tuberculosis, and provision of contraception. REFRESHER OR ONGOING TRAINING: None SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Health workers were initially given 10 packs, and restocked at monthly supply and monitoring visits. INCENTIVE: Monetary INCENTIVE DESCRIPTION: An incentive system was created for referral of severe cases of diarrhoea to district health centres, assigning five points for each case completing referral, with bonuses of 20 000 kyats for 100 points and 25 000 kyats for 150 points and above. RATIO CHW TO POPULATION: N.S. START DATE: 13/09/2010 END DATE: 07/2011 CONTROL DESCRIPTION: Standard ORS access: The control condition was the situation before programme activities, which comprised access to various brands of

	ORS without zinc for sale in formal and informal general stores and access to ORS plus zinc only in larger urban areas from commercial pharmacies and government clinics.
Outcomes	OUTCOMES: Primary: difference in the per cent of cases of diarrhoea among children <5 years old treated with ORS plus zinc. Secondary: the difference in total ORS (i.e. with or without zinc supplementation) in the intervention vs. control village tracts, cost DURATION OF FOLLOW UP: 10 months
Funding	Grant from Population Services International - Myanmar to the University of California San Francisco (#A115100), with funding from the Bill and Melinda Gates Foundation and USAID.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)?	Low Risk	Quote: "The names of the two village tracts in a matched pair were written on pieces of paper; one person folded them, put them in a box, closed the lid and shook the box. A different person drew one of the folded pieces of paper; a third recorded the name of the village tract drawn as the control and the one remaining as intervention. After the process, no substitutions or changes were made to assignments."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	High Risk	Quote "The community-level nature of the intervention precluded blinding survey respondents to the presence of a health worker in their village tract. To the extent possible, interviewers were blinded to the study arm by not recording the assignment on survey materials, not including implementation staff on the survey team, not including interviewers in the randomisation and asking program staff to not discuss program presence with interview staff...Analysts were not blinded to study assignment."
Incomplete Outcome data	Low Risk	Comment: Intention to treat analysis
Selective reporting (reporting bias)	Unclear Risk	Comment: Protocol not available
Recruitment Bias	High Risk	Comment: Individuals were recruited to the trial after the clusters had been randomised
Baseline imbalance	Low Risk	Quote: "No baseline imbalance in factors that might influence the outcome of the intervention, including diarrhoea period prevalence, village characteristics, knowledge of caregivers, previous use of ORS, knowing where to obtain ORS and amount willing to pay for ORS."
Loss of clusters	Low Risk	Comment: 11 clusters lost on both sides, but intention to treat analysis
Unit of Analysis Issues	Low Risk	Quote: "accounting for clustering at the village tract level"

Azad 2010

Trial reg.	ISRCTN54792066
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Unions CLUSTERS: Intervention: 9; Control: 9
Population	LOCAL TERM FOR CHW: Facilitator

	RECRUITMENT CRITERIA: A local female peer facilitator who acted as a catalyst for community mobilisation. HEALTHCARE SETTING: Community PATIENT POPULATION: Women aged 15–49 years who had given birth during the study period INTERVENTION SAMPLE SIZE: 17514 CONTROL SAMPLE SIZE: 18599
Intervention	OBJECTIVE: Assess the effectiveness of a scaled-up development programme with women's groups to address maternal and neonatal care in three rural districts of Bangladesh. INTERVENTION DESCRIPTION: All clusters received health services strengthening and basic training of traditional birth attendants. In intervention clusters, a facilitator convened 18 groups every month to support participatory action and learning for women, and to develop and implement strategies to address maternal and neonatal health problems. The three main strategies implemented by women's groups were the creation of emergency funds managed by the groups, the raising of awareness of maternal and newborn health issues during meetings and in the community by use of materials provided by the groups, and the fostering of effective communication with health-care providers through meetings with group representatives. These strategies changed over time and varied between groups. KEY CHW COMPETENCY: The role of the facilitator was to activate and strengthen groups, to support them in identifying and prioritising maternal and neonatal problems, to help to identify possible strategies, and to support the planning, implementation, and monitoring of strategies in the community. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Facilitators received five training sessions that covered participatory modes of communication and maternal and neonatal health issues. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Locally recruited supervisors supported facilitators in preparing for meetings and liaising with community leaders. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:1414 (per group; every facilitator was responsible for 18 groups) START DATE: 01/02/2005 END DATE: 31/12/2007 CONTROL DESCRIPTION: Control and intervention clusters all received health services strengthening and basic training of traditional birth attendants.
Outcomes	OUTCOMES: Primary: neonatal mortality rate (NMR) Secondary: Maternal deaths, stillbirths, uptake of antenatal and delivery services, home-care practices during and after delivery (Any antenatal care, ≥ 4 ANC visits, Any iron tablets, Maternal tetanus-toxoid injection, Institutional deliveries Home deliveries: Attended by formal provider (doctor or nurse), Safe delivery kit used, Plastic sheet used, Cord cut with new or boiled blade, Appropriate cord care, Live births (home deliveries), Infant wiped within 30 min, Infant not bathed in first 24 h, Any of three infant illnesses (cough, fever, diarrhoea), Health-care-seeking behaviour in event of infant illness, Exclusive breastfeeding for first 6 weeks) infant morbidity, health-care-seeking behaviour, perinatal mortality, and early and late NMR. DURATION OF FOLLOW UP: 3 years
Funding	Women and Children First, the UK Big Lottery Fund, Saving Newborn Lives, and the UK Department for International Development.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence	Low Risk	Quote: "For each district, cluster names were written

generation (selection bias)		on pieces of paper, which were folded and placed in a bottle. The first three cluster names drawn from the bottle were allocated to the intervention group and the remaining three to control."
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with sequence generation
Blinding (performance and detection bias)	High Risk	Quote: "Neither study investigators nor participants were masked to treatment assignment. "
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat." Comment: Exclusion (0%) and attrition (14.2%) low
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in trial registration are reported
Recruitment bias	Low Risk	Comment: Women recruited after randomisation but recruitment offered to all eligible women in the village
Baseline imbalance	High Risk	Quote: "The control clusters included three tea garden estates that had substantially worse health and socioeconomic indicators than did the rest of the study area. In these areas, surveillance started late because of entry restrictions...baseline differences in maternal education, maternal age, and household assets between intervention and control clusters: mothers in the intervention clusters were slightly younger, less educated, and had fewer household assets. We adjusted for these covariates"
Loss of clusters	Low risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "The resulting regression model was used to calculate ratio residuals for each cluster, which were then used in place of cluster-level observations for a stratified t test described above for unadjusted analysis."

Baqui 2008 (incl. Baqui 2016; Baqui 2009; LeFevre 2013)

Trial reg.	NCT00198705
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT (3-arm) UNIT OF RANDOMISATION: Unions (i.e. the smallest local gov't entities n Bangladesh) in Sylhet district w/average population of 20 000 each CLUSTERS: Intervention A: 8; Intervention B: 8, Control: 8
Population	LOCAL TERM FOR CHW: Community health worker RECRUITMENT CRITERIA: At least a 10th grade education (Baqui 2009) HEALTHCARE SETTING: Household PATIENT POPULATION: Ever-married women of reproductive age (15–49 years old). Mean age ~28 years, ~3 years of schooling INTERVENTION SAMPLE SIZE: A: 36059, B: 40159 CONTROL SAMPLE SIZE: 37598
Intervention	OBJECTIVE: Evaluate community-based intervention package through government and non-government organisation infrastructures to reduce neonatal mortality. INTERVENTION DESCRIPTION: (A) Home-care arm: female community health workers identified pregnant women, made two antenatal home visits to promote birth and newborn-care preparedness, made postnatal home visits to assess newborns on the first, third, and seventh days of birth, and referred or treated sick neonates. (B) Community-care arm: birth and newborn-care preparedness and care seeking from qualified providers were promoted solely through group sessions held by female and male community mobilisers. KEY CHW COMPETENCY: Identified pregnant women, made two antenatal home visits to promote birth and newborn-care preparedness, made postnatal home visits to assess newborns on the first, third, and seventh days of birth, and referred or

	treated sick neonates TRAINING DURATION (DAYS): 42 TRAINING DESCRIPTION: Hands-on supervised training in a tertiary-care hospital and in households. The training included skills development for behaviour-change communication, provision of essential newborn care, clinical assessment of neonates, and management of sick neonates with an algorithm adapted from the integrated management of childhood illness. REFRESHER OR ONGOING TRAINING: Midway through the intervention, a three-day refresher training was conducted (Baqui 2009) SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: The supervisors' work schedule was organized so that they spent at least two days per month accompanying each CHW in their respective group, evaluating her work performance using a structured checklist for observations and providing immediate feedback. CHWs also had the opportunity to discuss field problems and to obtain feedback in fortnightly group meetings with the senior supervisors. (Baqui 2009) INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:4000 START DATE: 06/2003 END DATE: 12/2005 CONTROL DESCRIPTION: Families in the comparison arm received the usual health services provided by the government, non-government organisations, and private providers.
Outcomes	OUTCOMES: Primary: NMR. Secondary: antenatal visits from trained providers, use of iron and folic-acid supplements, use of clean cord-cutting instruments, delays in the newborn's first bath, initiation of breastfeeding within 1 h of birth, tetanus-toxoid immunisation coverage, cause-specific mortality, cost. DURATION OF FOLLOW UP: 3 years
Funding	United States Agency for International Development by cooperative agreements with the Johns Hopkins Bloomberg School of Public Health, International Centre for Diarrhoeal Disease Research, Bangladesh, and the saving newborn lives programme by Save the Children (US) with a grant from the Bill and Melinda Gates Foundation.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Computer-generated pseudo-random number sequence without stratification or matching"
Allocation concealment (selection bias)	Low Risk	Quote: "The computer-generated randomisation was implemented by a study investigator who had no role in the implementation of the study."
Blinding (performance and detection bias)	Low Risk	Quote: "The nature of the intervention meant masking was unachievable..." Quote: "team of data collectors, supervisors, and researchers who had no role in the implementation of the intervention mapped and listed..."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Analysis was by intention to treat
Selective reporting (reporting bias)	Unclear Risk	Comment: Protocol not mentioned
Recruitment bias	Low Risk	Comment: attempts were made to cover all live births in each cluster
Baseline imbalance	Low Risk	Quote "Mother's age and education, birth order, child's sextant household wealth were similar at baseline across study arms for a sample of all women who had a live birth during 2002" Comment: No baseline differences in NMR in three trial

		arms
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of Analysis Issues	Low Risk	Quote: "Analysis was by intention to treat at the cluster level."

Bari 2006

Trial reg.	None
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Union CLUSTERS: Intervention: 6, Control: 6
Population	LOCAL TERM FOR CHW: Community Health Worker (CHW) RECRUITMENT CRITERIA: CHWs resided in the population they would serve HEALTHCARE SETTING: Household PATIENT POPULATION: Neonates INTERVENTION SAMPLE SIZE: CONTROL SAMPLE SIZE:
Intervention	OBJECTIVE: Improve maternal and newborn-care practices and care-seeking for maternal and newborn conditions through: (a) behaviour change communication, (b) identification and referral of sick newborns in the community, and (c) strengthening of neonatal care in health facilities. INTERVENTION DESCRIPTION: CHWs delivering a package of maternal and newborn-care interventions in the home KEY CHW COMPETENCY: The CHWs carried out bi-monthly pregnancy surveillance and registration of married WRA and made home-visits in the third and the eighth month of pregnancy to counsel families BNCP. After delivery, the CHWs made home-visits to promote evidence-based domiciliary newborn care and to identify and refer sick newborns and mothers on day 0 (day of birth), 3, 6, and 9. CHWs also facilitated referrals. For families that refused referral and for babies who had at least one danger sign of very severe disease, or any two from possible very severe disease, the CHWs carried out home-treatment of cases with suspected serious infection with oral cotrimoxazole. TRAINING DURATION (DAYS): 30 TRAINING DESCRIPTION: Training to equip them to provide a package of maternal and newborn care. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: The Field Supervisors supervised the daily activities of CHWs in the community and also reviewed the key performance and process indicators relating to coverage and quality every fortnight and provided feedback to the CHWs. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:4000 START DATE: 01/2003 END DATE: 09/2005 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Care-seeking from qualified providers, care-seeking from unqualified providers DURATION OF FOLLOW UP: 2.5 years
Funding	Infectious Disease Initiative of the Wellcome Trust–Burroughs Wellcome Fund. Additional support was provided by DFID; the United States Agency for International Development, Office of Health, Infectious Diseases, and Nutrition, Global Bureau through the Global Research Activity Cooperative Agreement (No. GHS-A-00-03-00019-00); the Government of Bangladesh (Improved Health for the Poor); and Save the Children-USA through a grant from the Bill and Melinda Gates Foundation.

<i>Risk of Bias</i>		
Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "randomly allocated to each study arm..." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified.
Blinding (performance and detection bias)	High Risk	Quote: "Data were drawn from records maintained by the CHWs, referral hospital registers"
Incomplete outcome data (attrition bias)	Unclear Risk	Comment: Follow-up rate not clear
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned
Recruitment bias	Low Risk	Quote: "Identified through a systematic random-sampling method..."
Baseline imbalance	Low Risk	Quote: "At the time of the baseline household survey in January 2003, there was no significant difference between the intervention arm and the comparison arm of the study in the proportion of sick newborns who were taken outside the home for care from any qualified provider (including Kumudini Hospital), to the Kumudini Hospital specifically, or to an unqualified provider"
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of Analysis Issues	Unclear Risk	Comment: Not mentioned

Bari 2011 (incl. Sadruddin 2012)

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: C-RCT (equivalence trial) UNIT OF RANDOMISATION: Union councils CLUSTERS: Intervention: 14, Control: 14
Population	LOCAL TERM FOR CHW: Lady Health worker RECRUITMENT CRITERIA: Is available for sick visits whenever needed, actively visits five households daily and covers all households monthly. HEALTHCARE SETTING: Household PATIENT POPULATION: 2-59 months of age with severe pneumonia INTERVENTION SAMPLE SIZE: 1995 CONTROL SAMPLE SIZE: 1477
Intervention	OBJECTIVE: To determine whether CCM of severe pneumonia by LHW using oral amoxicillin was equivalent to current standard of care. INTERVENTION DESCRIPTION: Treat pneumonia with oral amoxicillin KEY CHW COMPETENCY: Treat pneumonia with oral amoxicillin TRAINING DURATION (DAYS): 90 TRAINING DESCRIPTION: undergo three months of classroom training followed by 12 months of field supervision. REFRESHER OR ONGOING TRAINING: Refresher training occurred every 4 months for intervention and control cluster LHWs. The contents of these refresher trainings included assessment, classification and management of acute respiratory infections including pneumonia, identification of recurrence of pneumonia, antibiotic doses by age, definitions of treatment failure, adverse events and referrals. In addition verbal scenarios, question and answer session, strength and weaknesses (based on already collected data on case report forms) of LHWs were discussed. LHWs performing well were publicly appreciated, while LHWs not performing well were given extra training and were motivated. Pre-post tests were regular part of these refresher courses.

	SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: clinically supervised by a Lady Health Visitor and administratively supervised by a Lady Health Supervisor. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:1000 START DATE: 04/2008 END DATE: 12/2009 CONTROL DESCRIPTION: LHWs provided one dose of oral cotrimoxazole and referred to a health facility (standard of care).
Outcomes	OUTCOMES: Primary: treatment failure up to day 6. Secondary: clinical relapse between day 6 and 14, cost DURATION OF FOLLOW UP: 1.5 years
Funding	United States Agency for International Development

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote " using stratified randomisation.14 Strata were defined by child population, child mortality and literacy. We randomised 28 of the 44 Haripur union councils. "
Allocation concealment (selection bias)	Unclear Risk	Comment: Done simultaneously with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "Participants, care givers, and assessors were not blinded to study therapy."
Incomplete outcome data (attrition bias)	High Risk	Quote: "Per-protocol analysis"
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Unclear Risk	Comment: Not specified
Baseline imbalance	Low Risk	Quote: "Treatment arms were similar with respect to demographic characteristics and most indicators of baseline disease severity"
Loss of clusters	Low Risk	Comment: No cluster lost
Unit of Analysis Issues	Low Risk	Quote: "adjusting for clustering within arms using generalized estimating equations."

Bojang 2011

Trial reg.	None
Country	The Gambia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Catchment population of RCH trekking teams CLUSTERS: Intervention: 13 clusters; Control: 13 clusters
Population	LOCAL TERM FOR CHW: Village health workers (VHWs) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Children up to and including the age of 6. Exclusion criteria included known allergy to any antimalarial drug or the presence of acute or chronic, clinically significant pulmonary, cardiovascular, hepatic, or renal disease. INTERVENTION SAMPLE SIZE: 6250 CONTROL SAMPLE SIZE: 6076
Intervention	OBJECTIVE: Compare two approaches to the delivery of IPTc to Gambian children: distribution by VHWs or through reproductive and child health (RCH) trekking teams. INTERVENTION DESCRIPTION: Treatment with a single dose of SP plus three doses of AQ were given at monthly intervals during the malaria transmission season

	by VHWs KEY CHW COMPETENCY: Treatment with a single dose of SP plus three doses of AQ TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: Before the start of the study, VHWs were trained for 5 d on how to distribute monthly IPTc using drug treatment charts, recognise signs and symptoms of malaria and their relation to severity of illness, and maintain a compliance ledger. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: None INCENTIVE DESCRIPTION: VHWs are meant to be supported by their community and do not receive any payment from the Ministry of Health but VHWs received US\$11 per month (US\$ 2006) during the period of drug administration RATIO CHW TO POPULATION: 1:>400 START DATE: 09/2006 END DATE: 06/2007 CONTROL DESCRIPTION: Treatment with a single dose of SP plus three doses of AQ were given at monthly intervals during the malaria transmission season by RCH trekking teams. RCH trekking teams visited each of the study villages approximately once a month to administer routine EPI vaccines and vitamin A. During the course of the intervention each trekking team in the RCH arm of the trial comprised an average of ten staff members, four of whom were provided by the study team to assist with the intervention. Trekking teams comprised staff with varying levels of training ranging from state enrolled nurses to volunteers. Each member of the team employed by the Ministry of Health received a salary supplement of US\$28 per month (US\$ 2006) (about 47% of their salary) during the 3 months of the intervention as an incentive for the extra hours of work that administration of IPTc incurred. Trekking team drivers received a supplement of US\$14 per month.
Outcomes	OUTCOMES: Primary: incidence of malaria. Secondary: coverage of IPTc, mean haemoglobin (Hb) concentration, and the prevalence of asexual malaria parasitaemia at the end of malaria transmission period. DURATION OF FOLLOW UP: 4 months
Funding	Gates Malaria Partnership, which receives financial support from the Bill & Melinda Gates Foundation and by the Medical Research Council, UK.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Within each stratum, clinic catchment areas were randomised, using permuted blocks using the random number generator in Stata version 10 (Stata Corporation)."
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified
Blinding (performance and detection bias)	Unclear Risk	Comment: Not specified
Incomplete outcome data (attrition bias)	Low Risk	Quote: "The primary analysis included all children in the correct age range who were present in the study area at the initial census and were issued with a malaria card, regardless of the number of treatments received."
Selective reporting (reporting bias)	High Risk	Comment: protocol indicated trial would measure the following outcomes that were not reported - the proportion of children who received three IPT courses on schedule; - the proportion of children who received partial or off-schedule IPT courses

		- the proportion of children with no IPT.
Recruitment bias	Unclear Risk	Comment: Not specified
Baseline imbalance	Unclear Risk	Quote: "The randomisation was well balanced with respect to distance from the nearest health centre, population of children up to 6 y of age, proximity to a main road, and quality of the road, factors that might influence access to the intervention and to malaria diagnosis and treatment." "Observations made during the cross-sectional survey undertaken at the end of the malaria transmission season suggested that a slightly higher proportion of children in the VHW arm slept under a bed net compared to those in the RCH arm and that more children in the VHW group were fully vaccinated compared to those in the RCH arm."
Loss of clusters	Low Risk	No cluster lost
Unit of analysis issues	Low Risk	Comment: rate differences were adjusted for clustering

Bolton 2003 (incl. Bass 2006; Betancourt 2012)

Trial reg.	None
Country	Uganda
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Villages CLUSTERS: Intervention: 15, Control: 15
Population	LOCAL TERM FOR CHW: Group leader RECRUITMENT CRITERIA: A local person of the same sex as the group HEALTHCARE SETTING: Community PATIENT POPULATION: Adult men or women who met DSM-IV criteria for major depression or subsyndromal depression INTERVENTION SAMPLE SIZE: 163 CONTROL SAMPLE SIZE: 178
Intervention	OBJECTIVE: Test the efficacy of group interpersonal psychotherapy in alleviating depression and dysfunction and to evaluate the feasibility of conducting controlled trials in Africa. INTERVENTION DESCRIPTION: The intervention villages received group interpersonal psychotherapy for depression as weekly 90-minute sessions for 16 weeks. During each session, the group leader reviewed each participant's depressive symptoms. The participant was then encouraged to describe the past week's events and to link those events to his/ her mood. The group leader then facilitated support and suggestions for change from other group members. KEY CHW COMPETENCY: Provide group interpersonal psychotherapy for depression TRAINING DURATION (DAYS): 14 TRAINING DESCRIPTION: Intensive training in group interpersonal psychotherapy by 2 of the authors REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Weekly payment of an unspecified amount for the duration of the intervention. RATIO CHW TO POPULATION: N.S. START DATE: 02/2002 END DATE: 06/2002 CONTROL DESCRIPTION: Study participants in the control villages did not receive IPT; however, participants in both control and intervention villages were free to seek whatever other interventions they wished throughout the study

Outcomes	OUTCOMES: Depression and dysfunction severity scores on scales adapted and validated for local use; proportion of persons meeting DSM-IV major depression diagnostic criteria. Follow-up: depression diagnosis, depressive symptoms and functional impairment. DURATION OF FOLLOW UP: 2 weeks post intervention, 6 month follow-up
Funding	Ruth and David Levine and by funds from the International Program on Refugee Trauma, New York, NY.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Random assignment was performed by enumerating the villages and using a random- number table to determine study allocation. "
Allocation concealment (selection bias)	Unclear Risk	Comment: Eligibility criteria were modified to exclude people whose age varied widely from the rest of those selected in each village to ensure better outcomes with group IPT (based on previous experiences)
Blinding (performance and detection bias)	Low Risk	Quote: "Because interviewers likely knew which villages had received the intervention, all interviewees were taken to the same town and interviewed there. They were instructed not to mention their village during the interview or whether they had received IPT."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "The second and third analyses were on an intention-to- treat basis to allow for possible selection bias by participants or interviewers."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned
Recruitment bias	Low Risk	Quote: "After creation of the lists of eligible participants, 8 of the 15 male villages and 7 of the 15 female villages were randomly assigned..."
Baseline imbalance	High Risk	Quote: "There was a significant difference in the proportions who met the original depression diagnostic criteria, both among those who completed the study and all those on the original lists of eligible participants" Comment: Data only adjusted for sex, years of education, and age
Loss of clusters	Low Risk	Comment: No clusters lost.
Unit of analysis issues	Low Risk	Quote: "Since the unit of analysis is the participant and the study design is a cluster randomised trial, we used a 2-level model to allow for within-village (cluster) correlation and between-village variability."

Bolton 2007

Trial reg.	NCT00280319
Country	Uganda
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Participant
Population	LOCAL TERM FOR CHW: Facilitator RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household/Community PATIENT POPULATION: Aged 14-17 years, living in displaced persons camps in northern Uganda, and meet study-specific criteria for depression INTERVENTION SAMPLE SIZE: 105 CONTROL SAMPLE SIZE: (A) 104, (B) 105
Intervention	OBJECTIVE: To assess effect of locally feasible interventions on depression, anxiety, and conduct problem symptoms among adolescent survivors of war and displacement in northern Uganda. INTERVENTION DESCRIPTION: Intervention A: Psychotherapy-based

	intervention [group interpersonal psychotherapy]. Intervention groups met weekly for 16 weeks. 16 weekly group meetings, lasting 1.5 to 2 hours each. These were preceded by 1 or 2 individual meetings in which the intervention was explained and a treatment plan was generated by the IPT-G facilitators. Interpersonal psychotherapy is a time-limited, manualised intervention developed in the United States for ambulatory, nonpsychotic unipolar depression and later adapted for use with adolescents with depression, individually and in groups. Interpersonal psychotherapy assumes that depressive episodes are triggered by difficulties in 1 or more of 4 interpersonal areas: grief, interpersonal disputes, role transitions, and interpersonal deficits. Interpersonal psychotherapy focuses on improving depressive symptoms and functioning by identifying the interpersonal problem(s) most relevant to the current depression and then assisting the individual in building skills to manage those problems. Groups consisted of 6 to 8 adolescent boys or girls and a facilitator of the same sex. KEY CHW COMPETENCY: Facilitate IPT-G sessions TRAINING DURATION (DAYS): 14 TRAINING DESCRIPTION: 2 weeks of on-site training by Columbia University faculty. A treatment manual specifying IPT-G strategies and techniques was adapted for local use. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: Facilitators received weekly direct supervision by World Vision Uganda staff with prior IPT-G experience. These supervisors in turn received weekly phone supervision from the US-based trainer and were in weekly contact with study personnel. IPT-G supervisors also provided weekly written reports that were reviewed and discussed with study staff during the phone meetings for adherence to the treatment model and to monitor human subjects protection. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:9 START DATE: 05/2005 END DATE: 12/2005 CONTROL DESCRIPTION: (A) Wait-list control group (B) Intervention B: Activity-based intervention [creative play]. Two individuals (1 man, 1 woman) from War Child Holland jointly facilitated all creative play sessions. They did not meet inclusion criteria so this arm was excluded.
Outcomes	OUTCOMES: Primary: a decrease in score (denoting improvement) on a depression symptom scale. Secondary: improvements in scores on anxiety, conduct problem symptoms, and function scales. DURATION OF FOLLOW UP: 1 month post intervention
Funding	World Vision and War Child Holland

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "random allocation was done by computerized generation of a random number between 1 and 400 for each eligible participant, ordering them by number and assigning the first third to IPT-G, the second third to CP and the final third to the wait-control group. "
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified who allocated them and whether the allocation was concealed to them. There were no major differences in baseline prognostic variables or outcome measures
Blinding (performance and detection bias)	Low Risk	Quote: "Interviewers were blinded to interviewees' intervention status."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Analysis was by intent-to-treat

Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in trial registration were reported
---	----------	--

CDI Study Group 2010 (incl. UNICEF/UNDP/World Bank/WHO TDR 2008)

Trial reg.	None
Country	Cameroon, Nigeria and Uganda
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Health district CLUSTERS: Intervention: 28; Control: 7
Population	LOCAL TERM FOR CHW: Community implementer RECRUITMENT CRITERIA: Selected by community HEALTHCARE SETTING: Community/household PATIENT POPULATION: Specific to intervention: children aged 6 to 59 (vitamin A supplementation); children less than 5 years of age (home management of malaria; insecticide-treated nets;) and pregnant women (bed-nets; insecticide-treated nets); TB patients (DOTS) INTERVENTION SAMPLE SIZE: N.S. CONTROL SAMPLE SIZE: N.S.
Intervention	OBJECTIVE: To determine the extent to which the community-directed approach used in onchocerciasis control in Africa could effectively and efficiently provide integrated delivery of other health interventions. INTERVENTION DESCRIPTION: In the trial districts four other established interventions – vitamin A supplementation, use of insecticide-treated nets, home management of malaria and short-course, directly-observed treatment for tuberculosis patients – were progressively incorporated into a community-directed intervention (CDI) process. In CDI, the community takes charge of the process, usually through a series of community meetings where the roles and responsibilities of the community are discussed and the community decides how, when and where the intervention will be implemented and by whom; how implementation will be monitored, and what support (financial or otherwise), if any, will be provided to implementers. The community then collectively selects the implementers. KEY CHW COMPETENCY: Deliver vitamin A supplementation, insecticide-treated nets, home management of malaria and short-course, directly-observed treatment for tuberculosis patients TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Trained by facility-based health workers REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Monitored by facility-based health workers and community members INCENTIVE: Determined on-community-by-community basis INCENTIVE DESCRIPTION: Determined on-community-by-community basis RATIO CHW TO POPULATION: N.S. START DATE: 2005 END DATE: 2007 CONTROL DESCRIPTION: The four interventions continue to be delivered in the "conventional way."
Outcomes	OUTCOMES: Vitamin A supplementation: Percentage of children aged 6 to 59 months who received vitamin A during the last treatment round; Insecticide-treated nets: Percentage of households having at least one insecticide-treated net, Percentage of children less than 5 years of age who slept under an insecticide-

	treated net the night before the study interview, and Percentage of pregnant women who slept under an insecticide-treated net the night before the study interview. Home management of malaria: Percentage of children less than 5 years of age who had had fever sometime during the prior 2 weeks and who had received appropriate treatment (i.e. treatment with a nationally- recommended antimalarial drug, at the correct dosage schedule, within 24 hours of the onset of fever). Short-course, directly-observed treatment: Treatment completion rate. Cost. DURATION OF FOLLOW UP: 3 years
Funding	APOC, the Bill & Melinda Gates Foundation, The World Bank, and TDR

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "In each site, four of the health districts were randomly designated as trial districts for CDI implementation and one was randomly designated as a comparison district" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified
Blinding (performance and detection bias)	Unclear Risk	Comment: Not specified
Incomplete outcome data (attrition bias)	High Risk	Quote: "For evaluation purposes, 10 villages were randomly selected in each district. Five households were randomly selected from each community, so that each study site had a total of 50 evaluation villages and 250 evaluation households." Comment: It is unclear what proportion of the total study population this represents
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Low Risk	Comment: Households randomly selected for evaluation
Baseline imbalance	Unclear risk	Comment: Not specified
Loss of clusters	Unclear Risk	Comment: Not specified
Unit of analysis issues	Unclear Risk	Comment: Not specified

Chang 2010 (Chang 2013)

Trial reg.	NCT00675389
Country	Uganda
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Clinics CLUSTERS: Intervention: 10, Control: 5
Population	LOCAL TERM FOR CHW: Community-based peer health workers (PHW) RECRUITMENT CRITERIA: Being a PLHIV on ART, good ART adherence for at least six months, and literacy HEALTHCARE SETTING: Household PATIENT POPULATION: All adult patients at the 15 mobile clinic sites who were either already on ART at the start of the trial or were started on ART at any time during the trial (~66% female, ~35 years old) INTERVENTION SAMPLE SIZE: 970 CONTROL SAMPLE SIZE: 366
Intervention	OBJECTIVE: To assess the effect of community-based PHW on AIDS care of adults in Rakai, Uganda. INTERVENTION DESCRIPTION: PHW tasks included clinic and home-based provision of counselling, clinical, adherence to ART, and social support. At the

	clinic, PHW tasks included providing ART counselling and support in group and individual sessions. For their home visit tasks, PHWs were initially assigned about 15 patients each who were visited biweekly. At these visits, PHWs were tasked to record on a standardized form a review of symptoms, a patient self-report of adherence, and to perform and record a pill count. PHWs were asked to counsel and educate their patients on ART adherence and general HIV/AIDS-related issues during these home visits. If patients were thought to need urgent care, PHWs were asked to alert RHSP staff and facilitate transfer to a higher level of care. KEY CHW COMPETENCY: Record on a standardized form a review of symptoms, a patient self-report of adherence, and to perform and record a pill count. Counsel and educate their patients on ART adherence and general HIV/AIDS-related issues TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: Training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. Trainers included RHSP staff and experienced PHWs from an urban-based Ugandan ART programme REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Day-to-day supervision of PHW activities was largely performed by a single RHSP staff member working part-time. INCENTIVE: Monetary and non-monetary INCENTIVE DESCRIPTION: To assist with their duties and encourage retention, PHWs were each given a bicycle, identifying t-shirts, basic supplies, and an initial monthly allowance of about 12.5 USD. RATIO CHW TO POPULATION: 1:15 (Over the study period, the CHW to patient ratio grew from a mean of around 1:15 (range 1:9–26) to about 1:28 (range 1:18–42)). START DATE: 05/2006 END DATE: 07/2008 CONTROL DESCRIPTION: The control arm continued with the usual standard of care. However, standard of care did change over the study period, as a number of changes unrelated to the PHWs were subsequently implemented by RHSP in both the PHW and control arms. These programmatic changes included a peer educator programme to promote use of preventive care services in mid-2006, the use of viral load results to guide care in late 2006, more focused ART-related health messaging in early 2007, and the use of enhanced adherence counselling, chart stickers to help identify patients failing virologically, and second-line ART provider talks in mid- to late 2007.
Outcomes	OUTCOMES: Primary: adherence and cumulative risk of virologic failure (.400 copies/mL). Secondary: virologic failure, cost. DURATION OF FOLLOW UP: 3 years and 4 months
Funding	Doris Duke Charitable Foundation, The Division of Intramural Research, The National Institute for Allergy and Infectious Diseases, National Institutes of Health, and a National Institutes of Health Training Grant (2 T32-AI07291) and Career Development Grant (1 K23-MH086338).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "We assigned clusters using unmatched, unrestricted random allocation by a drawing of lots"
Allocation concealment (selection bias)	Low Risk	Quote: "Study investigators (LWC, JK) generated the allocation sequence and implemented the randomisation"
Blinding (performance and detection bias)	High Risk	Quote: "This study was open label and unblinded."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Analysis was by intention to treat.
Selective reporting	High Risk	Comment: Protocol specifies the following outcomes

(reporting bias)		that were not reported: morbidity (new opportunistic infection or drug toxicity).
Recruitment bias	Low Risk	Comment: Not specified
Baseline imbalance	Low Risk	Quote: "Socio-demographic characteristics, immunologic and clinical stage of disease, the proportion of patients on ART, the median duration of time patients were on ART prior to the start of the trial, and the pre-trial 24 week and 48 week virologic failure rates appeared well balanced between arms."
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of analysis issues	Low Risk	Quote: "Efficacy analyses for cumulative risk of virologic failure and for virologic outcomes at specific time points from ART initiation were by intention to treat ...robust variance estimates appropriate for cluster-randomised trials"

Chang 2015

Trial reg.	NCT0136 6690
Country	Uganda
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Individual
Population	LOCAL TERM FOR CHW: Peer supporters RECRUITMENT CRITERIA: PLHIV HEALTHCARE SETTING: Household PATIENT POPULATION: HIV-infected adults (≥ 18 years) who had recently (≤ 8 weeks) received HIV Counselling and Testing INTERVENTION SAMPLE SIZE: 221 CONTROL SAMPLE SIZE: 221
Intervention	OBJECTIVE: Assess the effectiveness of a peer support intervention on HIV care engagement and preventive care utilization among pre-ART adults in Rakai District, Uganda INTERVENTION DESCRIPTION: Peers were asked to make an initial visit to assigned participants within 2 weeks of randomisation, and then visit each participant monthly. During visits, peers performed 3 categories of activities: (1) Assessment-Peers assessed participants' clinical status, clinic attendance, basic care package adherence, and sexual behaviours using a structured interview form; (2) Support-Peers provided psychosocial support, encouragement, information on and reminders of clinic appointments and the basic care package and counselling on reducing risky sexual behaviours; (3) Access-Peers provided triage to higher level care if necessary. KEY CHW COMPETENCY: Provide assessment, support, and access to a higher level of care if necessary TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: Training on confidentiality, disclosure, counselling techniques, motivational interviewing principles and skills, engagement in care, the BCP, ART, and condoms REFRESHER OR ONGOING TRAINING: Regularly conducted field-based skills reinforcement trainings. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Supervised by two part-time coordinators INCENTIVE: Monetary INCENTIVE DESCRIPTION: Peers were remunerated (~10 USD/month) and given a transport stipend (~1 USD/participant visited) for what was considered part-time work. RATIO CHW TO POPULATION: 1:20 START DATE: June, 2011 END DATE: July, 2013 CONTROL DESCRIPTION: Standard of care.
Outcomes	OUTCOMES: Engagement in care (defined as HIV clinic attendance, either self-

	reported or based upon clinical records), basic care package, preventive care use, risky sexual behaviours, and ART initiation DURATION OF FOLLOW UP: 1 year
Funding	The National Institute of Mental Health, National Institutes of Health (K23MH086338), the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institute of Health, and the Johns Hopkins University Center for AIDS Research (P30AI094189).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "A list of random numbers with block sizes of 2, 4, and 6 was computer-generated."
Allocation concealment (selection bias)	Low Risk	Quote: "This allocation sequence was detailed on cards in sequentially numbered, opaque, sealed, and stapled envelopes...After completing the baseline survey, the next envelope was opened to obtain arm assignment."
Blinding (performance and detection bias)	High Risk	Quote: "Participants were necessarily unmasked. Outcome assessment was also unblinded because intervention-specific responses were required in interview questionnaires." Comment: Only likely to affect survey-based, self-reported outcomes (e.g. currently taking ART, current adherence to each of the components of the basic care package, condom use etc.) Clinic-based outcomes (e.g. first entry into care, CD4 cell counts, and ART initiation) are not likely to be affected
Incomplete outcome data (attrition bias)	Low Risk	Quote: "modified intention-to-treat wherein all participants who had data at follow-up were included in analyses." Comment: Attrition low (11%)
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes in protocol were reported

Chatterjee 2014

Trial reg.	ISRCTN56877013
Country	India
Methods	STUDY DESIGN: Stratified RCT UNIT OF RANDOMISATION: Patients
Population	LOCAL TERM FOR CHW: Community health workers RECRUITMENT CRITERIA: At least 10 years of schooling and good interpersonal skills. HEALTHCARE SETTING: Household PATIENT POPULATION: Eligible patients aged 16–60 years had a primary diagnosis of schizophrenia according to the ICD-10-DCR; an illness duration of at least 12 months of at least moderate severity overall, as rated by the clinical global impression-schizophrenia scale INTERVENTION SAMPLE SIZE: 187 CONTROL SAMPLE SIZE: 95
Intervention	OBJECTIVE: To compare the effectiveness of a collaborative community-based care intervention with standard facility-based care. INTERVENTION DESCRIPTION: Collaborative community-based care plus facility-based care: The intervention was delivered in three phases: (1) the intensive engagement phase (0–3 months), including six to eight home visits made by community health workers; (2) the stabilisation phase (4–7 months), with sessions delivered once every 15 days; and (3) the maintenance phase (8–12 months), with sessions delivered once a month. Psycho-educational information for participants and caregivers, adherence management and health promotion strategies, stigma and discrimination coping, linkages to self-help groups and networks with community

	agencies were provided. KEY CHW COMPETENCY: Deliver a structured package of psychosocial interventions. TRAINING DURATION (DAYS): 42 TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: Continued education and training of the health workers was "an essential part" of the intervention. SUPERVISION FREQUENCY: Quarterly SUPERVISION DESCRIPTION: Community health workers at each site were supervised by intervention coordinators. Coordinators were psychiatric social workers trained in the necessary supervision and monitoring skills. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:15-25 START DATE: 01/01/2009 END DATE: 31/12/2011 CONTROL DESCRIPTION: Facility-based care alone: Care was delivered by the psychiatrists through consultations, lasted 10–15 min, and all people with schizophrenia were prescribed antipsychotic drugs. Psychiatrists also provided information about the illness, encouraged adherence to drugs, and discussed specific concerns raised.
Outcomes	OUTCOMES: Primary: change in schizophrenia symptoms and disabilities. Secondary outcomes were adherence to antipsychotic treatment and experiences of stigma and discrimination. DURATION OF FOLLOW UP: 1 year
Funding	Wellcome Trust

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "randomly assigned (2:1), via a computer-generated randomisation list with block sizes of three, six, or nine,"
Allocation concealment (selection bias)	Low Risk	Quote: "For each site, the randomisation list was generated independently by the trial statistician and transferred to the site data manager before recruitment."
Blinding (performance and detection bias)	Low Risk	Quote: "Outcome assessors were masked to group allocation."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Modified intention to treat
Selective reporting (reporting bias)	High Risk	Comment: Protocol indicated QALY would be reported

Chinbuah 2012 (Chinbuah 2013)

Trial reg.	TDR/UNDP Trial registration. A: 20189
Country	Ghana
Methods	STUDY DESIGN: Stepped-wedge C-RCT (3-arm) UNIT OF RANDOMISATION: Communities were grouped into clusters of ~100 children under 5 years of age. CLUSTERS: Intervention A: 37 Intervention B: 39 Control: 38
Population	LOCAL TERM FOR CHW: Community health workers RECRUITMENT CRITERIA: Lay persons resident in their various communities were selected through an interactive process based on criteria set by community members during an open community meeting or by their community leaders.

	HEALTHCARE SETTING: Household PATIENT POPULATION: Children 2–59 months INTERVENTION SAMPLE SIZE: Intervention A: 3994 Intervention B: 4043 CONTROL SAMPLE SIZE: 4296
Intervention	OBJECTIVE: Evaluate the impact of adding an antibiotic, amoxicillin (AMX) to an antimalarial, artesunate amodiaquine (AAQ+AMX) for treating fever among children 2–59 months of age within the HMM strategy versus standard care on all-cause mortality INTERVENTION DESCRIPTION: CHWs dispensed an antimalarial drug alone (artesunate amodiaquine [AAQ], Intervention A), or together with an antibiotic (artesunate amodiaquine plus amoxicillin [AAQ+AMX], intervention B) to children with a fever. Fever was not assessed with a thermometer. The CHWs were trained to assess for signs of severe illness. Drugs were provided at no cost to caregivers. KEY CHW COMPETENCY: Dispense antimalarial and antibiotic TRAINING DURATION (DAYS): 3 TRAINING DESCRIPTION: CHWs were trained using modified Integrated Management of Childhood Illnesses (IMCI) modules and video that had been developed with the Ministry of Health/Ghana Health Service for health workers. Training was conducted by experienced Ghana Health Service national IMCI trainers together with some study team members. Training was held separately for CHWs from different arms and was identical apart from one aspect: CHWs from AAQ+AMX clusters received additional training on how to administer amoxicillin and on how to teach caregivers from those clusters to administer amoxicillin at home. REFRESHER OR ONGOING TRAINING: Two CHW review trainings were held during the period of the study. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: Field supervisors who checked drug supplies and forms using a standard checklist. INCENTIVE: None INCENTIVE DESCRIPTION: Nominated CHWs were informed that their participation would be voluntary and without any financial remuneration. RATIO CHW TO POPULATION: N.S. START DATE: 01/2006 END DATE: 12/2009 CONTROL DESCRIPTION: Standard care
Outcomes	OUTCOMES: Primary outcome: all-cause mortality. Secondary outcomes were to determine the impact of the intervention on some morbidity indicators (anaemia, severe anaemia and severe illness). DURATION OF FOLLOW UP: 6 months
Funding	UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the Knowledge Enriches Programme of the Netherlands Organization for the Advancement of Tropical Research, WOTRO (NWO) in the Netherlands.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "computer generated randomisation"
Allocation concealment (selection bias)	Low Risk	Quote: "Clusters were...assigned random numbers using the Excel random function into the three arms"
Blinding (performance and detection bias)	Unclear Risk	Quote: "During the first year baseline information was collected. Mortality was assessed through bi-annual censuses conducted routinely by the HDSS team complimented by a Community Key Informant system (CKI)."

		Comment: Unclear if HDSS team was blinded
Incomplete outcome data (attrition bias)	Low Risk	Comment: No children excluded from analysis
Selective reporting (reporting bias)	Unclear Risk	Comment: Unable to access existing protocol
Recruitment bias	Low Risk	Quote: "Almost all caregivers agreed to the general participation of their eligible children under 5 years of age in the trial. The decline rate was 0.2%."
Baseline imbalance	Unclear risk	Quote: "Clusters randomised to the different arms were similar in all but two aspects: in the prevalence of household ownership treated nets and parasitemia in the children."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Comment: Adjusted for clustering

Chongsuvivatwong 1996

Trial reg.	None
Country	Thailand
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village CLUSTERS: Intervention: 15; Control: 15
Population	LOCAL TERM FOR CHW: Village health volunteers RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Rural PATIENT POPULATION: Children <5 years INTERVENTION SAMPLE SIZE: 664 CONTROL SAMPLE SIZE: 649
Intervention	OBJECTIVE: To test whether a 5-yr education programme on management of cases w/ARI by the Thai Ministry of Health could reduce the average number of sick days of target children. INTERVENTION DESCRIPTION: All children <5 visited by team of home visitors once a week for a maximum of 19 weeks KEY CHW COMPETENCY: To be able to detect serious ARI (defined as a respiratory rate of >50/min, or with a convulsion, or refusal to eat and drink), to give initial care by sponging and to advise referral of cases to a health centre where antibiotics (either ampicillin or cotrimoxazole) would be given and referral to the hospital is made if necessary. TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: Health workers trained by paediatricians at the regional hospital collectively conducted a workshop for study villages. Approximately 10 health volunteers from each study villages participated in the workshop. Methods of training included a video programme lecture, demonstration and simulation. The medium of teaching was a combination of Thai and local Malay dialect REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: NS. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/1991 END DATE: 05/1991 CONTROL DESCRIPTION: No intervention
Outcomes	OUTCOMES: Number of sick days, care-seeking DURATION OF FOLLOW UP: The average follow-up period was 17 weeks
Funding	Ford Foundation

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: villages were "randomly allocated" Comment: Not clear how this was done.
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified
Blinding (performance and detection bias)	Unclear Risk	Comment: unclear if outcome assessors were blinded
Incomplete outcome data (attrition bias)	Low Risk	Comment: 89% retention
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol referenced
Recruitment bias	Low Risk	Comment: Not specified
Baseline imbalance	Low Risk	Quote: "both groups of villages were well balanced in size of target population, sex distribution and person-time observation."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Since children in the same village might have a common infective agent, and thus the probability of getting infection and the duration of illness might be correlated with the same village, the generalized estimating equation method was carried out...to adjust for this correlation wherever possible"

Clarke 2005 (incl. Clarke 2005B, Clarke 2006)

Trial reg.	None
Country	South Africa
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Farm CLUSTERS: Intervention: 106; Control: 105
Population	LOCAL TERM FOR CHW: Lay health worker (LHW) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: NSP adult TB patients INTERVENTION SAMPLE SIZE: 66 CONTROL SAMPLE SIZE: 82
Intervention	OBJECTIVE: The aim of the study was to evaluate the effect of LHW recruitment, training and deployment on successful treatment completion of NSP TB patients among permanent farm dwellers. Subsidiary aims included assessment of impact on TB case finding and retreatment cases. INTERVENTION DESCRIPTION: LHWs conducted monthly weighing and TB screening of all permanent farm dweller families. They referred individuals with two or more signs and symptoms suggestive of TB to the clinic for further investigation. After laboratory confirmation of diagnosis, TB patients were issued with at most 1 month's anti-tuberculosis treatment. Where patients chose to receive DOT from the LHW, treatment adherence was recorded by the chosen LHW on cards provided by the clinics. In cases of self-supervision, LHWs played a mentoring role, visiting the patient and encouraging and monitoring treatment adherence regularly. In all cases, LHWs addressed treatment non-adherence promptly through a process of particularised motivation. KEY CHW COMPETENCY: Weigh and screen farm workers for clinical features suggestive of TB, explain the sputum specimen collection procedure and inform patients of the results, supervise drug treatment TRAINING DURATION (DAYS): 35 TRAINING DESCRIPTION: Five one-week modules: Becoming a community LHW (including capacity building skills), TB (using the standardised provincial

	training manual), Family health (including HIV/AIDS, minor ailments, foetal alcohol syndrome), First aid, and Home-based care (personal capacity building in bereavement stages). LHWs were trained to keep detailed records of their client contacts, providing data regarding their caseload and type of activities. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Trainers visited each LHW on the farm where they were working in a four- to six-week cycle to establish any problem areas that the LHWs experienced, to review and discuss the monthly weighing reports, collect the monthly report reflecting what health-related issues the LHW had attended to since the previous visit, and encourage the LHW to persevere. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Where LHWs were not employed in another capacity on the farm, farmers were requested to give them an incentive. RATIO CHW TO POPULATION: N.S. START DATE: 01/05/2000 END DATE: 31/10/2001 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Case finding and cure rates of adult new smear- positive (NSP) TB cases DURATION OF FOLLOW UP: 1.5 years
Funding	The Boland District Municipality (the local public health provider)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Farms were grouped into nine strata, using the boundaries of the nine health areas comprising the study district. Farms were allocated a number within the strata. All the numbers were randomly drawn from containers and allocated sequentially to the intervention or control group."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Comment: Unblinded, but the outcomes and the outcome measurement were not likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias)	Low Risk	Quote: "28 (26%) intervention farms failed to implement the intervention...TB data from these 28 farms were included in the analysis on an intention-to-treat basis."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned
Recruitment bias	Low Risk	Comment: Not specified
Baseline imbalance	Low Risk	Quote: "Population size, number of workers (Table 2), and TB-related variables (Table 1) on intervention and control farms were similar at baseline."
Loss of clusters	Low Risk	Comment: Analysis by intention to treat
Unit of analysis issues	Low Risk	Quote: "adjusted for clustering within farms."

Colbourn 2013 (incl. Colbourn 2013B; Colbourn 2015)

Trial reg.	ISRCTN18073903
Country	Malawi
Methods	STUDY DESIGN: C-RCT(2x2 factorial) UNIT OF RANDOMISATION: Health centre catchment areas CLUSTERS: Intervention A: 15 Intervention B: 15 Intervention C: 14 Control: 17
Population	LOCAL TERM FOR CHW: Women's Group Facilitator

	RECRUITMENT CRITERIA: Recruited from the local communities. Community members and the local chiefs assisted in identifying potential candidates using criteria including: maturity and motherhood, opinion leadership, previous experience and having received at least some secondary education. HEALTHCARE SETTING: Community PATIENT POPULATION: Pregnant women INTERVENTION SAMPLE SIZE: Intervention A: 4948 live births, Intervention B: 5107 live births CONTROL SAMPLE SIZE: Control A: 4766 live births, Control B: 5165 live births
Intervention	OBJECTIVE: Intervention A: Village women's groups which followed an 'action cycle', to identify and prioritize maternal and neonatal health problems, decide upon local solutions, advocate for, implement and evaluate such strategies Intervention B: Women's group and coaching of staff in quality improvement methodology, such as developing change ideas, conducting small tests of change using Plan-Do-Study-Act cycles, to improve care at health centres N.B. 50% of the groups had maternal and neonatal health task forces (MNHTF) added to them by the MaiKhanda programme in an attempt to enhance antenatal coverage, maternal and neonatal health (MNH) knowledge, and facility delivery. MNHTF aimed to identify high-risk pregnant women; promote delivery at a health facility, and antenatal and postnatal care; and provide health education. INTERVENTION DESCRIPTION: A rural participatory women's group community intervention and a quality improvement intervention at health centres KEY CHW COMPETENCY: Facilitate a learning action cycle TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: The WGFs received initial training and support from CMOs to facilitate the women's group action cycle. In addition they were provided with a manual that outlines the meetings and phases of the cycle and presents Participatory Rural Appraisal (PRA) tools that can be used to facilitate discussion. To assist the discussions they were also provided with a pack of picture cards illustrating some of the main maternal and neonatal health problems, contributing factors and preventative and management activities. REFRESHER OR ONGOING TRAINING: The WGFs received on-going training and support from CMOs SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Supported by nine staff (CMOs) INCENTIVE: Monetary INCENTIVE DESCRIPTION: The women's group facilitators were given an allowance of 4000K (approx. £13 / US\$22) per month towards the costs of bicycle repairs. N.B. Not a salary RATIO CHW TO POPULATION: N.S. START DATE: 01/06/2007 END DATE: 31/12/2010 CONTROL DESCRIPTION: Control A: Neither intervention, Control B: Facility intervention alone
Outcomes	OUTCOMES: Primary: maternal, perinatal and neonatal mortality. Secondary: Percentage of deliveries at a health facility, number of women's groups mobilized annually, percentage of pregnant women attending women's groups, cost. DURATION OF FOLLOW UP: 2 years
Funding	The Health Foundation, London, UK

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Clusters were allocated to each, both or no intervention with a random number sequence generated in Stata 7"
Allocation concealment	Low Risk	Quote: "To ensure concealment of intervention allocation, identification numbers were assigned for each

(selection bias)		cluster and a random number generated for each...the sequence was concealed until interventions were assigned."
Blinding (performance and detection bias)	High Risk	Quote: "Given the nature of the interventions, neither participants nor those administering the interventions were blinded to group assignment. Those assessing the outcomes were also not blinded to group assignment. However, the analysis plan was pre-specified and detailed in a Stata .do file before the final analysis was carried out."
Incomplete outcome data (attrition bias)	High Risk	Quote: "During the baseline period, the recorded loss to follow-up to birth outcomes was 19% and during the intervention it was 29%, with higher rates in later months...there was little difference in loss-to-follow-up between arms"
Selective reporting (reporting bias)	Low Risk	Comment: Primary outcomes listed in protocol were reported
Recruitment bias	Low Risk	Quote: "A population of approximately 4000 was randomly chosen as follows. We enumerated the size of the catchment populations of health surveillance assistants and then randomly chose HSAs from each cluster until a population of roughly 4000 was gained using the Complex Samples procedure in SPSS 14"
Baseline imbalance	Low Risk	Quote: "Observed differences in baseline mortality, skilled birth attendance, urban, access by tar road, type of facility and tobacco estates were small and the variables were included in adjusted models if associated with the mortality outcome in question. Results of adjusted models were only slightly different to unadjusted models, and are not reported"
Loss of clusters	High Risk	Comment: One cluster not followed up due to lack of a health surveillance assistant
Unit of analysis issues	Low Risk	Quote: "Adjusting for clustering"

Cooper 2009 (incl. Murray 2015)

Trial reg.	ISRCTN25664149
Country	South Africa
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother
Population	LOCAL TERM FOR CHW: Health visitor RECRUITMENT CRITERIA: The home visitors all lived in Khayelitsha and were mothers themselves. Two had completed schooling; none had education or training beyond school. HEALTHCARE SETTING: Household PATIENT POPULATION: Women in the last trimester of their pregnancy (on the basis of the accounts of their gestation that women had received from the antenatal clinic). INTERVENTION SAMPLE SIZE: 220 CONTROL SAMPLE SIZE: 229
Intervention	OBJECTIVE: To assess the efficacy of an intervention designed to improve the mother-infant relationship and security of infant attachment in a South African peri-urban settlement with marked adverse socioeconomic circumstances. INTERVENTION DESCRIPTION: The intervention was manualised. It included key principles of the World Health Organizations 'Improving the Psychosocial Development of Children' and 'The Social Baby.' The intervention was delivered in hour-long sessions. The women in the intervention group were visited in their homes twice antenatally, weekly for the first eight weeks postpartum, fortnightly for a further two months, and then monthly for two months (that is, 16 sessions in total,

	ending at five months post-partum). Previously untrained lay community workers provided support and guidance in parenting. The purpose of the intervention was to promote sensitive and responsive parenting and secure infant attachment to the mother. KEY CHW COMPETENCY: Encourage the mother in sensitive, responsive interactions with her infant. A major aspect was the use of particular items from the neonatal behavioural assessment schedule to sensitise the mother to her infant's individual capacities and needs. TRAINING DURATION (DAYS): 21 (over a four month period) TRAINING DESCRIPTION: The content and conduct of the intervention was specified in a manual, which was used to train the therapists. The aim of the intervention was to encourage the mother in sensitive, responsive interactions with her infant. They received training over a four-month period in basic parenting and counselling skills, as well as in the specific mother-infant intervention. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: An experienced community clinical psychologist provided them with group supervision throughout the study, offering session-by-session supervision. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:55 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Women in the control group received no therapeutic input from the research team. Women in the control group received the normal service provided by the local infant clinic, including fortnightly visits by a community health worker, who assessed the physical and medical progress of mothers and infants, and encouragement of mothers to take their infant to the local clinic to be weighed, have their physical health assessed, and be immunised.
Outcomes	OUTCOMES: The primary outcomes at six and 12 months were those aspects of the mother-infant relationship previously found to be associated with both socioeconomic adversity and maternal depression (that is, maternal sensitivity and intrusiveness); at 18 months the primary outcome was infant attachment. The secondary outcome was maternal depression (a dichotomous variable for depressive disorder and a continuous variable for depressive symptoms) assessed at six and 12 months. Infant cognitive development (Bayley Scales Mental Development Index [MDI]) was also assessed at 18 months. DURATION OF FOLLOW UP: 1.5 years
Funding	Wellcome Trust (grant B574100). MT was supported by a fellowship from the Vlotman Trust.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "we randomly assigned pregnant women to an intervention group or a no treatment control group." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not described
Blinding (performance and detection bias)	Low Risk	Quote: "We made independent assessments, blind to treatment groups"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "By the final assessment at 18 months postpartum, 107 of the 449 women were lost to follow-up, leaving 342 women (76% of the original sample); most of the losses occurred, as anticipated, in the six month period after delivery...The women lost to follow-up did not differ from those retained in terms of demographic characteristics (parity, marital status),

		apart from age; those lost were somewhat younger than those retained (mean 24.9 (SD 4.4) years versus 26.1 (5.8); $t=2.07$, $df=447$, $P<0.05$), and neither did they differ on any of the minimisation criteria or on area of residence.”
Selective reporting (reporting bias)	Low Risk	Comment: All outcomes pre-specified in protocol were reported

Coutinho 2005

Trial reg.	None
Country	Brazil
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother-newborn dyad
Population	LOCAL TERM FOR CHW: Community health agents RECRUITMENT CRITERIA: Personal breastfeeding experience was not a prerequisite. HEALTHCARE SETTING: Household PATIENT POPULATION: Mothers and their newborn children INTERVENTION SAMPLE SIZE: 175 CONTROL SAMPLE SIZE: 175
Intervention	OBJECTIVE: To compare the hospital-based intervention (BFHI training of maternity staff) with a combined hospital-based and community-based intervention (BFHI training and postnatal home visits). INTERVENTION DESCRIPTION: Maternity staff from two hospitals was trained according to the Baby-Friendly Hospital Initiative. Intervention mothers giving birth at these hospitals were assigned ten postnatal home visits to promote and support breastfeeding. CHWs were expected to visit ten times—four times during the first month (on days 3, 7, 15, and 30), every 2 weeks during the second month, and once a month during the third to sixth months. Each mother was to be given the illustrated booklet. At each visit, the home visitors were expected to encourage exclusive breastfeeding for 6 months and continued breastfeeding for at least 2 years, to answer questions and discuss doubts, and to use the booklet as a basis for discussions of key topics relevant to the infant’s age. Whenever possible, they observed the positioning of the infant at the breast, flow of milk, and the baby’s satisfaction; if there were difficulties that they could not resolve, they were expected to refer the mother for more specialist help at the hospital. If other family members were present, their attitude towards exclusive breastfeeding was assessed and their support was sought, including help with household chores. Each visit had a mean duration of 30 min, with the initial visits taking longer than later ones. KEY CHW COMPETENCY: Breastfeeding home visits TRAINING DURATION (DAYS): 8 TRAINING DESCRIPTION: The home visitors received the same 20 h training as the maternity staff plus 5 days in which they studied Helping mothers to breastfeed in depth and practised how to discuss key topics with mothers by use of an illustrated booklet. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:35 START DATE: 03/2001 END DATE: 08/2001 CONTROL DESCRIPTION: No home visits
Outcomes	OUTCOMES: Rates of exclusive breastfeeding from birth to 6 months. DURATION OF FOLLOW UP: 6 months
Funding	Superintendência para o Desenvolvimento do Nordeste, Brasil (Agency for the Development of the Northeast, Brazil)

<i>Risk of Bias</i>		
Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "he random numbers were generated by the project manager,"
Allocation concealment (selection bias)	Low Risk	Quote: "Concealment was achieved by drawing numbers from envelopes at the time of assignment"
Blinding (performance and detection bias)	Low Risk	Quote: "Breastfeeding practices were studied...by researchers unaware of group allocation. "
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analyses were by intention to treat."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified

Darmstadt 2010

Trial reg.	NCT00198627
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Union CLUSTERS: Intervention:6; Control: 6
Population	LOCAL TERM FOR CHW: Community Health Workers RECRUITMENT CRITERIA: Educated to secondary school certification or higher HEALTHCARE SETTING: Household PATIENT POPULATION: Women of reproductive age (15–49). INTERVENTION SAMPLE SIZE: 4616 live births CONTROL SAMPLE SIZE: 5241 live births
Intervention	OBJECTIVE: To evaluate a delivery strategy for newborn interventions in rural Bangladesh. INTERVENTION DESCRIPTION: Community health workers (n=48) identified pregnant women; made two antenatal home visits (scheduled at 12–16 and 32–34 weeks of gestation) to promote BNCP; made four postnatal home visits (on days 2, 5 and 8) to negotiate preventive care practices and to assess newborns for illness; and referred sick neonates to a hospital and facilitated compliance. In the intervention arm, each of the 12 unions had six CHW areas KEY CHW COMPETENCY: Identification of high risk neonates; promotion of early, exclusive breastfeeding, including feeding of colostrum and avoidance of prelacteal feeds; Temperature maintenance, including monitoring axillary temperature; teaching appropriate bathing and skin care practices; clean umbilical cord care, and management of umbilical cord and superficial skin infections; identification and management of neonates with suspected serious bacterial infections. TRAINING DURATION (DAYS): 36 TRAINING DESCRIPTION: CHWs were trained on pregnancy surveillance, counselling and negotiation skills, essential newborn care, neonatal illness surveillance and management of illness based on a clinical algorithm adapted from IMCI REFRESHER OR ONGOING TRAINING: After initial training and evaluation, routine monitoring and refresher training were provided each fortnight. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: Field supervisors used a standard checklist to monitor CHWs' neonatal assessment and management during fieldwork, at scheduled and random evaluations. In each union, one field supervisor oversaw six CHWs by meeting with them for about 6 hours each fortnight to review all data collection forms and instructive case histories and to provide refresher training.

	INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:4000 START DATE: 01/2004 END DATE: 12/2006 CONTROL DESCRIPTION: Newborns in the comparison arm received the usual health services provided by the government, non-governmental organizations and private providers.
Outcomes	OUTCOMES: Primary: neonatal mortality. Secondary: antenatal and immediate newborn care behaviours (Had ≥ 1 antenatal care visits from a qualified provider Received ≥ 1 tetanus immunization, Received ≥ 2 tetanus immunizations, Received iron supplementation, Delivered at medical facilities, Sterile cord cut, Drying/wiping the baby before delivery of placenta, Delaying bath to the 3rd day or later, Breastfeeding initiation within 1 hour after birth, Breastfeeding prior to any food/liquid, Received any treatment Received treatment from a qualified provider) knowledge of danger signs, and care-seeking for neonatal complications. DURATION OF FOLLOW UP: 2 years
Funding	Wellcome Trust - Burroughs Wellcome Fund Infectious Disease Initiative 2000 and the Office of Health, Infectious Diseases and Nutrition, Global Health Bureau, USAID through the Global Research Activity Cooperative agreement with the Johns Hopkins Bloomberg School of Public Health (award HRN-A-00-96-90006-00). Support for data analysis and manuscript preparation was provided by the Saving Newborn Lives programme through a grant by the Bill & Melinda Gates Foundation to Save the Children-US.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Randomly allocated to either comparison or intervention arm using a computer-generated pseudo-random number sequence without stratification or matching"
Allocation concealment (selection bias)	Unclear Risk	Comment: Not stated.
Blinding (performance and detection bias)	High Risk	Quote: "Blinding was unachievable given the nature of the intervention"
Incomplete outcome data (attrition bias)	Low Risk	Comment: Attrition was 3.8% and 4.2% from intervention and control arm respectively
Selective reporting (reporting bias)	High Risk	Comment: Reported outcomes that were not pre-specified and did not report all outcomes that were (e.g. effectiveness of the community transport system)
Recruitment bias	Unclear Risk	Comment: Not stated.
Baseline imbalance	Low Risk	Quote: "Indicators of maternal and newborn care practice and knowledge were similar in the intervention and comparison arms at baseline. "
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of analysis issues	Low Risk	Quote: "Robust standard errors were adjusted for clustering on each union."

Datiko 2009 (incl. Datiko 2010)

Trial reg.	NCT00803322
Country	Ethiopia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Kebeles CLUSTERS: Intervention: 30; Control: 21
Population	LOCAL TERM FOR CHW: Health extension workers (HEWs) RECRUITMENT CRITERIA: N.S.

	HEALTHCARE SETTING: Health post (operational unit for HEWs) PATIENT POPULATION: TB suspects, who had cough for two weeks or more, were referred for further investigations. A smear-positive pulmonary TB case was defined by two positive sputum smears or one positive smear and x-ray findings consistent with active TB. INTERVENTION SAMPLE SIZE: 230 CONTROL SAMPLE SIZE: 88
Intervention	OBJECTIVE: Assess whether involving HEWs in TB control improved smear-positive case detection and treatment success rates in southern Ethiopia. INTERVENTION DESCRIPTION: HEWs in the intervention kebeles identified suspects, collected sputum, and provided directly observed treatment. The HEWs in the intervention kebeles advised people with productive cough of 2 weeks or more duration to attend the health posts. During health education sessions at health posts, HEWs informed people living in the kebele about TB and advised them to come to a health post if they had productive cough of 2 weeks or more duration. TB suspects who came to the health posts were told about community-based TB care. HEWs collected spot-morning-spot sputum specimens, and labelled and transported them to the diagnostic units every month for examination for acid-fast bacilli by direct microscopy. Smear-positive patients in the intervention kebeles received standard DOTS under the direct observation of HEWs. TB patients visited health posts daily during the intensive phase and once a month in the continuation phase. KEY CHW COMPETENCY: Identify suspects, collect sputum, and provide directly observed treatment. TRAINING DURATION (DAYS): 365 (prior to study); 2 (as part of study) TRAINING DESCRIPTION: The training focused on symptoms and transmission of TB, how to identify TB suspects, how to collect, label, store and transport sputum specimens, administer DOT, and follow patients during treatment. The messages and the content of our training were similar to the curriculum of training HEWs. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: HEWs received on job training about how to collect sputum samples and support patients to adhere to treatment. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Salary from the government RATIO CHW TO POPULATION: 1 health post: 5000 people START DATE: 09/2006 END DATE: 04/2008 CONTROL DESCRIPTION: HEWs in the in the control kebeles did not received on job training about how to collect sputum samples and how to support patients to adhere to treatment. However, they provided health services, including health education about TB, to the people living in their kebeles.
Outcomes	OUTCOMES: Case detection rate, treatment success rate, costs. DURATION OF FOLLOW UP: 7 months
Funding	The University of Bergen

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote "Using a table of random numbers"
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "Neither the general health workers nor TB programme managers were blinded to the allocation." Comment: Low risk for sputum specimens
Incomplete outcome data (attrition bias)	Low Risk	Quote: "All the smear-positive patients were analyzed"

Selective reporting (reporting bias)	Low Risk	Comment: Outcomes specified in protocol were reported.
Recruitment bias	Low Risk	Quote: "All the smear-positive patients [in intervention and control areas] were analyzed"
Baseline imbalance	Low Risk	Quote: "The baseline demographic and clinical characteristics were similar in the two groups."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "weighted by cluster size"

Dias 2008

Trial reg.	NCT00479271
Country	India
Methods	STUDY DESIGN: RCT (crossover) UNIT OF RANDOMISATION: person with dementia-caregiver dyad
Population	LOCAL TERM FOR CHW: Home Care Advisors (HCA) RECRUITMENT CRITERIA: Knowledge of the local language, being literate, preferably passed higher secondary school, and motivated to be involved in the community care of older people. HEALTHCARE SETTING: Household PATIENT POPULATION: CDR mild and moderate dementia. Exclusion criteria were: CDR severe dementia or severe co-morbid physical health conditions. The principal caregiver, as identified by the family, was enrolled for the trial. INTERVENTION SAMPLE SIZE: 41 CONTROL SAMPLE SIZE: 40
Intervention	OBJECTIVE: To develop and evaluate the effectiveness of a home based intervention in reducing caregiver burden, promoting caregiver mental health and reducing behavioural problems in elderly persons with dementia. INTERVENTION DESCRIPTION: Community based intervention provided by a team consisting of HCA who were supervised by a counsellor and a psychiatrist, focusing on supporting the caregiver through information on dementia, guidance on behaviour management, a single psychiatric assessment and psychotropic medication if needed. KEY CHW COMPETENCY: Basic education about dementia (what is the disease, its course, its features etc.); education about common behaviour problems and how they can be managed; support to the caregiver, for example for an elderly caregiver living alone with the patient, in activities of daily living; referral to psychiatrists or the family doctor when behaviour problems are severe and warrant medication intervention; networking of families to enable the formation of support groups; advice regarding existing government schemes for elders TRAINING DURATION (DAYS): 7 TRAINING DESCRIPTION: The HCA were trained in key skills including listening and counselling skills, bereavement counselling, stress management and health advice for common health problems REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: The HCA were supported, and supervised, by the two part-time specialists: two psychiatrists (one supporting each team) and one counsellor (supporting both teams). They shared experiences, supported one another, and problem solved difficult situations. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 10/2003 END DATE: 02/2005 CONTROL DESCRIPTION: The control arm dyads received only education and information regarding dementia and were then placed in a waiting list to receive the intervention after 6 months. They were free to utilize the existing health services

	during this time.
Outcomes	OUTCOMES: Primary: caregiver mental health (GHQ score). Secondary: perceived burden (Zarit Burden score), distress due to problem behaviours (NPI-D) and severity of the behavioural problems in the person with dementia (NPI-S), and functional ability of the subject (EASI). DURATION OF FOLLOW UP: 6 months
Funding	World Health Organization

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "by an independent person, based on simple random number tables"
Allocation concealment (selection bias)	Low Risk	Comment: The allocation was done by an 'independent person'
Blinding (performance and detection bias)	Low Risk	Quote: "Outcome evaluations were masked to the allocation status."
Incomplete outcome data (attrition bias)	Low Risk	Quote: " There was no significant difference in the baseline characteristics of those who died or were alive till the end of the trial"
Selective reporting (reporting bias)	High Risk	Comment: The protocol mentions "costs of illness" as being an outcome but it is not reported

Eloff 2014

Trial reg.	None
Country	South Africa
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother-child dyad
Population	LOCAL TERM FOR CHW: Community care workers RECRUITMENT CRITERIA: Individuals who exhibited good interpersonal skills and had at least 12 years of education HEALTHCARE SETTING: Community Group PATIENT POPULATION: HIV-positive women and their children, aged 6–10 years. INTERVENTION SAMPLE SIZE: 199 CONTROL SAMPLE SIZE: 191
Intervention	OBJECTIVE: Assess the efficacy of an intervention designed to promote resilience in young children living with their HIV-positive mothers INTERVENTION DESCRIPTION: The intervention consisted of 24 weekly group sessions led by community care workers. Mothers and children were in separate groups for 14 sessions, followed by 10 interactive sessions. The primary focus was on parent–child communication and parenting. The study included 12 groups, with approximately 15 participants in each group. Participants were reimbursed for travel expenses and meals were provided. KEY CHW COMPETENCY: Lead weekly group sessions TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: The training included information about HIV and AIDS and skills for facilitating groups, counselling and identification and management of children's emotional and behavioural problems. A manual identified specific objectives for each session and provided an outline for experiential learning activities such as games, group discussions and behavioural modelling. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Supervised by a social worker. INCENTIVE: N.S.

	INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Individuals randomised to the standard care condition were provided with information about local resources available for assistance.
Outcomes	OUTCOMES: Maternal psychological and parenting characteristics, child psychological and behavioural functioning DURATION OF FOLLOW UP: 18 months
Funding	National Institute of Mental Health (Grant Number: 5R01MH076442)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "...using a block randomisation process and table of random numbers with 30 individuals in each block"
Allocation concealment (selection bias)	Unclear Risk	Comment: Not described
Blinding (performance and detection bias)	Unclear Risk	Comment: Not described
Incomplete outcome data (attrition bias)	High Risk	Quote: "161 (81%) mothers in the intervention condition and 169 (88%) mothers in the standard care condition completed at least one follow-up interview and were included in the analysis.. Women lost to follow-up were less likely to be living with a partner (29 versus 50%, P 1/4 0.03) and had known of their HIV status for a shorter period of time (median 9 versus 13 months, P 1/4 0.013)." Comment: Those lost to follow-up were excluded from the analysis
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned, 16 variables reported

Ertl 2011

Trial reg.	NCT00552006
Country	Uganda
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Participant
Population	LOCAL TERM FOR CHW: Lay counsellors RECRUITMENT CRITERIA: Community-based lay therapists without a mental health or medical background. HEALTHCARE SETTING: Household/camp PATIENT POPULATION: Former Northern Ugandan child soldiers aged 12 to 25 years with PTSD INTERVENTION SAMPLE SIZE: 29 CONTROL SAMPLE SIZE: (A) 28, (B) 28
Intervention	OBJECTIVE: To assess the efficacy of a community-based intervention targeting symptoms of PTSD in formerly abducted individuals. INTERVENTION DESCRIPTION: Provide narrative exposure therapy (n = 29). KEY CHW COMPETENCY: Provide narrative exposure therapy TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Treatment fidelity and therapeutic competence were monitored by case discussions in supervision meetings

	INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 11/2007 END DATE: 10/2009 CONTROL DESCRIPTION: (A) Waiting list (n=28) (B) CHWs provide academic catch-up programme (n=28). This arm was included as a control as, in line with the protocol, CHWs were not trained to provide the intervention in question.
Outcomes	OUTCOMES: Primary: Change in PTSD severity. Secondary: depression symptoms, severity of suicidal ideation, feelings of guilt, and perceived stigmatization. DURATION OF FOLLOW UP: 1 year
Funding	Vivo and by the Deutsche Forschungsgemeinschaft and the Ein Herz für Kinder foundation.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "participants were randomised" Comment: no additional details provided
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	Low Risk	Quote: "Pre-treatment assessments as well as follow-up assessments at 3 months, 6 months, and 12 months after treatment were conducted by 13 clinical psychologists blinded to treatment conditions."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Analysis by intention to treat
Selective reporting (reporting bias)	Low Risk	Comment: All outcomes reported in trial registration were reported

Fottrell 2013 (incl. Clarke 2014)

Trial reg.	ISRCTN01805825
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Union CLUSTERS: Intervention: 9, Control: 9
Population	LOCAL TERM FOR CHW: Facilitator RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Community PATIENT POPULATION: Ever-married women of reproductive age INTERVENTION SAMPLE SIZE: 9106 CONTROL SAMPLE SIZE: 8602
Intervention	OBJECTIVE: Assess the effect of a participatory women's group intervention with higher population coverage on neonatal mortality in Bangladesh. INTERVENTION DESCRIPTION: Women's groups at a coverage of 1 per 309 population that proceed through a participatory learning and action cycle in which they prioritize issues that affected maternal and neonatal health and design and implement strategies to address these issues. KEY CHW COMPETENCY: Facilitate a learning action cycle TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S.

	RATIO CHW TO POPULATION: N.S. START DATE: 07/2009 END DATE: 06/2011 CONTROL DESCRIPTION: All unions, irrespective of allocation arm, received health system– strengthening initiatives from the Diabetic Association of Bangladesh–Perinatal Care Project. These included the provision of basic medical equipment, traditional birth attendant training in essential newborn care, physician training, and establishing links between communities and health services.
Outcomes	OUTCOMES: Primary: Neonatal mortality. Secondary: Early and late neonatal mortality rate, Stillbirth rate, Pregnancy-related mortality ratio and maternal mortality ratio, Health care use, Home care practices, Postpartum psychological distress (SRQ-20). DURATION OF FOLLOW UP: 2 years (except SRQ-20, collected for four months each year for three years)
Funding	Big Lottery Fund International Strategic Grant, Wellcome Trust Strategic Award

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Were randomly allocated to either intervention or control by blindly pulling pieces of paper, each representing 1 union, from a bottle. The allocation sequence had been decided on before drawing the papers."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "...the implementation and in-country monitoring and evaluation teams, were blind to the allocation arm"
Incomplete outcome data (attrition bias)	Unclear Risk	Quote: "the intention-to-treat analysis only included those who permanently resided in the study areas, as determined by the registration process. Analysis of secondary outcomes only included women who were successfully interviewed."
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes reported protocol were reported
Recruitment bias	Low Risk	Quote: "Participants were assigned to the cluster in which their birth or death was registered"
Baseline imbalance	Low Risk	Quote: "Socio-demographic indicators and primary and secondary out- come indicators were similar in both arms,"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: " Adjusted for clustering"

Fylkesnes 2013

Trial reg.	ISRCTN53353725
Country	Zambia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Villages CLUSTERS: Intervention: 18; Control: 18
Population	LOCAL TERM FOR CHW: Counsellor RECRUITMENT CRITERIA: The counsellors (four women and four men) were selected among already trained local lay counsellors. These lay health workers were individuals selected by the local communities for voluntary work at the clinics. HEALTHCARE SETTING: Home PATIENT POPULATION: Men and women aged 16 years or above INTERVENTION SAMPLE SIZE: 836 CONTROL SAMPLE SIZE: 858
Intervention	OBJECTIVE: Investigate the acceptance of HB-VCT, equity in uptake and negative life events

	INTERVENTION DESCRIPTION: Adults in the intervention clusters were offered HB-VCT by local lay counsellors. In the intervention arm, lay counsellors offered home-based VCT to all adult household members. This consisted of pre-test counselling, HIV rapid testing using the finger-prick method, and post-test counselling, all provided in the home during one session (one hour average duration). The participants were offered counselling in privacy, but couples were encouraged to be counselled and to disclose together. The counsellors used bicycles for transport and home visits were repeated until contact was achieved and attempts were made on weekdays and weekends. KEY CHW COMPETENCY: Provide HB-VCT TRAINING DURATION (DAYS): 10 TRAINING DESCRIPTION: The counsellors received supplementary training on general counselling principles, HIV testing, the principles on which the HB-VCT model was grounded, and their applications when approaching households. The training was carried out in workshops REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: Quality assurance of the counselling was addressed by counsellors from the district's VCT team and the District Medical Officer. Two counsellors were responsible for regular supervision, and two counselling sessions were observed for each counsellor. The counsellors and supervisors met with the scientific coordinator once a week for reporting, clarifications and open discussions based on experiences INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 03/2010 END DATE: 01/2011 CONTROL DESCRIPTION: Both arms had access to standard HIV testing services. Both the intervention and control communities received standard care, comprising clinic-based VCT and routine HIV counselling and testing connected to antenatal care (both offered by the Government), and outreach VCT offered occasionally by NGOs.
Outcomes	OUTCOMES: Acceptance of counselling and testing, equity in uptake and negative life events following VCT. Acceptance of HB-VCT was measured in two ways; i) acceptance of counselling and ii) acceptance of testing. DURATION OF FOLLOW UP: 6 months (from end of intervention)
Funding	Norwegian Programme for Development, Research and Education, the Research Council of Norway, and the Swedische Norwegian Regional, HIV/AIDS Team for Africa.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Randomisation was performed by a person who was blinded to the identity of the clusters by randomly selecting one cluster from each pair with computer."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "Another limitation was the lack of "blinding" of research assistants"
Incomplete outcome data (attrition bias)	Low Risk	Comment: Intention-to-treat analysis
Selective reporting (reporting bias)	Low Risk	Comment: Primary outcome reported in study protocol reported. Secondary outcome "changes in sexual risk behaviour" not reported
Recruitment bias	Low Risk	Comment: All people in study communities were screened for participation
Baseline imbalance	Low Risk	Quote: "Few baseline differences were identified between

		the intervention and control clusters with regard to the most relevant socio-demographic characteristics and previous HIV testing (Table 1). The only statistically significant difference was a lower proportion of single individuals in the intervention communities reporting having been tested (13% vs. 23%, p 1/4 0.010)."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Unclear Risk	Comment: Not mentioned

Gardner 2003 (incl. Walker 2004; Walker 2010)

Trial reg.	None
Country	Jamaica
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: LBW-term infant
Population	LOCAL TERM FOR CHW: Community health workers RECRUITMENT CRITERIA: Paraprofessionals with at least 4 years of secondary education HEALTHCARE SETTING: Home PATIENT POPULATION: LBW-term infants (weight <2500g) INTERVENTION SAMPLE SIZE: 66 CONTROL SAMPLE SIZE: 69
Intervention	OBJECTIVE: To determine whether early psychosocial intervention with LBW-term infants improved cognition and behaviour and to compare LBW-term with NBW infants. INTERVENTION DESCRIPTION: The intervention comprised weekly home visits by paraprofessionals for the first 8 weeks of life aimed at improving maternal-child interaction. KEY CHW COMPETENCY: Perform psychosocial stimulation intervention TRAINING DURATION (DAYS): 14 TRAINING DESCRIPTION: Training in child development and the intervention. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:20 START DATE: 03/1999 END DATE: N.S. CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Problem solving/cognition (2 means-end tests: cover and support) and 4 behaviour ratings. Six year follow-up: Cognition (Preschool and Primary Scale of Intelligence, 3rd edition (WPPSI-III); Picture Vocabulary Test (PPVT)); Memory (Corsi blocks test); Attention (Test of Everyday Attention for Children); Early reading ability (Early Reading Assessment); Behaviour (Strengths and Difficulties Questionnaire (SDQ)); Stimulation in the home (Middle Childhood Home Observation for the Measurement of the Environment (MC-HOME); Anthropometry (height-for-age and body mass index z-scores). DURATION OF FOLLOW UP: 7, 15, 24 months. 6 years
Funding	Nestlé Foundation

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	High Risk	Quote: "systematic random assignment"
Allocation concealment	Unclear Risk	Comment: Not specified

(selection bias)		
Blinding (performance and detection bias)	Low Risk	Quote: "All videotapes were scored by one of the testers. Another observer who was also unaware of the children's group scored a random 25% of the tests. The reliabilities were high, with intraclass correlations of 0.98 for the total intention scores for both sets of problems."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Approximately 10% of the children did not complete the fifth trial in at least one test because they became distressed or sleepy (a similar percentage in each group)." Comment: attrition low
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified

Gökçay 1993

Trial reg.	None
Country	Turkey
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: "Visiting areas" (Map of Istanbul was divided based on population) CLUSTERS: Intervention: 12, Control: 12
Population	LOCAL TERM FOR CHW: Lady home visitors (LHV) RECRUITMENT CRITERIA: All LHV's were primary school graduates HEALTHCARE SETTING: Household PATIENT POPULATION: Women and children under 5 years INTERVENTION SAMPLE SIZE: 873 CONTROL SAMPLE SIZE: 713
Intervention	OBJECTIVE: Compare the performance of midwives with that of LHV INTERVENTION DESCRIPTION: carries out to identify and offer guidance on family planning, antenatal care well-child control, immunizations and other mother and child health related issues to those at high risk KEY CHW COMPETENCY: (a) Identify those at risk, namely those who were not attending any health institutions and/or those in health need; (b) inviting these families to the clinic and (c) offering guidance on family planning, antenatal care, well-child control, immunizations and other mother and child health related problems TRAINING DURATION (DAYS): 21 TRAINING DESCRIPTION: Training on MCH issues, communication skills, and the tasks involved in the home visits REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Monthly salary (1/2 that of facility-based midwives) RATIO CHW TO POPULATION: 1:291 START DATE: 06/1989 END DATE: 06/1991 CONTROL DESCRIPTION: Home visits by midwives
Outcomes	OUTCOMES: Numbers of pregnant women identified in the last trimester, pregnant women receiving antenatal care, deliveries attended by a health care provider, modern family planning users, continuation rate of the new users, pregnancy rate of the new users, immunization status of children, well-child visits of the children DURATION OF FOLLOW UP: 1 year
Funding	Pathfinder International Foundation

Risk of Bias

Item	Reviewers'	Support for judgement
------	------------	-----------------------

	judgement	
Random sequence generation (selection bias)	Unclear Risk	Quote: "One midwife or one LHV was randomly assigned to each visiting unit"
Allocation concealment (selection bias)	Unclear Risk	Comment: Not stated
Blinding (performance and detection bias)	Unclear Risk	Comment: Not stated
Incomplete outcome data (attrition bias)	High Risk	Comment: Attrition rate not clear but reported that 497 households not evaluated as moved away from project area. Lay visitors followed up 473 households initially.
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified
Recruitment bias	High risk	Comment: Participants were recruited post randomisation
Baseline imbalance	Low Risk	Quote: "The areas resembled each other in such things as family size, use of clinics, etc."
Loss of clusters	Low Risk	Comment: no clusters lost
Unit of analysis issues	Unclear Risk	Comment: Not mentioned

Grantham-McGregor 1991 (incl. Chang 2002, Chang 2010, Gertler 2014, Grantham-McGregor 1993, Grantham-McGregor 1997, Walker 1991, Walker 1996, Walker 2000, Walker 2005, Walker 2006, Walker 2011)

Trial reg.	None
Country	Jamaica
Methods	STUDY DESIGN: RCT (2x2 factorial) UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Community health aide RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: N.S. PATIENT POPULATION: Growth-retarded (stunted) children aged 9-24 months INTERVENTION SAMPLE SIZE: Intervention A: 32, Intervention B: 32, Intervention C: 32 CONTROL SAMPLE SIZE: 33
Intervention	OBJECTIVE: The effects of nutritional supplementation, with or without psychosocial stimulation of growth-retarded (stunted) children aged 9-24 months were assessed INTERVENTION DESCRIPTION: (A) Supplemented only, (B) Stimulated only, (C) Supplemented and stimulated. The supplement comprised 1 kg milk-based formula per week for 2 years, and the stimulation weekly play sessions at home with a community health aide. KEY CHW COMPETENCY: Supplementation: Provide 1 kg milk-based formula per week. Stimulation: The community health aides visited the homes for 1 h per week and taught the mothers how to play with their children in such a way as to promote their development. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: A supervisor monitored 10% of all visits. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Free medical care was available to all children. The children in the four stunted groups were visited every week by a community health

	aide, and a history of the previous week's illnesses was recorded. Visits to the control group were intended to allow for any extra attention that the supplemented children might receive. Visiting was used instead of a placebo because it was considered unethical for health service personnel to give a true placebo (i.e. a low-calorie supplement) to undernourished children.
Outcomes	OUTCOMES: 12 months: length, weight, head circumference, MUAC, triceps, and subscapular skinfold thickness, WHZ. 2 years: Child development (Griffiths mental development scale). 4 years: Stanford Binet, Ravens matrices, tests of memory and language, grooved pegboard, WRAT, height and weight WAZ, HAZ, Arm circumference, Triceps skinfold, Subscapular skinfold, Cognitive function, school achievement, and fine motor tests. 8 years: WISC-R, Ravens Matrices, language, memory, attention, WRAT, Rutter parent and teacher scales, cognition, behaviour at home, behaviour at school, and school achievement. 15 years: WAIS IQ, Ravens Matrices, memory and language, vocabulary (PPVT), reading test, WRAT mathematics, depression, anxiety, self-esteem, antisocial behaviour (self-report), attention deficit, oppositional behaviour (Conners parent report). 18 years: WAIS IQ, WRAT—expanded form, grade level attained, examination passed, general knowledge, depression, anxiety, social inhibition, antisocial behaviour (self-report). DURATION OF FOLLOW UP: 12 months, 2 years, 5 years, 8 years, 12 years, 15 years, 18years, 20 years
Funding	Ford Foundation, USA, with assistance from the Population Council. S. G. McG. is partially supported by the Wellcome Trust, UK. The supplement was donated by Cow & Gate, UK, Grace Kennedy & Co Ltd, Jamaica Commodity Trading Co, and Seprod Ltd, Jamaica.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "The initial order of group assignment was determined randomly. For example, every fourth child was assigned to the supplemented and stimulated group."
Allocation concealment (selection bias)	Unclear Risk	Comment: Not stated
Blinding (performance and detection bias)	Low Risk	Quote: "All tests were done by one of two testers, who were trained by the same person, and who did not know the children's group assignment." Comment: Unclear for height and weight
Incomplete outcome data (attrition bias)	Low Risk	Quote: "2 children were lost from the stimulated group, and were omitted from the analyses. All other children had all five developmental assessments" Comment: ~2% attrition
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned

Haider 2000

Trial reg.	None
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Zones CLUSTERS: Intervention: 20; Control: 20
Population	LOCAL TERM FOR CHW: Peer counsellors RECRUITMENT CRITERIA: Women with personal breastfeeding experience, at least 4 years schooling, willingness to help other mothers breastfeed HEALTHCARE SETTING: Household PATIENT POPULATION: Women (aged 16–35 years) were enrolled during the last trimester of pregnancy. Women with documented records of heart disease, insulin-dependent diabetes mellitus, or eclampsia in a previous pregnancy, were not included. Mother-infant pairs were excluded in cases of multiple births, congenital anomalies, admission for intensive care, and birth weight below 1.8 kg.

	INTERVENTION SAMPLE SIZE: 363 CONTROL SAMPLE SIZE: 363
Intervention	OBJECTIVE: Asses whether the intervention of trained peer counsellors, mothers could be enabled to breastfeed exclusively for the recommended duration of 5 months. INTERVENTION DESCRIPTION: In the intervention group, 15 home-based counselling visits were scheduled, with two visits in the last trimester, three early postpartum (within 48 h, on day 5, between days 10 and 14), and fortnightly thereafter until the infant was 5 months old. KEY CHW COMPETENCY: Provide breastfeeding counselling TRAINING DURATION (DAYS): 5 (4 h daily for 10 days) TRAINING DESCRIPTION: Parts of the WHO/UNICEF breastfeeding counselling course and King's book were used as guides. Counselling skills were taught by demonstrations and role play and included listening to mothers, learning about their difficulties, assessing the position and attachment of babies during a breastfeed, building mothers' confidence, giving support, and providing relevant information and practical help when required. During a pilot of 2 weeks, prenatal and postpartum counselling were practised with pregnant mothers, and with mothers of newborn babies and infants aged 1–5 months, respectively. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Quarterly SUPERVISION DESCRIPTION: Their performance was monitored at least thrice over the total study period by two breastfeeding supervisors based at International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh INCENTIVE: Monetary INCENTIVE DESCRIPTION: The peer counsellors worked part time and were paid an honorarium (£16 [US\$22.50] per month). RATIO CHW TO POPULATION: 1:25 START DATE: 02/1996 END DATE: 05/1997 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Primary: prevalence of exclusive breastfeeding at 5 months. Secondary: time taken to initiate breastfeeding and proportion of mothers who fed colostrum or prelacteals after delivery. DURATION OF FOLLOW UP: 5 months
Funding	Swiss Agency for Development and Cooperation (SDC; 95–98 grant no t.311-BD 25).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Random number tables were used to select 40 of these zones and to allocate them into two groups"
Allocation concealment (selection bias)	Low Risk	Comment: Not mentioned
Blinding (performance and detection bias)	High Risk	Quote: "[Interviewers] were aware of which zones had peer counsellors and therefore of group assignment."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "There were no significant differences in the socioeconomic and demographic characteristics of the losses compared with those who completed 5 months' follow-up."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified
Recruitment bias	Low Risk	Quote: "Four field assistants identified the women who were in their last trimester of pregnancy by going door-to-door"
Baseline imbalance	Low Risk	Quote: "Households were similar in the intervention and control groups; about two-thirds of families lived in a

		partly concrete house with one room, and 95% had electricity. All the families had access to safe drinking water and almost all used slab latrines. Median family income was 4000 (range 0–90 000) Taka per month (53 Taka=US\$1.0) in the control area and 4900 (0–150 000) Taka in the intervention area. Most mothers were housewives, and there were no differences in maternal characteristics between the groups "
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of analysis issues	Low Risk	Quote: "Results are based on individual level analyses but with adjustment for the cluster level of randomisation of the intervention."

Jafar 2009 (incl. Jafar 2011, Jafar 2010, Jafar 2015)

Trial reg.	NCT00327574
Country	Pakistan
Methods	STUDY DESIGN: C-RCT (2x2 factorial) UNIT OF RANDOMISATION: 12 of 4200 low- to middle-income, geographic census-based clusters (mean household monthly income, \$70; about 250 households in each cluster) in Karachi, the most populous city in Pakistan (about 16 million inhabitants). CLUSTERS: Intervention A: 3; Intervention B: 3; Control A: 3, Control B: 3
Population	LOCAL TERM FOR CHW: Lay health worker, Community health worker RECRUITMENT CRITERIA: The education status of the workers employed was consistent with the requirements of the government-sponsored Lady Health Workers Programme of Pakistan (8 or preferably 10 years of schooling) HEALTHCARE SETTING: Household/clinic PATIENT POPULATION: Patients 40 years or older with hypertension (systolic blood pressure ≥ 140 mm Hg, diastolic blood pressure ≥ 90 mm Hg, or already receiving treatment). INTERVENTION SAMPLE SIZE: Intervention A: 348; Intervention B: 332 CONTROL SAMPLE SIZE: Control A: 335; Control B: 326
Intervention	OBJECTIVE: To assess the effectiveness of 2 community-based interventions on blood pressure in hypertensive adults. INTERVENTION DESCRIPTION: Family-based home health education (HHE) from trained community health workers, to improve population-level health literacy every 3 months and annual training of general practitioners (GPs) in hypertension management. The effect of these interventions were tested, alone and in combination, on blood pressure in adults with hypertension. (A) HHE only; (B) HHE + GP. The first HHE session, lasting 90 minutes, was held at a time when all members of the household could be present. Follow-up reinforcement visits of 30 minutes were made every 3 months. KEY CHW COMPETENCY: Delivering health messages TRAINING DURATION (DAYS): 43 TRAINING DESCRIPTION: Training in behaviour-changing communication strategies to convey standardized health education messages. The health messages included information on the deleterious effects of hypertension and nonpharmacologic interventions for preventing and controlling hypertension and cardiovascular disease, as well as advice on the importance of engaging in moderate physical activity; maintaining normal body weight; reducing salt intake; consuming a diet rich in fruit, vegetables, and low-fat dairy products; reducing intake of saturated and total fat (including suggestions on sample recipes for culturally acceptable and economically feasible food products); and smoking cessation REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Quarterly SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: Salary scales and assigned workload were similarly

	consistent with the requirements of the government-sponsored Lady Health Workers Programme of Pakistan RATIO CHW TO POPULATION: 1:110 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: (A) GP only; (B) No intervention
Outcomes	OUTCOMES: Primary: change in blood pressure from baseline to the last follow-up visit, all-cause mortality. Secondary: proportion of participants with controlled blood pressure (<140 mm Hg for systolic blood pressure and <90 mm Hg for diastolic blood pressure) at the last follow-up visit, BMI, waist hip ratio, per cent current tobacco users, total physical activity, LDL-cholesterol, cost. DURATION OF FOLLOW UP: 2 years + 5 year follow-up
Funding	Wellcome trust

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "We used computer-generated codes to randomly assign 3 clusters each to the following groups: HHE alone, GP alone, HHE and GP combined, and no intervention."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "The assessors were blinded to the randomisation status of each cluster...the key outcome (blood pressure) was measured by using an automated device, which minimized the risk for biased outcome assessment."
Incomplete outcome data (attrition bias)	High Risk	Quote: "Follow-up blood pressure measurements were missing for 22% of patients...However, these missing visits were balanced among intervention groups." Comment: Intention to treat analysis
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in the protocol were reported
Recruitment bias	Low Risk	Quote: "All households in each cluster were visited, and we obtained informed consent for screening from all adults 40 years and older, whose blood pressure was then measured 3 times with a calibrated automated device in the sitting position after 5 minutes of rest. Those with known hypertension were invited to participate."
Baseline imbalance	Low Risk	Quote: "Age, sex, and baseline blood pressure-adjusted... blood pressure and percentage of patients with controlled hypertension did not differ among groups"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "In the initial analysis, we investigated clustering of participants at the family level. Because this was not significant, we only accounted for clustering by household at the census level as a random effect in the main analysis; this was also the unit of randomisation..."

Jaffar 2009 (incl. Amuron 2011; Lugada 2010; Nakigozi 2015; Woodd 2014)

Trial reg.	ISRCTN 17184129
Country	Uganda
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Geographical areas along or, in the case of a few

	large subdistricts, by known barriers within the subdistrict CLUSTERS: Intervention: 22; Control: 22
Population	LOCAL TERM FOR CHW: Field officer RECRUITMENT CRITERIA: Most field officers had degree qualifications or college diplomas HEALTHCARE SETTING: Household PATIENT POPULATION: All patients with WHO stage IV or late stage III disease or CD4-cell counts fewer than 200 cells per μL who started antiretroviral therapy between Feb 15, 2005, and Dec 19, 2006, were eligible, apart from those living on islands. INTERVENTION SAMPLE SIZE: 859 CONTROL SAMPLE SIZE: 594
Intervention	OBJECTIVE: Assess whether home-based HIV care was as effective as was facility-based care. INTERVENTION DESCRIPTION: For the home-based group, trained field officers travelling on motorcycles visited patients at home every month to deliver drugs, monitor participants with a checklist that included signs and symptoms of drug toxicity or disease progression, and provide adherence support. KEY CHW COMPETENCY: Deliver drugs, monitor participants with a checklist that included signs and symptoms of drug toxicity or disease progression, and provide adherence support. TRAINING DURATION (DAYS): 25 TRAINING DESCRIPTION: About principles of antiretroviral therapy and adherence support. REFRESHER OR ONGOING TRAINING: Yearly refresher courses SUPERVISION FREQUENCY: Daily SUPERVISION DESCRIPTION: At the end of every day, a medical officer reviewed the notes made by field officers and, when needed, asked officers to return to the patient's home to refer them. INCENTIVE: None INCENTIVE DESCRIPTION: No financial or other incentives were provided to patients or staff RATIO CHW TO POPULATION: N.S. START DATE: 15/02/2006 END DATE: 31/01/2009 CONTROL DESCRIPTION: In the facility-based group, patients obtained drugs every month from the clinic and had routine reviews with a medical officer and counsellor that were scheduled at 2 and 3 months after start of treatment and every 3 months thereafter.
Outcomes	OUTCOMES: Primary: rate of virological failure, defined as time (starting from 6 months) to a plasma RNA viral load of more than 500 copies per mL. Secondary: time to either detectable plasma viral load of more than 500 copies per mL at any visit from 12 months onwards in patients with viral loads of fewer than 500 copies per mL at 6 months, or an increase of 1000 copies per mL between two consecutive tests in those not achieving a viral load of fewer than 500 copies per mL at 6 months. Other secondary outcomes were all-cause mortality; virological failure as defined in the primary outcome or death; time to first admission; death, admission, or change to second-line antiretroviral therapy; outpatient attendance; adherence during the previous 28 days (measured with a standardised questionnaire); and costs incurred by the health service and patients, VCT uptake and HIV prevalence, adherence to clinic visits, HIV disease progression, and self-reported conduct of daily life. Sensitivity analysis conducted on outcomes for people who presented with low CD4 ⁺ cell counts (<50 cells/ml) and those who presented with higher CD4 ⁺ cell count DURATION OF FOLLOW UP: 2 years
Funding	US Centers for Disease Control and Prevention and UK Medical Research Council.

Risk of Bias

Item	Reviewers'	Support for judgement
------	------------	-----------------------

	judgement	
Random sequence generation (selection bias)	Low Risk	Quote: "were randomised to home-based or facility-based care by drawing sealed cards from a box."
Allocation concealment (selection bias)	Low Risk	Quote: "The cards were sealed in advance and labelled with the stratum number by the trial coordinator and TASO senior medical officer who organised the allocation event and placed the cards into the box for each stratum. Cards were drawn by two patient representatives, a TASO medical officer, TASO counsellor, and TASO field officer, each taking turns. This process was done in the presence of senior TASO staff, researchers, and local public health representatives."
Blinding (performance and detection bias)	Unclear Risk	Quote: "Independent research staff assessed adherence...Clinical data were transcribed from patient notes."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analyses were by intention to treat "
Selective reporting (reporting bias)	Low Risk	Quote: Outcomes listed in trial registration were reported
Recruitment bias	Low Risk	Quote: "participants were identified after they were tested for CD4-cell counts and clinically assessed"
Baseline imbalance	Low Risk	Quote: "44 geographical areas in nine strata, defined according to ratio of urban and rural participants and distance from the clinic... The two groups were well balanced according to baseline characteristics apart from CD4-cell count, which was lower in the home-based than in the facility-based group (table 1)...Analyses were...adjusted for baseline CD4-cell count and study stratum"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "The model was fitted with the assumption that the random cluster effects followed a γ distribution."

Kidane 2000

Trial reg.	None
Country	Ethiopia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Tabias (cluster of villages) CLUSTERS: Intervention: 12; Control: 12
Population	LOCAL TERM FOR CHW: Mother coordinators RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Children under 5 years INTERVENTION SAMPLE SIZE: 6383 CONTROL SAMPLE SIZE: 7294
Intervention	OBJECTIVE: Assess the effect on under-5 mortality of teaching mothers to promptly provide antimalarials to their sick children at home INTERVENTION DESCRIPTION: In the intervention tabias, mother coordinators were trained to teach other local mothers to recognise symptoms of malaria in their children and to promptly give chloroquine. The mother coordinators were supplied with chloroquine for distribution to all households and were responsible for reporting the use to the tabia coordinator and replenishing the drugs used. KEY CHW COMPETENCY: Recognise symptoms of malaria in their children and to promptly give chloroquine. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: In the intervention tabias mother coordinators taught

	neighbour-group mothers to recognise symptoms in their under-5 children that might be a result of malaria, to give the appropriate course of chloroquine for their age, to share chloroquine properly, and to recognise possible adverse reactions from the drug. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: In each tabia one mother coordinator was chosen to coordinate all other mother coordinators in the tabia. These coordinators collected the monthly reports on births, deaths, migration in and out of the community, and referrals, and checked whether drugs were short and reported any problems. Seven field supervisors from the Tigray Community-Based Malaria Control Programme were appointed to supervise the tabia coordinators through four to six visits per supervisor each month, and to directly supervise a sample of the mothers by visiting at least five of them per day. The supervisors along with mother coordinators and neighbourhood mothers verified births and deaths. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1: 583 START DATE: 01/01/1997 END DATE: 31/12/1997 CONTROL DESCRIPTION: In all tabias a list of children for every mother coordinator was prepared in a monthly report format. Mother coordinators were taught to keep track of and record, in this monthly format, all births and deaths, and were taught where to refer sick children.
Outcomes	OUTCOMES: Under-5 mortality DURATION OF FOLLOW UP: 12 months
Funding	UNDP/World Bank/WHO TDR

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "One tabia from each pair was allocated by random number to an intervention group and the other was allocated to the control group. "
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Comment: Blinding not possible but outcome was mortality and so unlikely to be influenced
Incomplete outcome data (attrition bias)	Low Risk	Comment: Intention to treat
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Unclear Risk	Comment: All <5yrs included.
Baseline imbalance	Unclear Risk	Comment: The average baseline mortality in the control tabias was 47.6 per 1000, compared to 60.8 per 1000 in the intervention tabias, and this difference is not statistically significant. Other baseline characteristics were not presented.
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Unclear Risk	Comment: Not stated.

Kirkwood 2013 (incl. Manu 2012, Pitt 2016, Vesel 2013)

Trial reg.	NCT00623337
Country	Ghana
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Zone CLUSTERS: Interventions: 49; Control: 49
Population	LOCAL TERM FOR CHW: Community-based surveillance volunteers (CBSV)

	RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: The trial included all pregnancies that ended in a live birth or stillbirth in the study period INTERVENTION SAMPLE SIZE: 9435 CONTROL SAMPLE SIZE: 9174
Intervention	OBJECTIVE: Test a home-visit strategy in sub-Saharan Africa by assessing the effect on all-cause neonatal mortality rate (NMR) and essential newborn-care practices. INTERVENTION DESCRIPTION: CBSVs in Newhints zones were trained to identify pregnant women in their community and to make two home visits during pregnancy and three in the first week of life to promote essential newborn-care practices, weigh and assess babies for danger signs, and refer as necessary. KEY CHW COMPETENCY: Identify pregnant women in their community and to make two home visits during pregnancy and three in the first week of life to promote essential newborn-care practices, weigh and assess babies for danger signs, and refer as necessary. TRAINING DURATION (DAYS): 9 TRAINING DESCRIPTION: CBSVs underwent training to identify pregnant women and newborn babies, and to counsel and solve problems relating to key essential newborn-care behaviours, weigh newborn babies, check them for danger signs, and if necessary refer them. REFRESHER OR ONGOING TRAINING: 2 days refresher with a focus on the newborn assessment procedures SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Monthly individual sessions with observation of home visit and supportive feedback; and group supervision meetings every 2 months INCENTIVE: Monetary and non-monetary INCENTIVE DESCRIPTION: GHC5 (=5USD) per month, polo shirt, and briefcase RATIO CHW TO POPULATION: START DATE: 01/11/2008 END DATE: 31/12/2009 CONTROL DESCRIPTION: Pregnant women and newborn babies living in the control zones continued to benefit from the routine maternal and child health care available, which consisted of antenatal clinics, access to free facility delivery, post-partum check-ups, infant welfare clinics, and routine CBSV activities for outreach maternal and child health and immunisation clinics. They also benefited from the hospital essential newborn-care strengthening and sensitisation activities that covered all health facilities in the trial area.
Outcomes	OUTCOMES: Primary: all-cause NMR, the percentages of mothers practising the Newhints recommended behaviours (Bed net in pregnancy, Saved money for delivery or emergency, Arranged transport to facility in advance, Baby delivered in a facility, Birth assistant washed hands with soap (home delivery), Early initiation of breastfeeding (<1 h of birth), Skin to skin contact (any and >2h), Delayed first bath (>6 h), Exclusive breastfeeding (26–32 days), Baby sleeping under bed net (8–56 days) Care-seeking, sick babies taken to hospital or clinic). Secondary outcomes were age- specific and cause-specific NMRs; the most important was the NMR after day 1 (days 2–28), Access to Care, Compliance with Referral, Cost. DURATION OF FOLLOW UP: 1 year
Funding	WHO, Save the Children's Saving Newborn Lives Programme from the Bill & Melinda Gates Foundation, and the UK Department for International Development

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Computer-generated restricted randomisation"

Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation; done by an independent epidemiologist
Blinding (performance and detection bias)	Low Risk	Quote "The surveillance consisted of home visits to all women of reproductive age (15-45 years) every 4 weeks by an independent group of resident research field workers"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analyses were by intention to treat"
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in protocol were reported
Recruitment bias	Low Risk	Comment: All births in study area were included
Baseline imbalance	Low Risk	Quote: "Computer-generated restricted randomisation was then done in a one-to-one ratio...using stratified sampling to ensure balance within districts and the four large towns; it was restricted to ensure comparability between intervention and control zones with respect to NMR (difference of less than two per 1000 live births), percentage of deliveries in a health facility (<2.5%), and percentage of deliveries in a private facility (<2.5%) by use of available surveillance data in each of the 3 years before the trial planning started (2004–06)."
Loss of clusters	Low Risk	Comment: No clusters lost.
Unit of analysis issues	Low Risk	Quote: "Random-effects logistic regression was used to account for the cluster-randomised design, with relative risks (RR) derived by use of the marginal standardisation technique, and the 95% CIs estimated with the delta method."

Kumar 2008 (incl. Kumar 2012; Willis 2012)

Trial reg.	NCT00198653
Country	India
Methods	STUDY DESIGN: C-RCT (3-arm) UNIT OF RANDOMISATION: Gram sabha (The basic geopolitical and administrative unit for village-level health planning and implementation) CLUSTERS: Intervention A: 13; Intervention B: 13; Control: 13
Population	LOCAL TERM FOR CHW: Saksham Sahayak RECRUITMENT CRITERIA: Recruited from the local community based on 12 years or more of education, proficient communication and reasoning skills, commitment towards community work, and references of community stakeholders. HEALTHCARE SETTING: Household PATIENT POPULATION: Pregnant women INTERVENTION SAMPLE SIZE: Intervention A: 1600; Intervention B: 1149 CONTROL SAMPLE SIZE: 1141
Intervention	OBJECTIVE: Test an intervention of behaviour change management, with a focus on prevention of hypothermia, aimed at modifying practices and reducing neonatal mortality. INTERVENTION DESCRIPTION: Two intervention groups: (A) received a preventive package of interventions for essential newborn care (birth preparedness, clean delivery and cord care, thermal care [including skin-to-skin care], breastfeeding promotion, and danger sign recognition) or (B) which received the package of essential newborn care plus use of a liquid crystal hypothermia indicator. In the intervention clusters, community health workers delivered the packages via collective meetings and two antenatal and two postnatal household visitations KEY CHW COMPETENCY: Deliver messages about newborn-care practices TRAINING DURATION (DAYS): 7 TRAINING DESCRIPTION: They received a combination of classroom-based and apprenticeship-based field training over 7 days on knowledge, attitudes, and practices related to essential newborn care within the community, behaviour change

	management, and trust-building. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Performance assessment of Saksham Sahayaks included feedback from community members, spot checks by their supervisors during home visits and community meetings to assess their level of community engagement, and monitoring by the supervisors of whether targets for home visits and community meetings were being met. INCENTIVE: Monetary INCENTIVE DESCRIPTION: US\$35–40 per month RATIO CHW TO POPULATION: N.S. START DATE: 01/2004 END DATE: 05/2005 CONTROL DESCRIPTION: Received the usual services of governmental and non-governmental organisations in the area
Outcomes	OUTCOMES: Changes in newborn-care practices (Place of delivery: Home/Health facility/On the way, Routine antenatal care check-up (≥ 1), Tetanus toxoid vaccination (≥ 2), Skilled birth attendant, Delivery in hands, Wiping of whole body, Bathing within 24 h, Skin-to-skin care, Cord cut with clean blade, Breastfeeding within 1 h of birth), and neonatal, stillbirth and perinatal mortality rate, compared with the control group, knowledge of maternal danger signs, self-reported complications), Maternal care practices (Routine prenatal care, Tetanus toxoid vaccination, Self-recognition of problems faced during pregnancy, Discussed problems with family members/neighbours), care-seeking from trained providers, maternal mortality ratio (MMR), perceived neonatal morbidity. DURATION OF FOLLOW UP: 16 months
Funding	USAID and Save the Children-US through a grant from the Bill & Melinda Gates Foundation.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Stratified cluster randomisation was done at Johns Hopkins University using Stata 7.0 (StataCorp, College Station, TX, USA) to allocate the 39 cluster units randomly to the three study groups, yielding three allocation sequences of 13 clusters each. "
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Unclear Risk	Comment: Not clear if the data collectors were different from those delivering the intervention.
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Primary analysis was undertaken as intention to treat at cluster level. "
Selective reporting (reporting bias)	Unclear Risk	Comment: Outcome measures not provided in protocol
Recruitment bias	Low Risk	Quote: "As part of the baseline survey, all pregnant women in the study area were identified...All live births were followed-up through the infant period, and all deaths were recorded."
Baseline imbalance	Low Risk	Quote: "There were more low-caste households in the essential newborn care arm than in the control arm. Since there is evidence that caste could be associated with neonatal mortality, we did an adjusted analysis which produced results similar to the unadjusted analysis."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "To account for clustering, point estimates for stillbirth rates, neonatal mortality rates, and perinatal mortality rates for each study arm were calculated as the mean of cluster event rates, giving an equal weight to each cluster"

Kweku 2009 (incl. Patouillard 2011)

Trial reg.	NCT00119132
Country	Ghana
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: village CLUSTERS: Intervention: 6; Control: 6
Population	LOCAL TERM FOR CHW: Community-based volunteers RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: All children aged 3–59 months resident in the 12 study villages INTERVENTION SAMPLE SIZE: 472 CONTROL SAMPLE SIZE: 490
Intervention	OBJECTIVE: Evaluate the coverage of IPTc that can be achieved by two different delivery systems in Ghana. INTERVENTION DESCRIPTION: Drugs were administered in May, June, September and October 2006 in both arms. In the intervention arm, IPTc was delivered by volunteers KEY CHW COMPETENCY: Administer doses of AQ plus SP TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Community volunteers received a payment of \$10 RATIO CHW TO POPULATION: N.S. START DATE: 05/2006 END DATE: 10/2006 CONTROL DESCRIPTION: IPTc was delivered by health workers at health centres or at Expanded Programme on Immunisation outreach clinics
Outcomes	OUTCOMES: Primary: coverage with IPTc. Secondary: Adherence to treatment, the incidence of clinical malaria, the incidence of anaemia, the prevalence of malaria parasitaemia at the beginning and at the end of the rainy season, the incidence of adverse events, cost. DURATION OF FOLLOW UP: 6 months
Funding	Funded by Gates Malaria Partnership based at the London School of Hygiene and Tropical Medicine.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "villages were assigned...by ballot."
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned
Blinding (performance and detection bias)	Unclear Risk	Comment: Not mentioned
Incomplete outcome data (attrition bias)	Low Risk	Quote: "For the analysis of coverage of IPTc, an intention to treat analysis was used in which all children who were enrolled and received the first dose of IPTc course one were included in the analysis."
Selective reporting (reporting bias)	High Risk	Comment: Primary outcome in protocol was "Mean Hb at the end of the high transmission season."
Recruitment bias	Low Risk	Quote: "All children aged 3–59 months (1039) resident in the 12 study villages were then enumerated in May

		2006. Children whose parents gave written consent (964) were enrolled into the study after a physical examination."
Baseline imbalance	Low Risk	Quote: "At enrolment, children in the two study groups were similar in regard to age, gender, anthropometric indices and malaria parasite prevalence"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Clustering adjusted for using logistic regression with robust standard errors."

le Roux 2010

Trial reg.	None
Country	South Africa
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother-child dyad
Population	LOCAL TERM FOR CHW: Mentor mother (MM) RECRUITMENT CRITERIA: Nominees were interviewed by supervisors for the Philani program. MMs chosen for the programme had children who were thriving; they demonstrated strong communication and interpersonal skills, commitment to community service and showed an organized and disciplined approach to tasks. HEALTHCARE SETTING: Household PATIENT POPULATION: Mother-child dyads where there was an underweight child. INTERVENTION SAMPLE SIZE: 536 CONTROL SAMPLE SIZE: 252
Intervention	OBJECTIVE: Philani, a paraprofessional home visiting program, was developed to improve childhood nourishment. The objective of this study is to evaluate whether the Philani programme can rehabilitate malnourished children in a timely manner. INTERVENTION DESCRIPTION: MMs visited dyads in the intervention condition for one year, supporting mothers' problem solving around nutrition. All children were weighed by MMs at baseline and three, six, nine and twelve month follow-ups. KEY CHW COMPETENCY: A typical MM visit at each home lasted 20 to 60 minutes. During the visits, the MM weighed the participating child and discussed his progress with the mother. The MM also made sure the mother had the social grants she might be entitled to, and that the mother understood proper nutrition and hygiene. MMs stressed the importance of breastfeeding, the proper time to introduce solids, frequent feeding, and a mixed diet including vegetables and fruit. She checked to see if immunizations were up to date and that the child was dewormed. In each MM's caseload there was likely to be one emergency a week (e.g., a child would be ill with high fever, have difficulty breathing, or appear severely dehydrated). These cases were brought to the Philani Health Clinic or the local public health clinic to receive immediate attention. MMs did not distribute food supplements. TRAINING DURATION (DAYS): 30 TRAINING DESCRIPTION: MMs who were selected received training in child health, nutrition and other related topics. Mentor Mothers received four phases of training: 1) watching experienced MMs implement the intervention in an inspiring manner, learning how to approach a family and build trust; 2) attending a month of training that covered nutrition; basic child health including HIV and TB, weighing of babies and completion of growth charts; how to recognize signs of abuse and crisis situations; and how to encourage depressed mothers to be more active and engaged with their children; 3) learning how to help mothers bond with their children and improve the consistency of healthy daily routines; and 4) implementing their first round of home visits independently in their neighbourhoods. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: A supervisor accompanied each MM at least one

	day a month on a random schedule to ensure that implementation proceeded as planned. The supervisor collaborated with the MM in problem-solving and generating action plans when problems occurred in the field. The quality of the implementation was monitored by reviewing the forms completed at each home visit, monitoring visitation patterns, collecting observations by supervisors, and brief ratings of home visits by the supervisors. INCENTIVE: Monetary INCENTIVE DESCRIPTION: MMs received a stipend of \$US 130/month from Philani to deliver home-based interventions and were expected to work four hours a day. RATIO CHW TO POPULATION: N.S. START DATE: 2006 END DATE: 2007 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Rehabilitation of low-weight babies and children DURATION OF FOLLOW UP: 12 months
Funding	World Childhood Foundation

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	High Risk	Quote: "Assignment to the intervention condition was based on a sequence decided a priori for every three dyads."
Allocation concealment (selection bias)	High Risk	Quote: "The MMs were given randomly sequenced numbered folders marked I for intervention and C for controls and supervisors made sure the folders were allocated in the correct order...Although the randomisation sequence in this study was determined a priori by the research team, the protocol was administered by MMs. The fact that children in the intervention arm of the study were significantly younger and more underweight indicates that the MMs may not have adhered to the assignments, but rather steered needier children into the immediate intervention. The direction of any bias caused by this deviation from random assignment is likely to be against the detection of an intervention effect: to the extent that children in the intervention had lower WAZ scores, they had further to go to achieve normative weights. Regardless, future research should use procedures that cannot be tampered with, such as neighbourhood rather than individual assignment."
Blinding (performance and detection bias)	Unclear Risk	Comment: Not enough information.
Incomplete outcome data (attrition bias)	High Risk	Quote: "We compared characteristics from Table 1 between dyads we were unable to follow after recruitment (10%; 82 of 788) vs. dyads with at least one follow-up (90%; 706/788). Dyads without follow-up were more likely to have older children (25.2 vs. 17.8 months, $t = 4.08$, $P < .01$), less likely to have the child supported by a nutrition programme (10% vs. 20%, $c2 = 4.80$, $P = .03$), more likely to have a mother who was employed (18% vs. 9%, $c2 = 6.11$, $P = .01$), and less likely to have a flush toilet versus less modernized toilet facilities (44% vs. 57%, $c2 = 4.36$, $P = .04$). Dyads in the control condition were more likely to miss follow-ups compared to those in the intervention condition (20% vs. 6%; $c2 = 35.38$, $P < .01$)."
Selective reporting	Unclear Risk	Comment: No protocol indicated

(reporting bias)		
le Roux 2011		
Trial reg.	NCT00995592	
Country	South Africa	
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother-child dyad	
Population	LOCAL TERM FOR CHW: Mentor Mother (MM) RECRUITMENT CRITERIA: One MM was recruited from each of 37 neighbourhoods in Cape Town, South Africa. Criteria included having thriving babies, demonstrating good communication and strong interpersonal skills, being committed to community service and showing a disciplined personal and professional lifestyle. HEALTHCARE SETTING: Household PATIENT POPULATION: Mothers with malnourished children under the age of six years (dyad). Malnourishment was defined as having a weight more than two standard deviations below the World Health Organization (WHO) normative mean weight for the child's age. Children in this range are in the bottom 2% in weight for age. INTERVENTION SAMPLE SIZE: 500 CONTROL SAMPLE SIZE: 179	
Intervention	OBJECTIVE: Philani, a paraprofessional home visiting program, was developed to improve nutritional outcomes for young children in South Africa. INTERVENTION DESCRIPTION: MMs provided the intervention home visits with growth monitoring to those dyads randomised to receive it. The frequency of MM visits was based upon need. Typical MM home visits lasted from 20 minutes to one hour. During the visits, the MM weighed the child and discussed developmental progress with the mother. The MM also ensured that the mother had applied for appropriate social grants and understood proper nutrition and hygiene. MMs stressed the importance of breastfeeding, the proper time to introduce solid food, frequent feeding and a mixed diet that includes fruits and vegetables. The MMs checked that immunizations were up to date and that the child had been dewormed. KEY CHW COMPETENCY: MMs were trained to conduct home visits to weigh children under six years old and to support mothers to problem-solve life challenges, especially around nutrition. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: MMs received four phases of training: (1) watching experienced MMs implement the intervention in an inspiring manner; (2) attending training sessions covering nutrition, basic child health, weighing babies and completion of growth charts, how to recognize danger signs and crisis situations and how to encourage depressed mothers to be more active and engaged with their children; (3) learning how to build trust with mothers and use the relationship to improve the consistency of healthy daily routines; and (4) implementing the first round of home visits independently in their neighbourhoods REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: The intervention supervisor visited at least one day per month on a random schedule to ensure that the implementation was proceeding as planned. The supervisor collaborated with the MM in problem-solving and generating action plans when problems occurred in the field. The quality of implementation was monitored by reviewing the forms completed at each home visit, monitoring visitation patterns, collecting observations by outreach supervisors and brief ratings of home visits by the outreach supervisors. INCENTIVE: Monetary INCENTIVE DESCRIPTION: They received a stipend of \$US 130/month and were expected to work for four hours per day. RATIO CHW TO POPULATION: START DATE: 07/2004	

	END DATE: 09/2005 CONTROL DESCRIPTION: Dyads in the control arm of the study did not receive these visits.
Outcomes	OUTCOMES: Weight-for-age Z-scores (WAZ) DURATION OF FOLLOW UP: 1 year
Funding	Centre for Health and Wellbeing through Professor Christina Paxon at the Woodrow Wilson School, Princeton University

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	High Risk	Quote: "Assignment to the treatment condition was based on a random sequence decided a priori for every three households enrolled into the study."
Allocation concealment (selection bias)	High Risk	Quote: "MMs were provided with randomly sequenced numbered folders containing the randomisation assignments. Once a dyad was enrolled the MM selected the next folder in her possession, and assigned the dyad accordingly...MMs implemented assignment of families to the treatment condition; the nearly significant difference in child weight at baseline between the intervention and control groups indicates that reassignment of needier families into the intervention probably took place, and prevents this from being a truly randomised study."
Blinding (performance and detection bias)	High Risk	Quote: "Because MMs conducted the intervention as well as collected data, assessments were not blinded as to study arm."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "We compared characteristics from Table 1 between mothers we were unable to follow after recruitment (6%; 39 of 679) and mothers with at least one follow-up assessment. Mothers in the control condition were more likely to be lost to follow-up compared to those in the intervention (11% vs. 4%; $\chi^2 = 13.24$, $df = 1$, $p < 0.01$). Mothers lost to follow-up were also likely to have lived in Cape Town for fewer years compared to mothers who were followed successfully (mean = 6.1 vs. 10.3; $t = -2.46$, $df = 415$, $p = 0.01$). None of the other demographic or background characteristics differed significantly by whether or not mothers were lost to follow-up. All observations were included in regression analyses; dyads lost to follow-up contributed only to weight estimation at the recruitment timepoint."
Selective reporting (reporting bias)	Unclear Risk	Quote: "Analyses on weight in kilograms and WAZ led to the same conclusions; therefore, we present results on WAZ only." Comment: Protocol notes WAZ as primary outcome and weight in kilograms as secondary outcome

Leite 2005

Trial reg.	None
Country	Brazil
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Mother and baby dyad
Population	LOCAL TERM FOR CHW: Lay counsellors RECRUITMENT CRITERIA: These lay counsellors have personal experience in breastfeeding their own children and comprise of a workgroup that is associated with the "Assis Chateaubriand Maternity School" Milk Bank at the Federal University of Ceará (MEAC-UFC) for more than five years.

	HEALTHCARE SETTING: Household PATIENT POPULATION: (a) Mothers who had given birth to unfavourably low weight babies (less than 3000 g); (b) healthy newborn babies with an expected discharge from the maternity facility, as determined by the nursery's medical team, within 5 d of birth; (c) mothers living in Fortaleza who had no plans to change residence within the next 6 months. INTERVENTION SAMPLE SIZE: 503 CONTROL SAMPLE SIZE: 500
Intervention	OBJECTIVE: To evaluate the effectiveness of home-based peer counselling to increase breastfeeding rates for unfavourably low birth weight babies. INTERVENTION DESCRIPTION: Breastfeeding counselling, conducted by lay counsellors from the community, during home visits carried out on days 5, 15, 30, 60, 90 and 120 after birth. KEY CHW COMPETENCY: a) Interviews with the mother; (b) observation of the home environment; (c) observation of all aspects involved with the breastfeeding, including technical ones, as well as the mother–child relationship; (d) identification of the difficulties faced by the mother during breastfeeding and the type of counselling provided. Each visit lasted, on average, 30–40 min. The counselling activity emphasized the following aspects: (a) careful listening to the mother; (b) correct positioning of the “grip”; (c) prevention and treatment of nipple fissures, and the teaching of breast and nipple exercises and techniques for milk expression; (d) instruction regarding breastfeeding; (e) counselling on use of pacifiers, bottles and bottle teats, as well as on the use of teas and water; (f) creation of a social and family support network. TRAINING DURATION (DAYS): TRAINING DESCRIPTION: Besides the periodic training and supervision activities, all members of this group have participated in a 20-h theory-practice course adapted from “Breastfeeding counselling: a training course” REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Each supervising mother received, as payment, approximately US\$4 for each home visit, which totalled about US\$100–120 per month, the equivalent of a minimum wage in Brazil at the time (1996–1997). RATIO CHW TO POPULATION: 1:25 START DATE: 11/1996 END DATE: 04/1997 CONTROL DESCRIPTION: The control group received no specific intervention. They were instructed in the maternity ward, by the interviewer, to seek out their local health service facility in case of any health problems.
Outcomes	OUTCOMES: Feeding methods in the fourth month of life. DURATION OF FOLLOW UP: 4 months
Funding	National Institute of Food and Nutrition (Ministry of Health), the Public Health School of Ceará and by the Division of Child Health and Development of the World Health Organization (WHO).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	High Risk	Quote: "Groups were formed consisting of 20 pairs (mother and baby), using a table of computer-generated random numbers"
Allocation concealment (selection bias)	Unclear Risk	Quote: "The interviewer handed the completed questionnaire to the study secretary, who then designated each mother to one of the study groups after opening a sealed envelope that contained the respective group code." Comment: Unclear if envelopes were opaque
Blinding (performance)	Low Risk	Quote: "The interviewers had not had any prior contact

and detection bias)		with the mothers and were also unaware as to the objectives of the research (blinding)"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "No difference was found in the variables studied for the pairs that had dropped out" Comment: Attrition low (15%)
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol indicated

Lewycka 2013

Trial reg.	ISRCTN06477126
Country	Malawi
Methods	STUDY DESIGN: C-RCT (2x2 factorial) UNIT OF RANDOMISATION: Clusters of villages. Clusters were defined on the basis of census enumeration areas CLUSTERS: Intervention A: 12, Intervention B: 12, Intervention C: 12; Control: 12
Population	LOCAL TERM FOR CHW: Volunteer peer counsellors/cluster facilitator RECRUITMENT CRITERIA: Counsellors were selected with communities, and were literate women aged 23–50 years with breastfeeding experience. Facilitators were local women aged 20–49 years, literate and with at least one child, selected from the community. HEALTHCARE SETTING: Household/Community PATIENT POPULATION: All women aged 10–49 years who consented to participate were enrolled INTERVENTION SAMPLE SIZE: A: 13678, B: 13683, c: 14022 CONTROL SAMPLE SIZE: 14548
Intervention	OBJECTIVE: Women's groups and health education by peer counsellors can improve the health of mothers and children. The study assessed their effects on mortality and breastfeeding rates in rural Malawi. INTERVENTION DESCRIPTION: Intervention A: Women's group and volunteer peer counselling; Intervention B: women's group only; Intervention C: volunteer peer counselling only. Women's group facilitators guided groups through a community action cycle to tackle maternal and child health problems. Counsellors identified pregnant women and made five home visits during and after pregnancy: in the third trimester, in the week after birth, and at 1 month, 3 months, and 5 months to support breastfeeding and infant care. KEY CHW COMPETENCY: Counsellors provided health education about exclusive breastfeeding, infant care, immunisations, prevention of MTCT, and family planning. They also supported women with breast problems and raised awareness of timely care-seeking. Through the four phases of the facilitated action cycle, members identified and prioritised maternal and child health problems, identified strategies to implement, planned and implemented them, and assessed them and made plans for the future. TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: Counsellors used an intervention manual describing visit content, and a simple picture book (adapted from manuals published by WHO, Save the Children, and Linkages). They were given an initial and annual refresher training, and attended monthly meetings. REFRESHER OR ONGOING TRAINING: Annual SUPERVISION FREQUENCY: Quarterly SUPERVISION DESCRIPTION: Counsellors were supervised by 24 government health surveillance assistants, who visited quarterly, documenting progress, achievements, and challenges, and by three MaiMwana officers. Health surveillance assistants received initial and yearly refresher training, attended monthly and quarterly meetings, and were given calendars and a supervisory checklist. INCENTIVE: Non monetary INCENTIVE DESCRIPTION: They were also given a bicycle, meeting allowances, registers, calendars, and supervision forms. RATIO CHW TO POPULATION:

	START DATE: 01/07/2005 END DATE: 31/01/2009 CONTROL DESCRIPTION: 12 clusters were allocated no intervention
Outcomes	OUTCOMES: Primary outcomes for the women's group intervention were maternal, perinatal, neonatal, and infant mortality rates (MMR, PMR, NMR, and IMR, respectively); and for the peer counselling were IMR and exclusive breastfeeding (EBF) rates. Secondary outcomes for the women's group intervention were maternal and infant morbidity, skilled antenatal, delivery, and postnatal care, tetanus toxoid immunisation, use of malaria prophylaxis, insecticide-treated bed nets during pregnancy, and PMTCT services, infant immunisations, early EBF, and reduced use of prelacteal feeds. For the volunteer peer counselling intervention, neonatal mortality and infant morbidity rates were secondary outcomes, and caretaker practices included duration of EBF, time to initiating breast-feeding, use of prelacteal feeds, time to weaning, management of breast problems, and family planning uptake, including condom use. DURATION OF FOLLOW UP: 12 months
Funding	Saving Newborn Lives with later contributions from the UK Department for International Development, Wellcome Trust, Institute of Child Health, and UNICEF Malawi.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "48 equal-sized clusters were randomly allocated to four groups with a computer-generated number sequence"
Allocation concealment (selection bias)	Low Risk	Quote: "Allocated clusters with a random number sequence generated in Stata"
Blinding (performance and detection bias)	Low Risk	Quote: "Allocation was masked for data analysis. Data were gathered independently of programme implementation and could not be used to inform the implementation of the interventions, and thus bias them."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat"
Selective reporting (reporting bias)	Unclear Risk	Comment: Outcomes in protocol reported, but protocol only released after data collection had finished.
Recruitment bias	Low Risk	Quote: "Women's migrations within and outside the study area were recorded to allow for both intention-to-treat and per-protocol analyses" Comment: All women in a cluster offered enrolment
Baseline imbalance	Unclear Risk	Quote: "Women in the four groups were similar in age, education, and marital status, with small differences in religious and tribal affiliations between groups, fewer farmers in areas with volunteer peer counselling only, and more primigravida in areas with women's group intervention only."
Loss of clusters	Low Risk	Comment: No clusters were lost.
Unit of analysis issues	Low Risk	Comment: Adjusted for clustering

Lubega 2015

Trial reg.	None.
Country	Uganda
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Community support agents RECRUITMENT CRITERIA: Influential community volunteers or experienced ART clients; minimum of primary education HEALTHCARE SETTING: Household PATIENT POPULATION: Newly screened (WHO stage 1 or 2) non-ART-eligible

	HIV-positive adult (aged >18 years) INTERVENTION SAMPLE SIZE: 200 CONTROL SAMPLE SIZE: 200
Intervention	OBJECTIVE: Assess the effect of providing home counselling visits by community support agents on 2-year retention in pre-ARV care INTERVENTION DESCRIPTION: Posttest counselling and monthly home counselling visits by community support agents to encourage them go back for routine pre-ARV care. At each visit, the CSA evaluated the impressions of the PLHIV on the last pre-ARV clinic visit service, counselled them on any negative perceptions on the service, and encouraged them to go back again on the appointment date. The CSA would also feedback the client's feelings about the service to the appropriate facility for the staffs' attention KEY CHW COMPETENCY: Posttest counselling TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: CSAs received orientation about this study and additional skills in counselling PLHIV to be retained under HIV care REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: CSAs are supervised by the "in charges" of their nearest health centre INCENTIVE: Non-monetary INCENTIVE DESCRIPTION: CSAs are purely volunteers and do not receive any salaries or other emoluments from the district. They are provided with a new bicycle purchased at US \$48 and an annual bicycle maintenance fee of US \$24 per agent. RATIO CHW TO POPULATION: 1:10 START DATE: After 12/2008 END DATE: N.S. CONTROL DESCRIPTION: Posttest counselling alone
Outcomes	OUTCOMES: Primary: Proportion of new PLHIV in either arm who attended their scheduled pre-ARV care visits for at least 6 of the anticipated 8 visits in the first 24 months after HIV diagnosis. Secondary: HIV status disclosure and indicators of safe sexual behaviour DURATION OF FOLLOW UP: 2 years
Funding	Sida and the European Commission through ARVMAC-INCO-DEV FP6 program and also in part (S.J.R.) from the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, and National Institutes of Health.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "computer-generated random number sequences"
Allocation concealment (selection bias)	Low Risk	Quote: "Assignment sequences were placed in opaque sealed envelopes and appropriate numbers sent to each of the recruitment centres. At each recruitment centre, participant enrolment was done by an assigned laboratory technician at the HIV/STI clinic who carried out the HIV testing on clients that wanted voluntary HIV counselling and testing. After a client had tested HIV positive, the technician at the centre asked the eligible client to pick and open an envelope from a batch of 4 for assignment to the appropriate arm. "
Blinding (performance and detection bias)	Unclear Risk	Quote: "Follow-up data were collected by a trained research assistant at each of the agreed participants nearest HC using interviewer-administered questionnaires" Comment: Unclear if they were blinded

Incomplete outcome data (attrition bias)	Low Risk	Comment: intention-to-treat analysis
Selective reporting (reporting bias)	Low Risk	Comment: Primary outcome same as that in trial registration

Lwilla 2003

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: TB treatment unit CLUSTERS: Intervention: 8; Control: 8
Population	LOCAL TERM FOR CHW: Community-based DOT (CBDOT) observers RECRUITMENT CRITERIA: The CBDOT observers who volunteered to work were interviewed and selected by the village leaders HEALTHCARE SETTING: Household PATIENT POPULATION: Patients with sputum-positive TB were recruited INTERVENTION SAMPLE SIZE: 221 CONTROL SAMPLE SIZE: 301
Intervention	OBJECTIVE: Compare CBDOT using a short-course drug regimen with institutional-based DOT (IBDOT) INTERVENTION DESCRIPTION: CBDOT using a short-course drug regimen. During house visits, the health workers monitored adherence to treatment by checking the treatment card and counting pills remaining from the patient's monthly drug supply. During the whole course of treatment, patients under CBDOT had to visit the health facility monthly for drug collection. In the continuation treatment phase of 6 months, patients from both approaches had to swallow the medications daily under self-supervision. Moreover, on completion of 2, 5 and 7 months of treatment, patients from both arms were asked to visit the diagnostic centres for a sputum check. KEY CHW COMPETENCY: Supervise the patient's intake of the TB medication TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Trained by the DTLC to supervise the patient's intake of the TB medications during the intensive treatment phase. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: During this phase, the CBDOT observers were visited fortnightly by the TB health worker in charge of the nearby health facility and once per month by the DTLC. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/1999 END DATE: 06/2000 CONTROL DESCRIPTION: Institutional-based DOT (IBDOT). In the intensive treatment phase (first 2 months), of the standard IBDOT, the patient had to visit the health facility daily to be observed while swallowing the medications.
Outcomes	OUTCOMES: Primary: conversion rate, defined as the proportion with AFB-negative sputum after 2 months of treatment. Secondary: cure rate at 7 months, defined by AFB-negative sputum at 2 months, remaining negative at month 5 and/or at month 7 after the start of treatment. DURATION OF FOLLOW UP: 7 months
Funding	The Tanzanian Health Users Research Fund, the Spanish Agency for International Co-operation (AECI), Fons Catalan and La Caixa

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection)	Unclear Risk	Quote: "Each pair was given a unique number (from 1 to 9), and one treatment unit within each pair was randomly

bias)		assigned to IBDOT or CBDOT." Comment: Unclear how randomisation sequence was generated.
Allocation concealment (selection bias)	Unclear Risk	Quote: "Some selection bias may have resulted in an overestimate of the benefits of CBDOT. SFDDH was the treatment facility with the single largest number of TB cases, with more than twice the number of the next largest treatment centre. SFDDH was also the designated district hospital and is likely to have attracted the more sick patients, with the highest risk of adverse outcome. Because FDDH was randomised to deliver care through IBDOT, this may well have biased the results in favour of CBDOT. For this reason, we are reluctant to place too great an emphasis on the differential mortality rates, although it was encouraging that exclusion of the SFDDH-containing cluster suggested that CBDOT may still afford a benefit in terms of reduced mortality. "
Blinding (performance and detection bias)	Low Risk	Comment: No blinding but outcome measures unlikely to be influenced by lack of blinding
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was performed by intention to treat and sensitivity analyses performed to evaluate the potential effect of missing responses."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned.
Recruitment bias	Low Risk	Comment: All patients received an offer to join the trial
Baseline imbalance	Low Risk	Quote: "The 18 TB treatment units were matched in pairs on the basis of available data on type of health facility and number of TB patients registered in the past year. "
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Endpoints were calculated for each of the eight paired clusters and the method described by Mantel and Haenszel (1959) used to obtain a pooled estimate for the completeness of follow-up, conversion rate and cure rate. The probabilities of converting or being cured were also compared between the groups using regression methods, the standard error of the coefficients for clustering on treatment units being corrected using the Huber-White-Sandwich method"

Manandhar 2004 (incl. Borghi 2005; Wade 2006)

Trial reg.	ISRCTN31137309
Country	Nepal
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village development committee CLUSTERS: Intervention: 12; Control: 12
Population	LOCAL TERM FOR CHW: Facilitator RECRUITMENT CRITERIA: Literate locally resident woman. Shortlists for this role were derived from nomination by community leaders, advertisement, and word of mouth, after which all potential candidates were interviewed. HEALTHCARE SETTING: Community PATIENT POPULATION: Married women of reproductive age. Inclusion criteria were: consent given for involvement; age 15–49 years; married; and potential to become pregnant. INTERVENTION SAMPLE SIZE: 2899 CONTROL SAMPLE SIZE: 3226
Intervention	OBJECTIVE: Assess whether a community-based participatory intervention could reduce the neonatal mortality rate INTERVENTION DESCRIPTION: The first step of the intervention was to discuss

	issues around childbirth and care behaviours in the community, which allowed facilitators to develop participatory learning skills and generated information on pregnancy and childbirth, covering beliefs and practices in both uncomplicated and complicated pregnancies. The facilitators then supported the women's groups through monthly meetings. This phase of ten meetings lasted almost a year. In the next steps of the intervention, the women's groups implemented and assessed their strategies. KEY CHW COMPETENCY: Support groups through an action-learning cycle in which they identified local perinatal problems and formulated strategies to address them TRAINING DURATION (DAYS): "Brief" TRAINING DESCRIPTION: Training in perinatal health issues; used manual based on Warimi project methodology REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: One supervisor provided support for every three facilitators by attending group meetings and making regular community visits. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/11/2001 END DATE: 31/10/2003 CONTROL DESCRIPTION: Health-service strengthening activities were undertaken in both intervention and control areas.
Outcomes	OUTCOMES: Primary: neonatal mortality rate. Secondary: stillbirths and maternal deaths, uptake of antenatal and delivery services, home care practices (Any antenatal care, Any iron and folic acid supplements, Any perceived maternal illness from 7 months gestation to 1 month postpartum, Visited health facility in event of illness, Institutional deliveries, Birth attended by government health provider, Birth attended by traditional birth attendant, Used a clean home delivery kit, Used a boiled blade to cut the cord, Attendant washed her hands, Cord undressed or dressed with antiseptic, Baby wrapped within 30 min, Baby bathed within 1 hour, Any of three infant illnesses (cough, fever, diarrhoea), Taken to health facility in event of illness, Breastfeeding initiated within 1 h, Discarded colostrum) infant morbidity, and health-care-seeking, Cost. DURATION OF FOLLOW UP: 2 years
Funding	DFID, with important support from the Division of Child and Adolescent Health, WHO, the United Nations Children's Fund, and the United Nations Fund for Population Activities.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "We randomly allocated one cluster in each pair to either intervention or control on the basis of a coin toss"
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Unclear Risk	Quote: "the trial allocation was not masked, but analysis of primary and secondary outcomes was not done until just before the data monitoring committee meeting at 30 months"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat."
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in the protocol were outcomes reported
Recruitment bias	Low Risk	Quote: "From March, 2001, to July, 2001, a team of 44 field interviewers visited every potential member of the cohort, reassessed her for inclusion, explained the study,

		asked for her consent, allocated her a unique identification number, and completed an individual questionnaire, which included questions on demography, education, maternity history, details of any preceding pregnancy, home-care practices, and use of health services for perinatal illness. ⁴ "
Baseline imbalance	Unclear Risk	Quote: "Some evidence exists of less poverty in intervention than control clusters: household asset scores and participant schooling—but not recalled annual food sufficiency— seem to slightly favour the pooled intervention clusters."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "taking account of clustering and the paired nature of the data, with hierarchical logistic models"

Mohan 2003

Trial reg.	None
Country	Iraq
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Volunteer RECRUITMENT CRITERIA: 15 women from a local NGO, the Iraqi Women's Federation were selected and assigned each to a TB centre coverage area. All the women were from the same district where the study was conducted. HEALTHCARE SETTING: Household PATIENT POPULATION: New smear-positive tuberculosis patients who had delayed collecting their drugs on one occasion. INTERVENTION SAMPLE SIZE: 240 CONTROL SAMPLE SIZE: 240
Intervention	OBJECTIVE: Over a period of 6 months the effect of home visits on compliance with DOTS, was studied on 480 new smear-positive tuberculosis patients who had delayed collecting their drugs on one occasion. INTERVENTION DESCRIPTION: Received home visits from trained personnel. The intervention patients received visits at home from a local volunteer in order to motivate them to attend the health centre daily. When a patient was recorded as late for collection of drugs, a woman was assigned to visit the patient at home and ask him/her to return to the centre. Home visitors also carried out health education for the patient and his/her family. If the patient was not at home, the volunteer returned another time and, if the patient refused to attend, she would try to convince him/her with the help of the family to report to the TB coordinator at the centre. KEY CHW COMPETENCY: Encourage patients to return to TB treatment at the clinic. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: The home visitors were trained under the guidance of the TB coordinator. They received standardized instructions and were assessed in a pilot study before the start of the main study. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: The TB coordinators (1 per district) were trained about the study design and how to identify latecomers, instruct the home visitors and carry out the follow-up. Another field worker was identified and trained to fill in the data collection forms. INCENTIVE: None INCENTIVE DESCRIPTION: All were volunteers who received no payment except an allowance for transport. RATIO CHW TO POPULATION: N.S. START DATE: 05/2001 END DATE: 05/2002

	CONTROL DESCRIPTION: No visits
Outcomes	OUTCOMES: TB smear conversion rates, rates of return, treatment success, default, treatment failure and death DURATION OF FOLLOW UP: 6 months
Funding	WHO Eastern Mediterranean Region (EMRO), Division of Communicable Diseases (DCD) and the WHO TDR: the EMRO/DCD/TDR Small Grants Scheme for Operational Research in Tropical and Communicable Diseases.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Equal number of patients were randomly allocated to the control or intervention group after the first episode of delaying. " Comment: Insufficient information to make a judgement
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified.
Blinding (performance and detection bias)	Low Risk	Quote: "The evaluation was blind as the information about outcome was collected by a field worker who did not know which group the patients was assigned to....Patients' attendance for treatment was noted from the TB centre records."
Incomplete outcome data (attrition bias)	Unclear Risk	Comment: Rate of loss to follow up unclear.
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified

More 2012

Trial reg.	ISRCTN96256793
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Settlement CLUSTERS: Intervention: 24, Control: 24
Population	LOCAL TERM FOR CHW: Sakhis RECRUITMENT CRITERIA: Local woman with secondary education and leadership skills, preferably married with children HEALTHCARE SETTING: Community PATIENT POPULATION: Women and infants INTERVENTION SAMPLE SIZE: 7944 CONTROL SAMPLE SIZE: 7759
Intervention	OBJECTIVE: Test an intervention in which urban slum-dweller women's groups worked to improve local perinatal health. INTERVENTION DESCRIPTION: A facilitator supported women's groups through an action learning cycle in which they discussed perinatal experiences, improved their knowledge, and took local action. The groups met fortnightly; the intervention followed a 36-meeting cycle that was predetermined in general but developed iteratively in detail KEY CHW COMPETENCY: Conduct meetings with women, attend planning and supervision meetings, and support group action. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: Sakhis attended regular training on a range of health care topics. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/10/2006

	END DATE: 30/09/2009 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: The primary outcomes: neonatal mortality, stillbirth rate, maternal morbidity, and extended perinatal mortality. Health care and morbidity outcomes: First antenatal visit before 3rd trimester, Antenatal care in public sector, Three or more antenatal care visits, Three packets of iron supplements, Rested more in 3rd trimester, Worked less in 3rd trimester, Ate more in 2nd and 3rd trimesters, Sentinel antepartum symptom, Sought clinical care for trigger symptom within 24 h, Institutional delivery, Breastfed within 24 h, Postnatal check, Exclusively breastfed for at least 28 d, Infant BCG, Any newborn problem, Sought clinical care for specified newborn illness within 24 h. DURATION OF FOLLOW UP: 6 wk. postpartum
Funding	ICICI Foundation for Inclusive Growth – Centre for Child Health and Nutrition and The Wellcome Trust.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "In a transparent process, social workers external to the trial drew lots to select 48 in blocks of eight per ward, and then to allocate four clusters per block to the intervention."
Allocation concealment (selection bias)	Low Risk	Quote: "The nature of the intervention precluded allocation concealment." Comment: Allocation concealment (selection bias) should not be an issue as in C-RCTs all clusters are randomised at once
Blinding (performance and detection bias)	Unclear Risk	Quote: "Analysts were blind to allocation"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "The analysis used records of individual births and was by intention to treat"
Selective reporting (reporting bias)	Unclear Risk	Comment: Trial retrospectively registered
Recruitment bias	Low Risk	Comment: Cluster-level consent for entire cluster population
Baseline imbalance	Low Risk	Quote: "A higher proportion of households in the control arm followed Islam (58%, compared with 33%). Control clusters had slightly more of the poorest quintile of women than intervention clusters (21% compared with 20%), but also of the least poor (23% compared with 17%)."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "The analysis...was... at cluster level."

Morrow 1999

Trial reg.	None
Country	Mexico
Methods	STUDY DESIGN: C-RCT (3-arm) UNIT OF RANDOMISATION: Cluster (4 city blocks) CLUSTERS: Intervention A: 13, Intervention B: 13, Control: 13
Population	LOCAL TERM FOR CHW: Peer counsellors RECRUITMENT CRITERIA: Three women who had previously worked for the Instituto Nacional de la Nutrición as field data collectors were trained to promote breastfeeding as peer counsellors. Each was a resident of the community in which they worked and had a commitment to breastfeeding, although they did not necessarily have previous personal breastfeeding experience. HEALTHCARE SETTING: Household PATIENT POPULATION: Mother-infant dyad INTERVENTION SAMPLE SIZE: Intervention A: 44, Intervention B: 52

	CONTROL SAMPLE SIZE: 34
Intervention	OBJECTIVE: Assess the efficacy of home-based peer counselling to increase the proportion of exclusive breastfeeding among mothers and infants residing in peri-urban Mexico City. INTERVENTION DESCRIPTION: Two intervention groups with different counselling frequencies: A) six visits and B) three visits. In the six-visit group, mothers were visited in mid and late pregnancy, in the first week and weeks 2, 4, and 8 post partum. In the three-visit intervention group, mothers were visited in late pregnancy, in the first week, and week 2 post partum. Home visits to pregnant women focused on the benefits of exclusive breastfeeding, especially during illness; basic lactation anatomy and physiology; positioning of the infant and “latching on”; common myths; typical problems and solutions; and preparation for birth. Postpartum visits with the mother focused on establishing a healthy breastfeeding pattern, addressing maternal concerns, and providing information and social support. Key family members who could provide support to mothers also were included in these counselling visits. KEY CHW COMPETENCY: Breastfeeding counselling TRAINING DURATION (DAYS): 67 TRAINING DESCRIPTION: Trained by La Leche League. These peer counsellors were trained and supervised by staff of La Leche League of Mexico and the physician study coordinator (MLG), who was also trained in lactation management. The peer-counsellor training consisted of 1 week of classes, 2 months in lactation clinics and with mother-to-mother support groups, and 1 day of observation and demonstration by visiting experts. REFRESHER OR ONGOING TRAINING: Finally, the peer counsellors practised in a neighbourhood nearby San Pedro Martir for 6 months before the intervention trial, during which the content of messages and problem-solving skills were refined. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Supervised by staff of La Leche League of Mexico and the physician study coordinator (MLG), who was also trained in lactation management. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:32 START DATE: 03/1995 END DATE: 12/1996 CONTROL DESCRIPTION: A control group that had no intervention.
Outcomes	OUTCOMES: Primary: exclusive breastfeeding, defined as giving maternal milk at the only infant food source in the previous week, with no other liquids or food given. Secondary: duration of breastfeeding, the proportion of infants who had an episode of diarrhoea in the first 3 months (cumulative incidence), and maternal satisfaction with counselling. DURATION OF FOLLOW UP: 6 months
Funding	Wellstart International's Expanded Promotion of Breastfeeding Program (USAID) cooperative agreement DPE-5966-A-00-1045-00) and the US National Institute of Child Health and Human Development (HD13021)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "13 clusters were allocated randomly to each study group, stratified by subdivision. This randomisation schedule was generated by computer. "
Allocation concealment (selection bias)	Low Risk	Comment: Allocation simultaneous with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "All data were collected through structured interviews of mothers residing in study households by two experienced staff other than the peer counsellors. The study hypothesis could not be concealed from these staff...Although peer counsellors taught intervention-

		group mothers that breastfeeding offers protection against diarrhoea and respiratory tract infections, there were fewer reports of infant diarrhoea in the intervention group than in the control group but no difference in non-specific respiratory signs and symptoms. Furthermore, most cases of infant diarrhoea were confirmed by a visit to a physician."
Incomplete outcome data (attrition bias)	Low Risk	Comment: 104/130 mother-infant pairs remained in the study at 6 months. Missing outcome data balanced in numbers across groups
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned
Recruitment bias	Low Risk	Quote: "Study mothers were identified by a semi-annual door-to-door census and continuous reporting of new pregnancies in the community by study staff and mothers. For more than 15 years, this method has been used to identify virtually all eligible mothers in the study community"
Baseline imbalance	Low Risk	Quote: "Study groups did not differ in baseline factors."
Loss of clusters	Unclear Risk	Quote: "Study participants were identified and enrolled in 31 of the 39 geographically defined clusters."
Unit of analysis issues	Low Risk	Quote: "Within-cluster correlation values were calculated for each study outcome and interview time and were found to be negligible: 0.011 at the 2-week visit and 0.074 at 3 months. Design effects were calculated and found to approach unity (no effect) for each study group. These results show that the cluster randomisation design achieved the equivalent of individual randomisation."

Mtango 1986

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village groups CLUSTERS: Intervention: 1, Control: 1
Population	LOCAL TERM FOR CHW: Village Health Worker (VHW) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Children <5 years INTERVENTION SAMPLE SIZE: 8028 CONTROL SAMPLE SIZE: 8098
Intervention	OBJECTIVE: Test control of ARI in children under five years of age. INTERVENTION DESCRIPTION: Community supported VHWs visited each family at their homes every six to eight weeks, giving health education on recognition and prevention of ARI, treating children with pneumonia on the spot with oral Cotrimoxazole or referring them to the next higher level of care. Training courses on ARI management were also given to health centre and dispensary workers in order to strengthen health services on the referral level. KEY CHW COMPETENCY: Giving health education on recognition and prevention of ARI, treating children with pneumonia or referring them TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Initial training REFRESHER OR ONGOING TRAINING: A year of in-service training and one month refresher course SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S.

	INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 06/1983 END DATE: 09/1985 CONTROL DESCRIPTION: No intervention, followed by deployment of CHWs in the second year
Outcomes	OUTCOMES: Under-5 mortality DURATION OF FOLLOW UP: 2 years
Funding	Not specified

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "The district's villages were divided randomly into two village groups" Comment: Not clear how this was done.
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	High Risk	Comment: Births and deaths registered by VHWs. Later resorted to a yearly census as VHWs were not reporting all events.
Incomplete outcome data (attrition bias)	Unclear Risk	Comment: Attrition rate at individual level not stated.
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified
Recruitment bias	Low Risk	Comment: VHW registered every birth in their catchment
Baseline imbalance	Unclear Risk	Comment: Baselines not compared
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "But having used a cluster sampling method, calculating the confidence interval of the difference in mortality for all clusters reveals no statistically significant difference between the intervention and control areas."

Neuner 2008

Trial reg.	None
Country	Uganda
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Counsellors RECRUITMENT CRITERIA: Nine refugees (five women, four men, mean age 27 years) from the community were trained as counsellors. Skills required to be accepted for the training included literacy in English and literacy in their mother tongue, as well as the ability to empathize with their clients and a strong motivation to carry out this work. HEALTHCARE SETTING: Community PATIENT POPULATION: Rwandan and Somalian refugees who were diagnosed with PTSD INTERVENTION SAMPLE SIZE: Intervention A: 111, Intervention B: 111 CONTROL SAMPLE SIZE: 55
Intervention	OBJECTIVE: The objective of this study was to examine whether trained lay counsellors can carry out effective treatment of PTSD in a refugee settlement. INTERVENTION DESCRIPTION: A) Strictly manualised narrative exposure therapy (NET) was compared with B) more flexible trauma counselling (TC). Both treatment types were conducted for six sessions (usually two sessions per week) that were between 1 and 2 hr. each in length. KEY CHW COMPETENCY: Provide strictly manualised NET or more flexible TC TRAINING DURATION (DAYS): 42

	TRAINING DESCRIPTION: The lay counsellor group of trainees was educated in a 6-week course in general counselling skills (e.g. active listening, empathy, verbalization, emotional processing, etc.) as well as specific abilities and methods that were needed for both treatment approaches. Each trainee was educated in all approaches. The trainers were five postdoctoral- and doctoral-level personnel from the University of Konstanz, Germany, the aid organization vivo, and Mbarara University of Science and Technology, in Uganda. Trainee counsellors were then closely tutored under supervision for both approaches before beginning to work autonomously with clients. REFRESHER OR ONGOING TRAINING: N.S SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: Case and personal supervision were still maintained thereafter on a weekly basis. Treatment adherence was monitored by case discussions in supervision meetings, by the direct observation of treatment sessions, and by a review of the testimonies and treatment protocols. No major deviation from the study protocol was noticed. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S START DATE: 2003 END DATE: 2004 CONTROL DESCRIPTION: A no-treatment monitoring group
Outcomes	OUTCOMES: PTSD diagnosis and severity DURATION OF FOLLOW UP: 9
Funding	This research was supported by Deutsche Forschungsgemeinschaft (DFG) and Bundesministerium für wirtschaftliche Zusammenarbeit (BMZ)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	High Risk	Quote: "The participants were assigned to the groups in the following way: The list of participants was ordered randomly; the first 4 were consecutively assigned to the NET, TC, NET, and TC groups; and the fifth was assigned to the control group. This procedure was repeated until all 277 participants were assigned." Comment: Alternate assignment; baseline imbalances in nationalities due to lack of stratification
Allocation concealment (selection bias)	High Risk	Comment: Allocation not concealed; participants were approached at home and allocated treatments after randomisation; baseline imbalances in prognostic variables.
Blinding (performance and detection bias)	Low Risk	Quote: "The participants who received treatment were reassessed 3 and 6 months later by the same local research assistants who had carried out the interviews for the survey. They were blind with respect to the particular treatment condition"
Incomplete outcome data (attrition bias)	Unclear Risk	Quote: "Aiming at an intention-to-treat analysis, we included in the outcome analysis all participants who were randomised." Comment: Dropouts more than 65% in all groups; significantly high differential dropout rates
Selective reporting (reporting bias)	Unclear Risk	Comment: No study protocol located.

Newell 2006

Trial reg.	None
Country	Nepal
Methods	STUDY DESIGN: C-RCT

	UNIT OF RANDOMISATION: District CLUSTERS: Intervention: 5, Control: 5
Population	LOCAL TERM FOR CHW: Female community health volunteer or a village health worker RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: New (i.e. never having previously received treatment for TB) sputum smear-positive patients who were aged 15 years or older, diagnosed at all health facilities within the study areas INTERVENTION SAMPLE SIZE: 549 CONTROL SAMPLE SIZE: 358
Intervention	OBJECTIVE: The Nepal national TB programme asked us to develop and test a strategy of direct observation of treatment for the hill districts of Nepal, where direct observation of treatment by health workers is not feasible. The study aimed to assess the success rates of two DOTS strategies developed for such areas. INTERVENTION DESCRIPTION: Community DOTS was defined as a strategy with drug taking supervised daily by a female community health volunteer or a village health worker KEY CHW COMPETENCY: Supervise drug taking TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: At the start of implementation in all community DOTS districts, all potential supervisors of directly observed treatment received training. The training focused on the DOTS strategy, the treatment process and its duration, the role of community supervisors in ensuring tuberculosis treatment completion, the side effects of tuberculosis drugs, and the process of referring patients for their management. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Drugs provided to the supervisor every month. The supervisor also traced patients who discontinued treatment. INCENTIVE: None INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 07/2002 END DATE: 07/2003 CONTROL DESCRIPTION: Family-member DOTS: a strategy with drug taking supervised daily by a household member selected by the patient, with drugs provided to the patient's supervisor every week. Government workers provided dedicated tracing.
Outcomes	OUTCOMES: Success rate (proportion of registered patients who achieved cure or completed treatment) DURATION OF FOLLOW UP: 8 months
Funding	DFID

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Nepal-based senior researcher established which strategy was allocated to which district by randomly drawing ten numbered identical pieces of paper from an opaque bag."
Allocation concealment (selection bias)	High Risk	Comment: Not specified.
Blinding (performance and detection bias)	Low Risk	Quote: "Researchers undertaking the primary analysis were unaware of the strategy identities."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat."
Selective reporting	Unclear Risk	Comment: No protocol available

(reporting bias)		
Recruitment bias	Low Risk	Quote: "This study was a pragmatic implementation of health-service provision strategies."
Baseline imbalance	Low Risk	Quote: "Clusters were well balanced at baseline"
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of analysis issues	Low Risk	Quote: "...takes into account the clustered nature of the trial using the coefficient of variation k"

Odendaal 2009

Trial reg.	None.
Country	South Africa
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Household
Population	LOCAL TERM FOR CHW: Paraprofessional visitors RECRUITMENT CRITERIA: Recruited from a cohort of community health workers employed by the collaborating NGO whose community work was external to the study. These health workers came from the community. HEALTHCARE SETTING: Household PATIENT POPULATION: Households were families with children 10 years or younger who lived in that house, visited there more than once a month, or stayed there during school holidays INTERVENTION SAMPLE SIZE: 112 CONTROL SAMPLE SIZE: 99
Intervention	OBJECTIVE: Assess the influence of a home visitation programme (HVP) on household hazards associated with unintentional childhood injuries in a South African low-income setting. INTERVENTION DESCRIPTION: Four intervention visits were conducted over 3 months, focusing on child development, and the prevention of burn, poison, and fall injuries. The HVP, a multi-component intervention, included educational inputs, provision of safety devices, and an implicit enforcement strategy. KEY CHW COMPETENCY: Share safety information (education), encourage the implementation of safe practices (enforce), distribute safety devices (engineer) TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Home visitors were jointly trained on home visitation and relationship-building skills and the four intervention topics REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: "Supervision meetings" INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:10 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: The control households received a courtesy visit after the post-intervention assessment, and were furnished with the same information and at least one of the safety devices.
Outcomes	OUTCOMES: Five aggregated hazard scores (Burns electrical, Burns paraffin, Burns safety practices, Poison, Falls) DURATION OF FOLLOW UP: 3 months
Funding	Western Cape Department of Health

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Randomisation of households to intervention or control group was performed" Comment: Unclear how this was done
Allocation concealment	Unclear Risk	Comment: Not specified

(selection bias)		
Blinding (performance and detection bias)	Unclear Risk	Quote: "Though the data collectors were not informed to the intervention or control status of households at post-intervention assessment, it is possible that they may have been alerted to the status of the household by the caregivers themselves or by observing the safety devices in the intervention households."
Incomplete outcome data (attrition bias)	Low Risk	Comment: intention-to-treat analysis and <10% attrition
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.

Ohnmar 2012

Trial reg.	ACTRN1261000070607
Country	Myanmar
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village CLUSTERS: Intervention: 21, Control: 17
Population	LOCAL TERM FOR CHW: Village volunteers RECRUITMENT CRITERIA: For health volunteers to meet eligibility criteria, they had to reside in the village, be of either gender, 18 years and above, having completed primary school and easily accessible to villagers, willing to work on a volunteer basis, accepted and respected by villagers, and willing to work under the supervision of health staff. In each village, the leader selected a volunteer who was then approved by the health staff. HEALTHCARE SETTING: Household PATIENT POPULATION: The study population included villagers currently living in the endemic area or its surroundings for at least one month. INTERVENTION SAMPLE SIZE: 268 CONTROL SAMPLE SIZE: 287
Intervention	OBJECTIVE: Determine whether the training of village volunteers and their regular supervision by health staff improve the coverage of timely diagnosis and treatment of malaria, to lower malaria-related mortality. INTERVENTION DESCRIPTION: The intervention consisted of training unpaid village volunteers in the provision of RDT and ACT and supervising them KEY CHW COMPETENCY: Patients who tested positive treated with a combination of AL. Patients with fever (with or without chills) whose test results were negative were administered chloroquine to suppress <i>P. vivax</i> malaria, after excluding local infection. The programme did not provide paracetamol, an antipyretic commonly used and available locally to treat fever. TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: A volunteer training module for malaria in Myanmar and the Job Aid illustration on the use of RDT, recommended by the WHO were used. The training topics covered communication skills, the nature, symptoms and modes of transmission of malaria, fevers, vulnerable populations, illustrations of RDT, treatment regimens, patient referral, prevention and control measures, thermometer reading, record-keeping and reporting. Referral to basic health staff was recommended for infants and pregnant women, and in the presence of symptoms such as sore throat, difficult urination, ear discharge, cough, loose motion and skin ulcers. The training did not cover the treatment of non-malarial diseases. REFRESHER OR ONGOING TRAINING: SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Midwives in the nearest health facility routinely monitored and supervised volunteers during their immunization visits to villages. Malaria supervisors occasionally visited volunteers. The first author assessed the regularity of activities through monthly review of calendar cards and reports and regularly visited these townships to collect forms from malaria supervisors but did not supervise their activities

	INCENTIVE: Non-monetary INCENTIVE DESCRIPTION: "Unpaid" "Volunteers were reimbursed for their travel costs (7.5 to 15 USD) and meals (17 USD) during their stay in Bago. Giveaways included a T-shirt, a cap and a handbag with a malaria volunteer logo." RATIO CHW TO POPULATION: START DATE: 08/03/2009 END DATE: 17/02/2010 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Primary: (i) whether villagers had received blood test-based treatment for fever, (ii) interval between onset of fever and the blood test, (iii) mortality. Secondary: variable was knowledge of malaria mode of transmission. DURATION OF FOLLOW UP: 6 months
Funding	WHO/SEARO-TDR Small Grants Scheme, New Delhi, India (T/16/72/83/MMR - SN1162).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Eligible villages (clusters) without existing health staff were then randomised into intervention and comparison groups." Comment: Unclear how done
Allocation concealment (selection bias)	High Risk	Quote: "The nature of the intervention allowed neither allocation concealment nor blinding."
Blinding (performance and detection bias)	High Risk	Quote: "The nature of the intervention allowed neither allocation concealment nor blinding."
Incomplete outcome data (attrition bias)	Low Risk	Comment: Intention-to-treat analysis was used
Selective reporting (reporting bias)	Unclear Risk	Comment: Retrospectively registered
Recruitment bias	Unclear Risk	Comment: Not specified.
Baseline imbalance	Low Risk	Quote: "The results of the randomisation process were not perfect as mean distance from home of the villagers to the health facility in the comparison group was around 1 km, far more than that of the intervention group. However, the results of multivariable analysis showed that the outcome was not significantly influenced by the distance and so the distance could not explain the difference between the intervention and the comparison groups"
Loss of clusters	Low Risk	Quote: "Of 21 intervention villages, two villages did not receive training due to a delay in communication from local health staff. One volunteer dropped out a week after training for personal reasons. Another did so after two months to avoid a personal conflict with a relative receiving temporary, paid RDT from PSI/Myanmar. All 21 villages were included in the intervention group in the analysis."
Unit of analysis issues	Low Risk	Quote: "Linear mixed model with random intercepts was employed to predict the binary outcome variable (diagnostic blood test for malaria) with the individual (subjects with febrile illness) as the first level and the cluster (village) as the second level."

Pachón 2002 (incl. Schroeder 2002, Sripaipan 2002)

Trial reg.	None
Country	Vietnam
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Community

	CLUSTERS: Intervention: 1, Control: 1
Population	LOCAL TERM FOR CHW: Local health volunteers RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household/community PATIENT POPULATION: Children age 5 to 25 months old and their caregivers INTERVENTION SAMPLE SIZE: 119 CONTROL SAMPLE SIZE: 119
Intervention	OBJECTIVE: Evaluate the effect of Save the Children's (SC) nutrition programme on the complementary food intake of young rural Vietnamese children. INTERVENTION DESCRIPTION: Malnourished children and their caregivers participated in intensive nutrition rehabilitation sessions that lasted two weeks per month and continued for up to nine months. Furthermore, all children less than three years of age participated in growth monitoring and promotion sessions every other month for two years and were dewormed once during the first six months of the program. Messages promoted at the growth monitoring and at the nutrition rehabilitation sessions included feeding children "good foods" (i.e. positive deviant foods) identified during the interviews with positive deviant families, increasing the quantity of foods given to children, feeding 5 to 6 times per day, continued breast-feeding, and health-seeking. KEY CHW COMPETENCY: Growth monitoring to identify malnourished children, a positive deviance informed approach to identify local foods or behaviours that are associated with better growth, and intensive feeding based on the "hearth model" TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Save the Children staff conducted a training of trainers that included management training for provincial, district, and commune partners and implementation training for designated district trainers. These, in turn, trained local health volunteers to implement the 10-month community empowerment and nutrition programme including the situation analysis and interventions in their communes. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: The supervisor reviewed all health volunteers' hamlet summaries and aggregated the commune's data, often plotting them on a prominently displayed chart or table in the People's Committee headquarters or commune health centre. INCENTIVE: None INCENTIVE DESCRIPTION: None RATIO CHW TO POPULATION: N.S. START DATE: 12/1999 END DATE: 12/2000 CONTROL DESCRIPTION: Usual care
Outcomes	OUTCOMES: Daily frequency of consuming intervention-promoted food and non-breast milk liquids and food, daily quantity of non-breast milk liquids and food consumed, daily energy intake, and proportion of children meeting daily median energy requirements, weight-for-age Z scores (WAZ), height-for-age Z scores (HAZ), weight-for-height Z scores (WHZ), level of reported childhood diarrhoea, symptoms of respiratory illness. DURATION OF FOLLOW UP: Six months with a re-survey at 12 months
Funding	Save the Children/US (Hanoi and Westport, Conn., USA), the USAID-funded LINKAGES Project (Washington, D.C.), Emory University's Rollins School of Public Health (Atlanta, Ga., USA), and the Research and Training Center for Community Development (Hanoi).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "...we randomly assigned, by coin toss, one group of six for the intervention and the other group of six as comparison communes."

Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "Blinding subjects or data gatherers to the intervention was not feasible."
Incomplete outcome data (attrition bias)	Low Risk	Comment: 95% retention
Selective reporting (reporting bias)	Unclear Risk	Comment: Protocol is unclear regarding what outcome variables are being assessed in what units as well as which are primary outcomes.
Recruitment bias	Low Risk	Quote: "120 children 5 to 25 months old were randomly selected from the intervention communes and another 120 children were randomly selected from the comparison communes."
Baseline imbalance	Unclear Risk	Comment: "No relevant difference at baseline except Weight-for-height"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	High Risk	Comment: No adjustment for clustering

Pence 2007 (incl. Pence 2005)

Trial reg.	None
Country	Ghana
Methods	STUDY DESIGN: C-RCT (2x2 factorial) UNIT OF RANDOMISATION: MOH sub-district CLUSTERS: Intervention A: 1, Intervention B: 1, Control A: 1, Control B: 1
Population	LOCAL TERM FOR CHW: Volunteers RECRUITMENT CRITERIA: The committees appoint and oversee a cadre of primarily male health volunteers. HEALTHCARE SETTING: Household/Community PATIENT POPULATION: Children 0-59 months INTERVENTION SAMPLE SIZE: (A) 8,218, (B) 18,350 CONTROL SAMPLE SIZE: (A) 8,109, (B) 18,457
Intervention	OBJECTIVE: The Navrongo Community Health and Family Planning Project is a longitudinal community trial of alternative organizational strategies for health service delivery in a rural, impoverished area of Ghana INTERVENTION DESCRIPTION: Intervention A: Training of a local health volunteer and community involvement in health delivery. Intervention B: Volunteers and nurse doorstep visitation and services. Volunteers visit households to talk about hygiene, child immunization, and other health issues, and to make it known that they are available for basic treatments and referrals to the clinics and hospital. They are supplied with paracetamol, chloroquine, aludrox, and multivitamins, but they do not carry antibiotics or vaccines. KEY CHW COMPETENCY: Visit households to talk about hygiene, child immunization, and other health issues, and to make it known that they are available for basic treatments and referrals to the clinics and hospital. They are supplied with paracetamol, chloroquine, aludrox, and multivitamins, but they do not carry antibiotics or vaccines TRAINING DURATION (DAYS): 30 TRAINING DESCRIPTION: Volunteers receive six weeks of initial training REFRESHER OR ONGOING TRAINING: Quarterly refresher training sessions SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/07/1993 END DATE: 30/04/2000 CONTROL DESCRIPTION: Control A: In one area, nurses are placed in communities with doorstep visitation and service responsibilities. Control B:

	Comparison group.
Outcomes	OUTCOMES: Under-5 mortality DURATION OF FOLLOW UP: 5 years
Funding	Rockefeller Foundation

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "The three interventions and the comparison condition were then randomly assigned" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	Unclear Risk	Comment: Used field workers to collect the information, unclear if they were blinded.
Incomplete outcome data (attrition bias)	Unclear Risk	Comment: Insufficient information to make an assessment
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified
Recruitment bias	Unclear Risk	Comment: Not clear.
Baseline imbalance	Low Risk	Quote: "We used Poisson regression models to compare area-level mortality rates before and during the intervention within each area while adjusting for baseline differences in mortality between areas and potential confounding variables."
Loss of clusters	Low Risk	Quote: No clusters lost
Unit of analysis issues	High Risk	Quote: "The study region could not reasonably be divided into more than four areas...[this prevents] the application of statistical methods to correct for intra-area correlation of errors." Comment: Results did not reflect the clustering

Penfold 2014 (incl. Hanson 2015)

Trial reg.	NCT01022788
Country	Tanzania
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Ward (Each ward consists of an average of five villages, approximately 8,000 people, and 260 births per year.) CLUSTERS: Intervention: 65, Control: 67
Population	LOCAL TERM FOR CHW: Volunteer RECRUITMENT CRITERIA: Over 800 women who volunteered and were not currently involved in other community activities were recruited from and by their communities (two per village in intervention wards) HEALTHCARE SETTING: Household PATIENT POPULATION: Women of reproductive age INTERVENTION SAMPLE SIZE: 2486 CONTROL SAMPLE SIZE: 2490
Intervention	OBJECTIVE: Test a sustainable and scalable home-based counselling intervention for delivery by community volunteers in rural southern Tanzania to improve newborn care practices and survival one year after full implementation. INTERVENTION DESCRIPTION: Trained volunteers made home visits during pregnancy and after childbirth to promote recommended newborn care practices including hygiene, breastfeeding and identification and extra care for low birth weight babies. For every pregnant woman identified in her village, a volunteer was expected to make three visits to her home during pregnancy and two in the early neonatal period, with additional visits for small babies. KEY CHW COMPETENCY: Promotion of behaviour change TRAINING DURATION (DAYS): 14 TRAINING DESCRIPTION: They were trained for five days by their district health

	teams and followed-up in their villages after starting work as volunteers conducting home visits. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Volunteers were supported through quarterly review meetings with district health leaders, monthly contacts with village executive officers, who facilitated the link with the community, and with health facility staff, who provided technical support INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S START DATE: 06/2010 END DATE: 09/2011 CONTROL DESCRIPTION: Routine care only
Outcomes	OUTCOMES: Primary: neonatal mortality, breast feeding within an hour of delivery, birth attendants for home deliveries washing hands before childbirth or wearing gloves, and babies fed only breast milk in the first three days. Secondary: other behaviours promoted during counselling to maximise newborn health (Childbirth in a health facility, Childbirth with a skilled attendant, Prepared soap, Prepared new or washed cloth for drying baby, Prepared cloth or mat for childbirth, Cleaned floor where childbirth to take place, Had plan in case of emergency childbirth, Attendant had clean hands during childbirth, Baby had cord cut with new or sterilised blade, Baby had cord tied with new thread, Baby dried <5 minutes after birth, Baby wrapped <5 minutes after birth, Baby bathed at least six hours after birth, Baby had nothing applied to umbilical cord) DURATION OF FOLLOW UP: 12 months, 36 months
Funding	Bill & Melinda Gates Foundation through the Saving Newborn Lives programme of Save the Children, UNICEF, the Laerdal Foundation and the Batchworth Trust.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "...using random numbers generated by Microsoft Excel."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "The data analyst was masked to the cluster allocation until analysis was complete."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "for those who gave birth in the preceding year in intervention and comparison areas, using the allocation given at the sub-village level (intention-to-treat analysis)"
Selective reporting (reporting bias)	Unclear Risk	Comment: Primary outcomes in protocol and study differed, though all outcome reported
Recruitment bias	Low Risk	Comment: A representative sample of intervention and control area households included in survey
Baseline imbalance	Low Risk	Quote: "Background characteristics were similar in intervention and comparison areas."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "...including a design effect of 1.5 to account for clustering."

Powell 1989

Trial reg.	None.
Country	Jamaica
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Child
Population	LOCAL TERM FOR CHW: Health paraprofessionals/community health aides

	RECRUITMENT CRITERIA: Most of the community health aides had not completed secondary education. HEALTHCARE SETTING: Household PATIENT POPULATION: Children 16-30 months INTERVENTION SAMPLE SIZE: 29 CONTROL SAMPLE SIZE: 29
Intervention	OBJECTIVE: Study home visiting and psychosocial stimulation with deprived urban children in Jamaica INTERVENTION DESCRIPTION: Weekly 1 hour visits by CHWs to provide psychosocial stimulation (i.e. demonstrate play; engage in language activities, songs, and crafts; encourage mothers to have friendly rather than authoritarian relationships with their children) KEY CHW COMPETENCY: Provide psychosocial stimulation TRAINING DURATION (DAYS): 56 TRAINING DESCRIPTION: The community health aides previously received 8 weeks of nutrition and healthcare training. They received a further 8 weeks training in child development, teaching techniques, and toy making REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: Throughout the program, 5% of visits were observed by the nurse-supervisor, who filled out an evaluation checklist. This was then discussed with the community health aides. The aides kept record books for each child and recorded every visit. Weekly workshops were held with each aide by the nurse. INCENTIVE: Monetary INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:13 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: No visits
Outcomes	OUTCOMES: Griffiths Mental Developmental subscales: hearing and speech, hand and eye coordination, and performance. DURATION OF FOLLOW UP: 12 months
Funding	Ford Foundation, US; the Overseas Development Agency; and the Wellcome Trust, UK

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Random assignment was used" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified
Blinding (performance and detection bias)	Low Risk	Quote: "Testers were all women who had been trained by the same person and were unaware of the child's group."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "No child withdrew from [the] study"
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.

Powell 2004 (incl. Baker-Hennigham 2005)

Trial reg.	None.
Country	Jamaica
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Nutrition clinics CLUSTERS: Intervention: 11, Control: 7
Population	LOCAL TERM FOR CHW: Community health aides

	RECRUITMENT CRITERIA: Community health aides are paraprofessionals employed in government health centres. HEALTHCARE SETTING: Household PATIENT POPULATION: Mothers of 139 undernourished children (WAZ -1.5) aged 9–30 months INTERVENTION SAMPLE SIZE: 70 CONTROL SAMPLE SIZE: 69
Intervention	OBJECTIVE: To determine the effect of early childhood stimulation with undernourished children and their mothers on maternal depression. INTERVENTION DESCRIPTION: Weekly home visits by community health aides for one year. Mothers were shown play activities to do with their child using home made materials, and parenting issues were discussed. The intervention focused on improving child development by improving mothers' knowledge and practices of child rearing and their parenting self-esteem. KEY CHW COMPETENCY: Mothers were shown play activities to do with their child using home made materials, and parenting issues were discussed. TRAINING DURATION (DAYS): 28 TRAINING DESCRIPTION: Four weeks pre-service training on health and nutrition and an additional two weeks training covering child development, parenting issues, and how to conduct the intervention REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Biweekly-Monthly SUPERVISION DESCRIPTION: The supervisor (a nurse) observed each aide conducting visits once a month and visited the health centre every fortnight to discuss the programme and review the records of each visit. INCENTIVE: Monetary INCENTIVE DESCRIPTION: CHWs are employed by the government RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Standard health and nutrition care only
Outcomes	OUTCOMES: Children's scores on the Griffiths mental development scales and mothers' knowledge and practices of childrearing measured by questionnaires. Frequency of maternal depressive symptoms was assessed by questionnaire. DURATION OF FOLLOW UP: 1 year
Funding	Thrasher Research Fund, USA, with subsidiary grants from the British High Commission–DFID, Jamaica and the University of the West Indies Mona Campus Research and Publication Fund. The Ministry of Health Jamaica supported the Community Health Aides. This work was undertaken in collaboration with Great Ormond Street Hospital for Children NHS Trust which receives a proportion of its funding from the NHS Executive.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "clinics were...randomly assigned" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: No information on how this was done
Blinding (performance and detection bias)	Low Risk	Quote: "Two interviewers administered the questionnaires in the study; they were unaware of the mothers' intervention status...The children were assessed by one of two persons who...were unaware of the children's group"
Incomplete Outcome data	Low Risk	Quote: "There were no differences between mothers lost to the study and mothers interviewed on any of the baseline characteristics."

		Comment: 10% attrition
Selective reporting (reporting bias)	High Risk	Comment: Final global development quotient scores for children in intervention and control not reported
Recruitment Bias	High Risk	Comment: Women recruited to study after clinics randomised
Baseline imbalance	Low Risk	Quote: "There were no statistically significant differences between the groups in socioeconomic background variables, parental characteristics and child age, developmental quotients, gender and nutritional status, and maternal depression on enrolment"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of Analysis Issues	Low Risk	Quote: "The effect of treatment was examined...using multilevel multiple regression analysis to take into account the hierarchical structure of the study"

Pradeep 2014

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Villages CLUSTERS: Intervention: 3, Control: 3
Population	LOCAL TERM FOR CHW: Community Health Workers RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Treatment naïve women diagnosed with depression, who did not receive any treatment in the last six months INTERVENTION SAMPLE SIZE: 122 CONTROL SAMPLE SIZE: 138
Intervention	OBJECTIVE: Primary: evaluate whether enhanced care resulted in greater number of treatment naïve women with depression living in the community sought help from the primary care center, and to examine whether adherence to antidepressant medication would be better in women receiving enhanced care compared to treatment as usual. Secondary: assess whether there was a change in the severity of depression and quality of life (QOL) before and after the treatment intervention INTERVENTION DESCRIPTION: CHWs visited patients immediately following the first medical consultation, educated the patient and her family members about depression and its treatment. They also emphasised taking antidepressant medication and continuing the treatment regimen. This was followed by another visit in the subsequent week to enquire about any possible side effects of medication and clarification of any doubts concerning the medical treatment of depression. This pattern of visits was maintained after every monthly consultation with the physician in the TI group. In addition, CHWs visited those patients who discontinued medication and / or those who did not visit the PHC for an initial consultation and encouraged them to resume treatment in the intervention group. KEY CHW COMPETENCY: Provided health education and adherence support TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:30 START DATE: 08/2006

	END DATE: 09/2009 CONTROL DESCRIPTION: Patients diagnosed with depression were encouraged to seek help from the physician at PHC with no additional input from the CHW.
Outcomes	OUTCOMES: Number of women who sought and completed treatment, number of clinic visits, duration of treatment with antidepressant, changes in severity of depression (HDRS) and changes in quality of life [WHO-QOL (Brev) scale]. DURATION OF FOLLOW UP: 6 months
Funding	Anuradha Foundation, Los Altos, CA, USA

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The six villages covered by the PHC were randomized into two groups of three villages each" Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear risk	Comment: Unclear how this was done.
Blinding (performance and detection bias)	Low risk	Quote: "The research assistant was ...blind to the two groups"
Incomplete outcome data (attrition bias)	Low risk	Comment: No attrition reported
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified
Recruitment bias	Low risk	Quote: "General Health Questionnaire (GHQ) -28 item versio7 was administered to all adult women (>18 yr) in each of the selected villages"
Baseline imbalance	Low risk	Quote: "At baseline, there were no significant differences between the TAU and the TI groups in terms of socio-demographic data, clinical characteristics and co-morbid diagnosis"
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Quote: "Adjustments were made in the statistical tests to account for the clustering effect of village for the primary outcomes. Intra-cluster correlation i.e. ICC (ρ) was calculated on the log transformed variable. Design effect was calculated by $1 + (n-1) \rho$ "

Rahman 2009

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village CLUSTERS: Intervention: 24, Control: 24
Population	LOCAL TERM FOR CHW: Lady Health Workers RECRUITMENT CRITERIA: Existing government CHWs HEALTHCARE SETTING: Household/community PATIENT POPULATION: Healthy women aged 17–40, married, in their last trimester of pregnancy that were registered with LHWs INTERVENTION SAMPLE SIZE: 177 CONTROL SAMPLE SIZE: 157
Intervention	OBJECTIVE: Adapted one parent-based programme that promoted early psychological development of children (Learning Through Play) for a rural Pakistani population and aimed to study if: (1) it was acceptable to community health workers; (2) the programme led to an improvement, after a period of 6 months, in mothers' knowledge and attitudes about early infant development; (3) it led to a reduction in the levels of maternal mental distress in the post-natal period. INTERVENTION DESCRIPTION: After receiving training, each LHW conducted a half-day workshop on the second birth month stage of development with a group

	of mothers recruited from her area (each group comprised 6–8 mothers). Each mother was given a ‘Learning Through Play’ calendar to take home. Thereafter, during their routine fortnightly home visits, LHWs spent 15–20 min with the mothers to discuss their child’s development, using the calendar as the basis for discussion. In addition, the mothers were encouraged to meet in groups on their own to support each other in the use of the techniques outlined in the calendar. (24 villages) KEY CHW COMPETENCY: Deliver Learning Through Play curriculum TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Twenty-four LHWs from the intervention villages were recruited to receive training in the ‘Learning Through Play’ programme from a trained psychologist. Each LHW was given the Urdu version of the manual for reference, along with suggested teaching guidelines. REFRESHER OR ONGOING TRAINING: 1 hour SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:8 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: All women in the intervention group then received the programme, while the control group had routine follow-up visits.
Outcomes	OUTCOMES: Primary: Mothers’ knowledge and attitudes about the second birth month stage of development was assessed using a specially developed Infant Development Questionnaire (IDQ). Secondary: The level of mental distress in mothers was measured using the World Health Organization’s Self-Reporting Questionnaire (SRQ-20). DURATION OF FOLLOW UP: 6 months
Funding	Not stated.

Risk of Bias

Item	Reviewers’ judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Randomisation was conducted by an independent Trial Centre in Rawalpindi. Villages were assigned by simple random allocation using a table of random numbers by a researcher not involved in the study and who was blind to the identity of the villages."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "All questionnaires were administered individually to mothers in groups by workers who were blind to the control-intervention status of the two groups."
Incomplete outcome data (attrition bias)	High Risk	Comment: Per-protocol analysis
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified
Recruitment bias	Low Risk	Quote: "With the assistance of the local health centres, all healthy women aged 17–40, married, in their last trimester of pregnancy that were registered with LHWs were approached to participate in the study...Out of 367 approached, 334 (91%) met the eligibility criteria and

		agreed to take part in the study."
Baseline imbalance	Low Risk	Quote: "Baseline characteristics...were broadly similar, although there were more primiparous mothers and fewer male babies in the intervention group. There were no statistically significant differences between the groups in age, years of education and family income."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Comment: ICC provided

Rahman 2008 (incl. Rahman 2012)

Trial reg.	ISRCTN65316374
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Union council CLUSTERS: Intervention: 20; Control: 20
Population	LOCAL TERM FOR CHW: Lady health workers RECRUITMENT CRITERIA: Existing government CHWs HEALTHCARE SETTING: Household PATIENT POPULATION: Married women (aged 16–45 years) in their third trimester of pregnancy with perinatal depression were eligible to participate. INTERVENTION SAMPLE SIZE: 463 CONTROL SAMPLE SIZE: 440
Intervention	OBJECTIVE: Integrated a cognitive behaviour therapy-based intervention into the routine work of community-based primary health workers in rural Pakistan and assessed the effect of this intervention on maternal depression and infant outcomes. INTERVENTION DESCRIPTION: The intervention consisted of a session every week for 4 weeks in the last month of pregnancy, three sessions in the first postnatal month, and nine 1-monthly sessions thereafter. Based on principles of cognitive behaviour therapy, the intervention was delivered by ordinary village-based primary health workers. KEY CHW COMPETENCY: Delivery psychological intervention (manualised intervention based on principles of cognitive behaviour therapy) TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: A manual was used (with step by step instructions for each session) to train the health workers and for them to keep for reference. The intervention, called the Thinking Healthy Programme, used cognitive behaviour therapy techniques of active listening, collaboration with the family, guided discovery, and homework, and applied these to health workers' routine practice of maternal and child health education. REFRESHER OR ONGOING TRAINING: 1 day refresher after 4 months SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Health workers in both groups received monthly supervision, and were monitored by the research team to ensure that they were attending the scheduled visits INCENTIVE: Monetary INCENTIVE DESCRIPTION: Government salary RATIO CHW TO POPULATION: 1:12 START DATE: 04/2005 END DATE: 03/2006 CONTROL DESCRIPTION: In the control group untrained health workers made an equal number of visits to the depressed mothers. In practice, the Lady Health Workers seldom provide such structured and monitored care in the community. The control group thus received what would be regarded as ideal care
Outcomes	OUTCOMES: Primary: infant weight and height at 6 months and 12 months. Secondary: maternal depression. DURATION OF FOLLOW UP: 12 months

Funding	Wellcome Trust.	
<i>Risk of Bias</i>		
Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "These administrative units were assigned by random allocation with a table of random numbers by a researcher who was not involved in the study and who was unaware of the identity of the Union Councils"
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "Participants only knew that they were receiving one of two types of visits from the health workers....interviewers were unaware of what group the participants were assigned to."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat"
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes mentioned in protocol were reported in trial
Recruitment bias	Low Risk	Comment: All women in third trimester of pregnancy were recruited to the trial (~10% not located/refusal)
Baseline imbalance	Low Risk	Quote: "The baseline characteristics in the two groups were similar"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Comment: ICCs listed

Ramaiah 2001

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Blocks (subunits of a district) CLUSTERS: Intervention: 20, Control: 20
Population	LOCAL TERM FOR CHW: Drug distributors RECRUITMENT CRITERIA: Selected by village leaders HEALTHCARE SETTING: Household PATIENT POPULATION: 1-60 years INTERVENTION SAMPLE SIZE: 270000 CONTROL SAMPLE SIZE: 250000
Intervention	OBJECTIVE: A new drug delivery strategy community-directed treatment (comDT) was developed and implemented and its effectiveness compared with that of the traditional health services-organized drug delivery INTERVENTION DESCRIPTION: Communities were given the responsibility to distribute diethylcarbamazine. ComDT envisages distribution of drugs by communities, who alone decide the timing, duration and mode of distribution, select drug distributors and maintain records, with a minimum role for health services. KEY CHW COMPETENCY: Distribute diethylcarbamazine TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Trained by VHNs in the mode of transmission of and the role of the drug in the control and elimination of lymphatic filariasis REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 12/1998 END DATE: 03/1999 CONTROL DESCRIPTION: Health services personnel from primary health centres

	distributed the drug.
Outcomes	OUTCOMES: Percentage of the population that complied with treatment, percentage of the population that received DEC DURATION OF FOLLOW UP: Immediate
Funding	UNDP/World Bank/WHO/TDR (Grant no. 970628).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "The study was implemented in four blocks and two blocks each were randomly allocated to ComDT and [control] arms." Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	Unclear Risk	Comment: Insufficient information to permit judgement
Incomplete outcome data (attrition bias)	Unclear Risk	Comment: Insufficient information to permit judgement
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol mentioned
Recruitment bias	Low Risk	Quote: "In each village, all streets were listed and four were randomly chosen to sample the households. In each selected street five households were randomly selected. Thus, a total of 20 households were sampled to collect data."
Baseline imbalance	Unclear Risk	Comment: Insufficient information to permit judgement
Loss of clusters	Unclear Risk	Comment: Insufficient information to permit judgement
Unit of analysis issues	Low Risk	Quote: "after adjusting them for clustering"

Rotheram-Borus 2014 (incl. le Roux 2013, le Roux 2014, Tomlinson 2015)

Trial reg.	NCT00996528
Country	South Africa
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Neighbourhood CLUSTERS: Intervention: 12; Control: 12
Population	LOCAL TERM FOR CHW: Community health workers RECRUITMENT CRITERIA: CHW were selected to have good social/communication skills, problem-solving skills, and thriving children (positive deviants). HEALTHCARE SETTING: Household PATIENT POPULATION: Pregnant women who were at least 18 years old and able to give informed consent INTERVENTION SAMPLE SIZE: 644 CONTROL SAMPLE SIZE: 594
Intervention	OBJECTIVE: To evaluate the effect of home visits by CHW on maternal and infant well-being from pregnancy through the first six months of life for women living with HIV and all neighbourhood mothers. INTERVENTION DESCRIPTION: Home visits by CHW (Philani Intervention Program) KEY CHW COMPETENCY: Provide: 1) foundational skills in behaviour change; 2) application of key health information about HIV, alcohol use, malnutrition, and general maternal and child health; and 3) coping with their own life challenges. TRAINING DURATION (DAYS): 30 TRAINING DESCRIPTION: CHW were trained for one month using an intervention manual, role-playing, and watching videotapes of common challenging situations that CHW might face during home visits.

	REFRESHER OR ONGOING TRAINING: Supervisors provided monthly CHW in-service trainings. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: Supervisors monitored implementation by reviewing charts and visit documentation on mobile phones, making random site visits once every two weeks. INCENTIVE: Monetary INCENTIVE DESCRIPTION: CHWs worked 20 hours weekly, and were paid R1250 a month (about 150 USD). RATIO CHW TO POPULATION: N.S. START DATE: 01/05/2009 END DATE: 01/09/2010 CONTROL DESCRIPTION: Comprehensive healthcare at clinics
Outcomes	OUTCOMES: Primary: a composite score that is calculated as the sum of indices for the presence or absence of maternal and child health and well-being across the five domains: 1) Child health status, 2) Healthcare and monitoring including the number of antenatal and postnatal clinic visits and additional care required if a mother is HIV-positive, 3) HIV-related preventive acts including disclosure of HIV test results to partners, requesting partners to test, consistent use of condoms, and the following among HIV-positive mothers: Use of single feeding method until 6 months, baby testing at 6 weeks, baby prescribed antibiotics at 6 weeks, maternal AZT during pregnancy, NVP at birth, and AZT after birth, and disclosure of baby's status to partner or family; 4) Mental health including the Edinburgh Postnatal Depression Scale, and 5) Social support including an adapted scale from Barrera and colleagues indicating instrumental, emotional, and childrearing support for healthy acts by family, neighbour, and community. Other outcomes: height-for-age, weight-for-age, height-for-weight-for age DURATION OF FOLLOW UP: 6 months (child growth outcomes), 8 months (all other outcomes)
Funding	NIAAA Grant # 1R01AA017104 and NIH grants MH58107, 5P30AI028697, and UL1TR000124. M. T. is supported by the National Research Foundation (South Africa).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "The UCLA team randomised neighbourhoods within matched pairs to PIP or SC using simple randomisation." Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified
Blinding (performance and detection bias)	Low Risk	Quote: "A driver transported all participants to a central assessment site, allowing interviewers to be blinded to condition. However, participants may have spontaneously talked about CHW to interviewers."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "There were no significant selection effects between mothers who were successfully reassessed over the six months post-birth and those who were not" Comment: Attrition <10%
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes noted in the protocol were reported
Recruitment bias	Unclear Risk	Comment: Women recruited post-randomisation
Baseline imbalance	Low Risk	Quote: At baseline: "Women were similar across conditions on: outcome-related measures...matching criteria...demographic characteristics...general health...One significant baseline difference...among women who had been pregnant before, SC had a higher mean number of previous births."

Loss of clusters	High Risk	Quote: "One matched pair was eliminated after 6 months of recruitment due to low numbers of pregnant women (total n=10 pregnant women), leaving 24 study neighbourhoods "
Unit of analysis issues	Low Risk	Comment: Adjusted for clustering

Saleem 2014

Trial reg.	NCT01128517
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: "Geographically-distinct areas" CLUSTERS: Total:10; breakdown not specified
Population	LOCAL TERM FOR CHW: Community health workers (CHWs) RECRUITMENT CRITERIA: Females with at least 10 years of schooling HEALTHCARE SETTING: Household PATIENT POPULATION: Infants INTERVENTION SAMPLE SIZE: 118 CONTROL SAMPLE SIZE: 94
Intervention	OBJECTIVE: This cluster-randomised interventional trial at peri-urban settings of Karachi was conducted to evaluate the impact of maternal educational messages regarding appropriate complementary feeding on the nutritional status of their infants after 30 weeks of educational interventions delivered by trained community health workers. INTERVENTION DESCRIPTION: Mothers in the intervention group received three education modules about breastfeeding (BF) and appropriate complimentary at a baseline visit and two subsequent visits 10 weeks apart from CHWs KEY CHW COMPETENCY: Breastfeeding and complimentary feeding education TRAINING DURATION (DAYS): 14 TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 2007 END DATE: N.S. CONTROL DESCRIPTION: The control group received advice about BF according to national guidelines. Infants' growth [weight, length, and mid-upper arm-circumference (MUAC), stunting, wasting, and underweight] were measured at four time points.
Outcomes	OUTCOMES: Weight (kg), Length (cm) MUAC (cm), Wasting (% SD weight-for-length z-score according to WHO growth charts for age and gender), Stunting (% SD length-for-age score according to WHO growth charts for age and gender); Underweight (%SD weight-for-age z-score according to WHO growth charts for age and gender) DURATION OF FOLLOW UP: 7.5 months after baseline
Funding	Aga Khan University Research Council and NIH-Fogarty research training fund

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Ten geographically-distinct areas were randomly assigned by using random number table to educational intervention and control."
Allocation concealment	Low Risk	Comment: Simultaneous with randomisation

(selection bias)		
Blinding (performance and detection bias)	Unclear Risk	Quote: "In order to minimize the bias, educational session was conducted, and infants' anthropometric measurements were taken by different teams and on different days."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "all the models were adjusted for age in days, vaccination status, and parents' educational status."
Selective reporting (reporting bias)	Unclear Risk	Comment: Trial retrospectively registered. Not all secondary outcomes listed in protocol (diarrhoea episodes, ARI) were reported
Recruitment bias	Unclear Risk	Quote: "Potential participants were enrolled in the study by using eligibility criteria and a screening questionnaire"
Baseline imbalance	Low Risk	Quote: "Our study has some limitations. Parents' educational status, vaccination status, MUAC, and mean age of infants at baseline were different between the two groups; however, the mixed model and GEE analyses allowed us to adjust for these variables in the final models."
Loss of clusters	Low Risk	Comment: No clusters were lost
Unit of analysis issues	Low Risk	Quote: "We took a design effect of 1.25 and adjusted for cluster effect by multiplying the maximum sample of 59 with 1.25."

Sankaranarayanan 2005 (incl. Ramadas 2003, Ramadas 2010, Sankaranarayanan 2000, Sankaranarayanan 2013, Subramanian 2009)

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Panchayaths (municipal administrative units) CLUSTERS: Intervention: 7, Control: 6
Population	LOCAL TERM FOR CHW: Health workers RECRUITMENT CRITERIA: Non-medical university graduates HEALTHCARE SETTING: Household PATIENT POPULATION: Healthy participants aged 35 years and older INTERVENTION SAMPLE SIZE: 96517 CONTROL SAMPLE SIZE: 95356
Intervention	OBJECTIVE: Assess the effect of visual screening on oral cancer mortality INTERVENTION DESCRIPTION: Oral visual inspection by trained health workers at 3-year intervals. Screen-positive people were referred for clinical examination by doctors, biopsy, and treatment. KEY CHW COMPETENCY: Count the households in study clusters and the household members; explain the study to eligible participants; obtain informed consent; obtain information on socio-demographic factors and personal habits by interviewing eligible participants, give messages aimed at preventing and reducing tobacco and alcohol use, undertake oral and visual inspection, identify lesions suggestive of being precancerous in the oral cavity and identify oral cancer. TRAINING DURATION (DAYS): 90 TRAINING DESCRIPTION: Training organized by the Community Oncology Division of the Regional Cancer Centre (RCC), Trivandrum. Intervention HWs were taught about cancer in general, and oral cancer in particular, in terms of risk factors, natural history, methods of diagnosis, treatment, and prognosis by audio-visual methods. Two manuals on visual inspection with colour photographs and descriptions of oral lesions were used for training and reference during screening. After the theoretic teaching, they were trained in the field on home visits, enumeration of subjects, interview methodology, performing oral examination in individuals, and recording height, weight, blood pressure, and respiratory peak flow measurements. The faculty for the training sessions consisted of a dental surgeon, a

	pathologist, a cancer surgeon, a radiation oncologist, a family physician, a social scientist, and a demographer from the Kerala Census Department REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 01/1996 END DATE: 10/2004 CONTROL DESCRIPTION: Standard care. Although control participants identified during house visits to control clusters were not screened, they continued to receive routine awareness messages and to use health-care facilities as usual.
Outcomes	OUTCOMES: Survival, case fatality, oral cancer mortality, oral cancer incidence, mortality from all-causes, cancer or tobacco-related cancer, cost. DURATION OF FOLLOW UP: 8 years, 15 years
Funding	Association for International Cancer Research (AICR), St Andrews, UK.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "These clusters were allocated...by restricted (blocked) randomisation"
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified.
Blinding (performance and detection bias)	Low Risk	Quote: "Three physicians (KR, GT, ST) assessed information on death and were unaware of screening status."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Low Risk	Comment: All households screened
Baseline imbalance	Low Risk	Quote: "The study groups were well balanced as indicated by the similar distribution of age, sex, religion, and socioeconomic status"
Loss of clusters	Low risk	Comment: No cluster lost
Unit of analysis issues	Low Risk	Quote: "analysed and compared by use of cluster analysis"

Selke 2010

Trial reg.	None
Country	Kenya
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Sublocation CLUSTERS: Intervention: 8; Control 16
Population	LOCAL TERM FOR CHW: Community Care Coordinators (CCCs) RECRUITMENT CRITERIA: CCCs were chosen from the HIV clinic population at Mosoriot. All CCCs chosen had a secondary education. To be considered for this position, a patient had to be clinically stable with self-reported 100% adherence to ART over the 6 months before recruitment and considered by the clinic staff to be a good role model and mentor for other patients. HEALTHCARE SETTING: Household PATIENT POPULATION: HIV-infected adults clinically stable on antiretroviral therapy (ART) INTERVENTION SAMPLE SIZE: 96 CONTROL SAMPLE SIZE: 112

Intervention	OBJECTIVE: To assess whether community-based care delivered by people living with HIV/AIDS could replace clinic-based HIV care. INTERVENTION DESCRIPTION: The intervention group received monthly Personal Digital Assistant supported home assessments by CCCs with clinic appointments every 3 months. KEY CHW COMPETENCY: During home visits, CCCs obtained and entered data concerning patient's symptoms into their PDAs along with vital signs and an assessment of adherence to ART and opportunistic infection prophylaxis (derived from in-home pill counts). They were encouraged to comply with all alerts from their PDAs and to dispense a one-month supply of the patient's medications (from a prefilled kit) if the visit triggered no alerts that would require the patient to be seen immediately in clinic. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Once recruited, CCCs underwent structured didactic training, which included the use of a PDA that was pre-programmed to collect a symptom review, vital signs (temperature, weight, and pulse oximetry), adherence, food security, and domestic violence information. REFRESHER OR ONGOING TRAINING: After the didactic training, each CCC underwent 2 months of clinical mentoring. After completing mentoring, CCCs entered the field where they conducted home visits with patients within their assigned sublocations. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 03/2006 END DATE: 04/2008 CONTROL DESCRIPTION: The control group received standard of care monthly clinic visits.
Outcomes	OUTCOMES: Viral load, CD4 count, Karnofsky score, stability of ART regimen, opportunistic infections, pregnancies, and number of clinic visits. DURATION OF FOLLOW UP: 12 months
Funding	Doris Duke Charitable Foundation. This research was also supported in part by a grant to the United States Agency for International Development and the Academic Model for Providing Access to Healthcare programs Partnership from the United States Agency for International Development as part of the President's Emergency Plan for AIDS Relief (PEPFAR).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "For each stratum, community names were placed in sealed opaque envelopes with 2 envelopes being assigned to the control group for every envelope assigned to the intervention group. "
Allocation concealment (selection bias)	Low Risk	Quote: "[Opaque] envelopes were then opened, and the list of intervention and control sublocations was generated."
Blinding (performance and detection bias)	High Risk	Comment: Data collected by CCCs
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat, in which all study participants were regarded as randomly assigned to their respective study arm based on their sublocation... At 12 months, there were no significant intervention-control differences in the per cent withdrawn from study or lost to follow-up"
Selective reporting (reporting bias)	Low Risk	Comment: All outcomes specified in protocol reported save hospitalizations

Recruitment bias	Unclear Risk	Comment: Insufficient information to permit judgement
Baseline imbalance	Low Risk	Quote: "Demographic characteristics did not differ between the intervention and control groups. The mean age for both groups was in the late thirties, and approximately three quarters were female (Table 1). Enrolment clinical status also did not differ between the 2 groups, with no significant differences in the per cent of individuals with a history of WHO Stage 3/4 disease, detectable viral load, or mean Karnofsky score. There was a slight trend toward a higher mean CD4 count in the intervention group at enrolment, with a mean of 305 cells per microliter in the intervention group versus 278 cells per microliter in the control group, respectively ($P = 0.09$). With regard to composition of ART regimen, more than 90% of patients in both groups were receiving a nonnucleoside reverse transcriptase inhibitor-based regimen. The per cent of patients receiving tuberculosis and cryptococcal prophylaxis or treatment were also similar between the 2 groups."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Unclear Risk	Comment: Insufficient information to permit judgement

Siabani 2015

Trial reg.	ANZCTR 365532
Country	Iran
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Community health volunteers (CHVs) RECRUITMENT CRITERIA: The selected CHVs were those judged most experienced in health education and who had been participating effectively in health programs during the preceding year, among 50 CHVs connected to a health delivery centre in Kermanshah. The CHVs had to have acceptable conversation skills, be able to speak in both Persian and Kurdish, and have been actively involved during the preceding year in assisting at the health centre. HEALTHCARE SETTING: Household PATIENT POPULATION: The inclusion criteria were being able to be contacted during the following four months, being 18 years of age or older, not suffering from a serious mental or cognitive disorder, not having a disabling physical problem, and CHF being in New York Heart Association Functional Class II or III. INTERVENTION SAMPLE SIZE: 76 CONTROL SAMPLE SIZE: Control A: 78, Control B: 77
Intervention	OBJECTIVE: This study aimed to evaluate the effectiveness of a home-based educational strategy using CHVs in improving self-care of patients with CHF in comparison with an educational strategy using formal health professionals at hospital, and also with a control group receiving usual care in western Iran. INTERVENTION DESCRIPTION: Face-to-face education programme by CHVs at the patients' homes. Each CHV educated seven or eight patients, one at a time for about two hours, at the patients' homes. Data obtained through interviewing patients before and two months after interventions were analysed KEY CHW COMPETENCY: Provide face-to-face education programme for improving self-care of patients with chronic heart failure (CHF). TRAINING DURATION (DAYS): 3 TRAINING DESCRIPTION: Female CHVs (27–53 years old) were trained in three workshops by a general practitioner. CHVs were educated about self-care activities and how to present the information to patients. The contents and educational materials were obtained from heart failure textbooks and related websites that had been translated into simple Persian language understandable for patients, even those who were illiterate.

	REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: CHVs were provided with the contact number of a trained general practitioner in case the CHV needed to ask questions during a home visit. During the educational programs, one of the investigators monitored the home-based education and randomly checked 25% of the home visits, after the educational programme was finished, in order to check the way the intervention was conducted. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Community health volunteers making home visits were reimbursed for travel costs only. RATIO CHW TO POPULATION: 1:8 START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: A: An education programme using paid formal health providers at hospital. B: Usual care, which involved health practitioners visiting and prescribing medicine and possibly providing some advice about how to consume those medicines.
Outcomes	OUTCOMES: Self-care components, comprising self-care maintenance, self-care management and self-care confidence, were the primary outcomes. DURATION OF FOLLOW UP: 2 months
Funding	Kermanshah University of Medical Sciences and the University of Sydney

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Randomised using a simple random number list generated in Microsoft Excel"
Allocation concealment (selection bias)	Low Risk	Quote: "The allocation was made using only study number. Only after the allocation had been made were the subject numbers re-linked to study names."
Blinding (performance and detection bias)	Low Risk	Quote: "The first data collection occurred prior to randomisation. Subjects were aware of where the training took place and what sort of training they received, but were not aware of the study hypothesis. The nurses involved in the data collection undertaken two months after the intervention did not have other involvement in the study. They were blinded to intervention allocation and did not know the study hypothesis, but it is possible that the intervention allocation may have become known as a result of discussions during the data collection process."
Incomplete outcome data (attrition bias)	High Risk	Quote: "One hundred and eighty-four of the initial 231 participants (80%) completed the study by participating in the second interviews, which were held two months after intervention: 55 people (71.4%) from the [control B], 67 (85.9%) from the [control A] and 62 (81.6%) from those educated by CHVs...Females were more likely to withdraw...Those who withdrew were on average four years older than those who completed"
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes listed in the protocol were reported

Sloan 2008

Trial reg.	None
Country	Bangladesh
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Unions

	CLUSTERS: Intervention: 21, Control: 21
Population	LOCAL TERM FOR CHW: Community nutrition workers RECRUITMENT CRITERIA: Existing community nutrition workers from Bangladesh Integrated Nutrition Programme HEALTHCARE SETTING: Household PATIENT POPULATION: Women aged 12–50 years INTERVENTION SAMPLE SIZE: 2121 CONTROL SAMPLE SIZE: 2044
Intervention	OBJECTIVE: Test whether community-based kangaroo mother care reduces the overall neonatal mortality rate by 27.5%, infant mortality rate by 25%, and low birth weight neonatal mortality rate by 30%. INTERVENTION DESCRIPTION: Community-based workers were taught to teach community-based kangaroo mother care to all expectant and postpartum women KEY CHW COMPETENCY: Teach CKMC to mothers TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: A physician who had participated in the pilot study trained 12 BRAC supervisors and, along with 1 supervisor, trained all 63 community nutrition workers and their 25 National Nutrition Programme supervisors serving the intervention group in 5 groups of 6 to 22 people REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Community nutrition workers were supervised, though a description was not provided. INCENTIVE: Monetary INCENTIVE DESCRIPTION: All participating community workers received \$7.50 a month, approximately equivalent to their compensation for their normal half-time NNP duties, for attending the monthly standardization sessions and collecting newborn weight within 72 hours of birth. To avoid implementation bias, no additional compensation for teaching CKMC was provided to the nutrition workers. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Standard care
Outcomes	OUTCOMES: Birth weight, CKMC behaviours (STS, Breastfed, Still breastfeeding at neonatal visit, Breastfed within 1 h of birth, Breastfed immediately after birth, Hours after birth first breastfed, Consumed something other than breast milk, Bathed on date of birth, Age first bathed, Slept with infant in first 2 d of life, Took infant to clinic/hospital for well-care check-up in neonatal period, Took infant to clinic/hospital for perceived newborn illness, Took infant to clinic/hospital for well-care check-up in infancy, Took infant to clinic/hospital for perceived infant illness), infants' weight, head or arm circumference, or reported morbidity, NMR, infant mortality, maternal outcomes DURATION OF FOLLOW UP: 1 year
Funding	Save the Children (USA) Saving Newborn Lives Through the Bill & Melinda Gates Foundation, and USAID and for the baseline survey from the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, Save the Children (USA) Saving Newborn Lives, and Population Council.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Half of the unions were randomly assigned to receive CKMC and half not." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not specified.
Blinding (performance and detection bias)	High Risk	Quote: "All community nutrition workers serving both the intervention and control group villages were responsible for obtaining birth weight "

Incomplete outcome data (attrition bias)	High Risk	Quote: "Intention-to-treat analyses (according to the groups to which the participants' villages and thus the participants were assigned) are presented. Intention-to-treat analyses include all CKMC infants regardless of whether they received the intervention." "Forty per cent overall and ~65% of newborns who died were not weighed at birth, and missing birth weight was differential by study group"
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Low Risk	Quote: "An independent research organization, Mitra and Associates, conducted a baseline survey of all households in the sample villages to identify eligible women (aged 12–50 years) and obtain their demographic and reproductive health information. Thereafter, consenting women who were identified to be pregnant through quarterly household surveillance were enrolled during 15 months. "
Baseline imbalance	High Risk	Quote: "Of the 8 study villages in Sylhet division (contributing to 11% of the total sample), only 2 were randomly allocated to the intervention group...the sample villages in Dhaka division are...somewhat socio-demographically distinct from those in Sylhet division, where villages are generally smaller, the population is less literate, and slightly more Hindus live. This imbalance is reflected in small differences in previous NMR."
Loss of clusters	Low risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote "terms were included for...generalized estimating equations (GEE) with robust variance estimation in adjusted analyses to account for design effect as a result of stratification and clustered allocation."

Soofi 2012

Trial reg.	NCT01192789
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Union councils, the smallest administrative unit of the district CLUSTERS: Intervention: 9, Control: 9
Population	LOCAL TERM FOR CHW: Lady health workers (LHWs) RECRUITMENT CRITERIA: Public-sector LHWs--Women, most of whom live in the local communities, with at least 8 years of formal schooling HEALTHCARE SETTING: Household PATIENT POPULATION: Age 2–59 months presenting to LHWs with severe pneumonia INTERVENTION SAMPLE SIZE: 2341 CONTROL SAMPLE SIZE: 2069
Intervention	OBJECTIVE: Aimed to establish whether community case identification and management of severe pneumonia by oral antibiotics delivered through community health workers has the potential to reduce the number of infants dying at home. INTERVENTION DESCRIPTION: LHWs undertook community case management of WHO-defined severe pneumonia. The children in intervention clusters with suspected pneumonia were screened by LHWs and those diagnosed with severe pneumonia were prescribed oral amoxicillin syrup (90 mg/kg per day in two doses) for 5 days at home. In both groups, follow-up visits at home were done at days 2, 3, 6, and 14 by LHW. KEY CHW COMPETENCY: Screened children for pneumonia and prescribe oral amoxicillin syrup

	TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: In collaboration with the national LHW programme managers, the study developed an enhanced acute respiratory infections training module and training video to emphasise classification of acute respiratory infections according to the WHO algorithm, identification of danger signs, and home pneumonia case management with oral amoxicillin. The faculty members of the Department of Paediatrics at Aga Khan University (Karachi, Pakistan) led these training sessions with trainers of the LHW programme. The LHWs in control clusters received a refresher training course that consisted of existing acute respiratory infection modules of national LHW programme curriculum. REFRESHER OR ONGOING TRAINING: Monthly refresher sessions SUPERVISION FREQUENCY: SUPERVISION DESCRIPTION: "The variation in LHW coverage is indicative of variations in health-system performance and is related to wide variations in LHW residence status, working hours, and local supervision." INCENTIVE: Monetary INCENTIVE DESCRIPTION: A monthly salary of 3800 Pakistani rupees (about US\$45). LHWs were reimbursed for additional travel costs arising from the study follow-up visits but did not receive any additional salary or other financial incentives. RATIO CHW TO POPULATION: 1:1000-1500 (~100 families) START DATE: 13/02/2008 END DATE: 15/03/2010 CONTROL DESCRIPTION: Children in control clusters were given one dose of oral cotrimoxazole and were referred to their nearest health facility for admission and intravenous antibiotics, as per government policy. Follow-up visits at home were done at days 2, 3, 6, and 14 by LHW
Outcomes	OUTCOMES: Primary: treatment failure up to day 6 after enrolment. Secondary: treatment failure or relapse between day 7 and day 14 in children who were well on day 6. DURATION OF FOLLOW UP: 2 weeks
Funding	USAID through grants to John Snow Incorporation and Boston University, USA.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Clusters were randomly allocated to intervention and control groups with a computer-generated randomisation sequence that was generated by an independent expert at WHO"
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Unclear Risk	Comment: Insufficient information to permit judgement
Incomplete outcome data (attrition bias)	Low Risk	Comment: Though analyses were done per-protocol, attrition was 0%
Selective reporting (reporting bias)	Low Risk	Comment: Primary outcome indicated in study registration was reported as the primary outcome.
Recruitment bias	Unclear Risk	Comment: Insufficient information to permit judgement
Baseline imbalance	High Risk	Quote: "At baseline, children seen by LHWs differed on several key demographic and clinical indicators that could have affected treatment failure."
Loss of clusters	Low Risk	Quote: "There was no exclusion and no attrition of clusters during the study."
Unit of analysis issues	Low Risk	Quote: "We calculated cluster-specific failure rates for each cluster and compared the mean differences between treatment and control groups. We also calculated

	individual-level risk differences, adjusting for clustering with a generalised estimating equation with an exchangeable correlation."
--	---

Stanton 2013

Trial reg.	NCT01108289
Country	Ghana
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Births attended by one CHO CLUSTERS: Intervention: 28, Control: 26
Population	LOCAL TERM FOR CHW: Community health officers (CHO) RECRUITMENT CRITERIA: CHOs are employees of the Ghana Health Services, trained for two years to provide outreach health services to the rural population, including childhood immunization, antenatal and postnatal care, and family planning. CHOs do not have midwifery skills. The majority of CHOs in this trial were recent CHO training programme graduates. HEALTHCARE SETTING: Household PATIENT POPULATION: Pregnant women INTERVENTION SAMPLE SIZE: 689 CONTROL SAMPLE SIZE: 897
Intervention	OBJECTIVE: Oxytocin is the drug of choice for prevention of PPH. Its use has generally been restricted to medically trained staff in health facilities. The study assessed the effectiveness, safety, and feasibility of PPH prevention using oxytocin injected by peripheral health care providers without midwifery skills at home births. INTERVENTION DESCRIPTION: One injection of oxytocin one minute after birth by CHO. The CHOs did not in any way manage the delivery; they (1) observed, (2) provided the intervention (according to random allocation), (3) measured the outcomes, and (4) provided early treatment for PPH and facilitated referral, when needed. KEY CHW COMPETENCY: Provide one injection of oxytocin [10 IU] one minute after birth TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: All CHOs were required to successfully complete an allocation-specific, five-day competency-based assessment and were trained to palpate the uterus for a twin prior to administration of oxytocin. Oxytocin in Uniject supplies were kept refrigerated at KHRC and distributed approximately monthly in small quantities to CHOs, who stored them outside of a temperature-controlled cold chain at their CHO compound. CHOs requested additional supplies as needed from Field Supervisors. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 21/04/2011 END DATE: 30/11/2012 CONTROL DESCRIPTION: No provision of prophylactic oxytocin. The control group CHOs did not provide this prophylactic injection to women they observed. All other CHO activities (outcome measurement, data collection, and early treatment and referral when necessary) were identical across the control and oxytocin CHOs.
Outcomes	OUTCOMES: Primary: PPH. Secondary: safety and feasibility of home-based PPH prevention. DURATION OF FOLLOW UP: 19 months
Funding	PATH with funding from the Bill & Melinda Gates Foundation. Collaborators include the Johns Hopkins Bloomberg School of Public Health, the Kintampo

	Health Research Center, and Research Triangle Institute.	
<i>Risk of Bias</i>		
Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "The randomisation sequence was determined using Stata (version 12). The random allocation sequence was generated by LCM, while enrolment of the clusters was done by SN and KHRC in collaboration with the GHS. "
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	High Risk	Quote: "The random allocation was not masked because the efficacy of oxytocin is not in doubt, use of a placebo was unacceptable to the GHS, and filling the Uniject device with saline has prohibitive testing requirements...Two to three days following delivery, field workers conducted separate follow-up interviews with all enrolled women and their traditional birth attendants (if one was present at birth) at their respective homes. These interviews were conducted as an independent assessment of the care provided to avoid relying on CHOs' documentation of their own behaviour."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis of primary and secondary outcomes was conducted following an intent-to-treat approach. "
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes reported in protocol were reported
Recruitment bias	Low Risk	Quote: "KHRC field workers were responsible for the identification of all pregnant women. At seven or more months' gestation, field workers sought written initial informed consent for the trial. Women agreeing to participate, should they choose to deliver at home, were asked to contact the field worker after the onset of labour...Field workers identified 5,919 pregnant women at seven months' gestation or more, nearly all (99.9%) of whom gave initial consent to participate in the study. "
Baseline imbalance	High Risk	Quote: "The randomisation process achieved balance between oxytocin and control clusters on a range of women's characteristics and on indicators of recruitment, enrolment, and measurement procedures. Two CHO behaviours differed between oxytocin and control groups. Among CHOs called by women prior to birth, oxytocin CHOs referred women to hospital prior to delivery in 198 of 1,148 (17.3%) deliveries as compared to 192 of 1,539 (12.5%) deliveries referred by control CHOs. Likewise, in 58 of 689 (8.4%) enrolled oxytocin deliveries and 278 of 897 (31.1%) enrolled control deliveries, CHOs tied on the drape after delivery of the baby and before delivery of the placenta, as opposed to before delivery of the baby, per protocol. The mean duration between birth and unfolding the drape, which could potentially affect blood-loss measurement, confirms a delay of one minute on average by control CHOs relative to oxytocin CHOs (4.4 minutes [SD: 5.7] versus 3.4 minutes [SD: 5.4], respectively)."
Loss of clusters	High Risk	Quote: "Two additional CHO catchment areas were selected to replace non-compliant CHOs both of which were in the oxytocin group...The new catchment areas were adjacent to those of the non-compliant CHOs with

		similar characteristics."
Unit of analysis issues	Low Risk	Comment: ICCs provided

Swart 2008

Trial reg.	None
Country	South Africa
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Blocks of households CLUSTERS: Intervention: 26, Control: 26
Population	LOCAL TERM FOR CHW: Home visitors (HV) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Households with children aged ≤ 10 years INTERVENTION SAMPLE SIZE: 208 households CONTROL SAMPLE SIZE: 202 households
Intervention	OBJECTIVE: To evaluate the effectiveness of the HV programme in reducing hazards for burns, poison ingestion, and fall-related injuries for low-income neighbourhood households with children aged ≤ 10 years. INTERVENTION DESCRIPTION: The HV programme was implemented in two low-income communities in South Africa. Intervention households received four visits, one every two weeks, by trained paraprofessionals that focused on a specific injury topic and consisted of: information dissemination about specific injury prevention practices; home inspection accompanied by information about home hazards; and the supply of safety devices. Each visit lasted approximately 40 minutes. KEY CHW COMPETENCY: Information dissemination about specific injury prevention practices; home inspection accompanied by information about home hazards; and the supply of safety devices. TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: All recruits received training in unintentional injury prevention methods. A manual for the home visitors training was developed and tested by the research team during the pilot, and was further refined for this trial REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Daily SUPERVISION DESCRIPTION: Two site coordinators met with the home visiting teams daily to provide debriefing and supervision. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Home visitors were paid ZAR30 (approximately US\$4) per visit. RATIO CHW TO POPULATION: N.S. START DATE: 10/2005 END DATE: 02/2006 CONTROL DESCRIPTION: Control households were given safety devices after the injury risk post-assessment.
Outcomes	OUTCOMES: Key outcomes developed to measure the presence of household hazards were scores for burns (safety practices, paraffin, and electrical), poisoning, and falls DURATION OF FOLLOW UP: 4 months
Funding	N.S.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Thereafter, eligible blocks were randomly allocated to the intervention and control groups (14 each for Nomzamo and 12 each for Slovo Park" Comment: Unclear how this was done.
Allocation concealment	Unclear Risk	Comment: Not mentioned

(selection bias)		
Blinding (performance and detection bias)	High Risk	Quote: "While data collectors were masked to group assignment at baseline, they might have been alerted to the intervention houses at post-intervention."
Incomplete outcome data (attrition bias)	Low Risk	Comment: 8% LTFU
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol available
Recruitment bias	Unclear Risk	Quote: "They were given computer generated lists of all house numbers in each block, sorted in random order, and instructed to select houses from the top down until the required total of eligible households were recruited from each site. Before a home was considered non-participating it was visited at least three times. After obtaining informed written consent from participating homes, data collectors implemented the baseline injury risk assessment."
Baseline imbalance	Low Risk	Quote: "The socio-demographic characteristics of the intervention and control households were similar except that more respondents from the intervention than control group were the children's primary caretaker...At baseline, the risk profiles were similar for both groups."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "To allow for the cluster design an inflation factor of 1.2, estimated from the pilot study, was used, indicating a sample size of 120 houses per intervention group."

Tian 2015

Trial reg.	NCT01503814
Country	Tibet (China) and India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Villages CLUSTERS: Intervention 23; Control: 24
Population	LOCAL TERM FOR CHW: Community health workers (CHWs) RECRUITMENT CRITERIA: In rural China, CHWs are village doctors who are not physicians but have received basic health professional training and are able to prescribe medications. In India, the CHWs were volunteer members of the community and not authorized to prescribe medicine. HEALTHCARE SETTING: Household/community PATIENT POPULATION: Individuals with high cardiovascular risk (aged ≥ 40 years with self-reported history of coronary heart disease, stroke, diabetes mellitus, and/or measured systolic blood pressure ≥ 160 mmHg). INTERVENTION SAMPLE SIZE: 1095 CONTROL SAMPLE SIZE: 991
Intervention	OBJECTIVE: This study aims to develop and evaluate a simplified cardiovascular management programme delivered by community health workers with the aid of a smartphone-based electronic decision support system. INTERVENTION DESCRIPTION: Participants in the intervention villages were managed by CHWs through an Android-powered app on a monthly basis focusing on the appropriate prescription of 2 medications (blood pressure-lowering agents and aspirin) and 2 lifestyle modifications (smoking cessation and salt reduction). KEY CHW COMPETENCY: Manage individuals at high CVD risk TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Before the start of the intervention, all CHWs in the intervention group received initial systematic training. REFRESHER OR ONGOING TRAINING: CHWs received refresher training every 3 to 4 months during the implementation of the intervention.

	SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: CHWs were supported by a partnership with licensed physicians in healthcare centres responsible for providing clinical guidance and prescribing medications. INCENTIVE: Monetary INCENTIVE DESCRIPTION: Performance feedback and performance-based incentives to the CHWs were also part of the intervention package in China, whereas Indian CHWs received fixed monthly honorariums. On average, the payment for CHWs in the intervention group for the entire duration of the intervention was US\$400 in China and US\$500 in India. RATIO CHW TO POPULATION: N.S. START DATE: 01/2012 END DATE: 003/2014 CONTROL DESCRIPTION: In the control group, the usual cardiovascular management programs continued without additional intervention. However, in India, as stipulated by the ethics committee, the medications used in the intervention arm were also made available free of charge to patients in the primary care facilities serving the control villages, typically located 15 to 40 minutes by motorcycle from the villages.
Outcomes	OUTCOMES: Primary: net difference between groups in the change in the proportion of patient-reported antihypertensive medication use from baseline to 1-year follow-up. Secondary: between-group differences in the changes from baseline to 1-year follow-up in (1) the proportion of high-risk individuals taking aspirin, (2) the mean systolic blood pressures of high-risk individuals, (3) the proportion of current smokers, (4) the proportion of high-risk individuals aware of the harms of a high-salt diet, (5) the proportion of high-risk individuals receiving monthly follow-ups from the CHWs, and (6) the proportion of high-risk individuals hospitalized. DURATION OF FOLLOW UP: 12 months
Funding	This study was funded in part with Federal funds from the US National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN268200900027C. Additional support was received from the UnitedHealth Group Chronic Disease Initiative

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Villages were randomly assigned to intervention or control with stratification by country. In China, 27 villages were randomly assigned to the intervention and control group with stratification by county and township. Villages in India were randomly assigned without any stratification. A study-independent staff generated the randomisation pattern through a central computerized process."
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned.
Blinding (performance and detection bias)	Low Risk	Quote: "single-blinded"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Intention to treat analysis...To further investigate the influence of loss to follow-up on the results, we compared all 15 available baseline characteristics between the subjects who were followed up and those lost to follow-up. We found that there were significant differences in 3 variables: age, body mass index, and sex; however, the absolute differences for these variables were not large"
Selective reporting (reporting bias)	Low Risk	Comment: Outcomes noted in the trial registration were reported.
Recruitment bias	Unclear Risk	Quote: "2086 individuals were identified as high-risk and recruited for the purposes of this study"

Baseline imbalance	Low Risk	Quote: "In China, there were no statistical differences between subjects in the intervention group in comparison with those in the control group. In India, with the exceptions of the history of coronary heart disease ($P<0.001$) and history of diabetes mellitus ($P=0.004$), no significant differences were found..."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "The intracluster correlation coefficient at baseline for the primary outcome of patient-reported antihypertensive medication use was 0.02. "

Tine 2011 (incl. Tine 2014)

Trial reg.	None
Country	Senegal
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: CHW CLUSTERS: Intervention: 4, Control: 4
Population	LOCAL TERM FOR CHW: Community health worker (CHW) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Children less than 10 years INTERVENTION SAMPLE SIZE: 500 CONTROL SAMPLE SIZE: 500
Intervention	OBJECTIVE: A randomised controlled trial was carried out to assess the impact of combining seasonal IPTc with home-based management of malaria by CHWs in Senegal. INTERVENTION DESCRIPTION: Eight CHWs in the eight villages around the Bonconto health post, were trained to diagnose malaria using RDT and provide prompt treatment with AL to children less than 10 years. Intervention CHWs also administered monthly IPTc with single dose of SP plus three doses of AQ. KEY CHW COMPETENCY: Diagnose malaria using RDT, provide prompt treatment with AL to children less than 10 years, administer monthly IPTc with single dose of SP plus three doses of AQ TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 10/2010 END DATE: 11/2010 CONTROL DESCRIPTION: Diagnose malaria using RDT and provide prompt treatment with AL to children less than 10 years only
Outcomes	OUTCOMES: Primary: incidence of single episode of malaria attacks over 8 weeks of follow up. Secondary: prevalence of malaria parasitaemia, and prevalence of anaemia at the end of the transmission season, malaria diagnostic accuracy; access to ACT treatment; SMC coverage; safety and drug tolerability. DURATION OF FOLLOW UP: 2 months
Funding	Malaria Capacity Development Consortium (MCDC)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "The CHWs were randomised using a random number generator from Excel software."

Allocation concealment (selection bias)	Unclear Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	High Risk	Comment: Primary data collected by CHWs
Incomplete outcome data (attrition bias)	Low Risk	Quote: "For the primary end point of incidence of malaria attacks over 8 weeks of follow up, analysis was by intention to treat including all children who attended the baseline survey. For the secondary end points of cross sectional prevalence of malaria and anaemia at the end of the transmission season, analysis was done by per protocol, including all children seen at cross sectional survey at the end of the transmission season."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified
Recruitment bias	Unclear Risk	Comment: Insufficient information to permit judgement
Baseline imbalance	Low Risk	Quote: "At baseline the two groups were similar in terms of demographic characteristics (age, gender and P. falciparum carriage). Prevalence of moderate anaemia and severe anaemia were similar in the two groups, as well as prevalence of under nutrition (stunting, underweight) [and bed net usage]"
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "To assess the impact of combining [home management] + IPTc, analysis was conducted at the individual level with adjustment for clustering using robust standard errors"

Tomlinson 2014 (incl. Ijumba 2015)

Trial reg.	ISRCTN41046462
Country	South Africa
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: "Subplaces" in the area CLUSTERS: Intervention: 15; Control: 15
Population	LOCAL TERM FOR CHW: Community health workers (CHWs) RECRUITMENT CRITERIA: CHWs were women residing in the neighbourhoods, who had at least at least grade-10 education level and were conversant in English and first language isiZulu. HEALTHCARE SETTING: Household PATIENT POPULATION: Pregnant women aged 17 years or older and their newborns residing in the clusters during the recruitment period. INTERVENTION SAMPLE SIZE: 1894 CONTROL SAMPLE SIZE: 2243
Intervention	OBJECTIVE: Evaluate an integrated, scalable package providing two pregnancy visits and five post-natal home visits delivered by CHWs INTERVENTION DESCRIPTION: Women in the intervention arm were scheduled to receive seven home-based visits: two during pregnancy and one within 48 h of delivery, during days 3–4 and 10–14, during weeks 3–4 and 7–8. Low birth weight neonates (≤ 2500 g) were to receive two extra visits within the first week. KEY CHW COMPETENCY: Provide a structured home visiting programme consistent with existing PMTCT, IMCI, lactation counselling and newborn care guidelines in which motivational interviewing techniques were used for breastfeeding counselling. TRAINING DURATION (DAYS): 10 TRAINING DESCRIPTION: Trained on home entry, motivational interviewing techniques, breast-feeding, disclosure, antenatal care, infant feeding with emphasis on exclusive breast-feeding, breast problems and diseases, interaction with newborns, baby blues and postnatal depression, and neonatal care, including danger

	signs in newborns and their mothers that might warrant a referral. The training was based on a manual compiled by the principal investigators drawing on several resources, including consultation with designated training authorities in South Africa and the WHO/UNICEF Breastfeeding Counselling Course. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Weekly SUPERVISION DESCRIPTION: Records of all births in this hospital were checked on a daily basis and CHW supervisors contacted the CHW via a mobile phone system when an intervention group participant gave birth. Attempts were made to complete all visits regardless of how many earlier visits might have been missed. CHW received feedback on their work from the supervisors during weekly supervision meetings. INCENTIVE: Monetary INCENTIVE DESCRIPTION: All CHWs were paid a salary in line with existing South African government stipends for CHWs (R 3500/\$US 300 per month). RATIO CHW TO POPULATION: START DATE: 01/06/2008 END DATE: 01/12/2010 CONTROL DESCRIPTION: In control clusters, CHWs provided information and support on accessing social welfare grants and conducted three home-based visits: one during pregnancy and two during weeks 4–6 and 10–12 post-delivery.
Outcomes	OUTCOMES: Primary: were levels of exclusive and appropriate infant feeding and levels of HIV-free survival at 12 weeks postnatally. Secondary: (i) uptake of a post-natal clinic visit within 7 days of life, (ii) coverage of care and behavioural indicators (antenatal HIV testing, uptake of cotrimoxazole amongst HIV-exposed infants at 12 weeks, family planning uptake at 12 weeks, 2 week prevalence of diarrhoea at 12 weeks, Use of dual protection methods for family planning (%)), (iii) post-intervention levels of maternal depressed mood at 12 weeks post-partum. Other: Preparations made for delivery (%), Knowledge of newborn danger signs: median score out of 8(IQR), Neonatal mortality (%), Initiation of breastfeeding within one hour of birth (%) Appropriate feeding in HIV-positive moms (exclusive breastfeeding or no breast feeding) (%), Infant weight-for-length z-scores at 12 weeks (mean, standard deviation), Infant weight-for-age z-scores at 12 weeks mean (mean, standard deviation), Infant length-for-age z-scores at 12 weeks (mean, standard deviation), Hospitalisation for diarrhoea (%) Mother disclosed status (%), Outcomes amongst HIV-positive women, Mother-to-child HIV transmission at 12 weeks (%) Infant HIV tested at 6 weeks (%) DURATION OF FOLLOW UP: 12 weeks post-natal
Funding	Save the Children Federation, Inc. (USA), South African Medical Research Council (sub-grant # 354).

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "We used simple computer-generated randomisation, with clusters assigned in a 1:1 allocation ratio. "
Allocation concealment (selection bias)	Unclear Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: " Trained data collectors blinded to arm visited "
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Intention to treat analysis was performed."
Selective reporting (reporting bias)	Low Risk	Comment: Primary outcomes listed in protocol were reported in the trial
Recruitment bias	Low Risk	Quote: "In intervention and control clusters, trained CHWs identified all pregnant women living in the cluster and informed them of the study."
Baseline imbalance	Low Risk	Quote: "A baseline survey was conducted to establish cluster homogeneity between 'subplaces' in the area

		(Figure 1). No significant differences were noted, and therefore, neither stratification nor matching was performed."
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote : "To adjust for the cluster design, the standard errors of the statistical effect measures of both regression models used were estimated using a robust cluster variance approach."

Tripathy 2010 (incl. Houweling 2013)

Trial reg.	ISRCTN21817853
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Clusters (parts of districts) CLUSTERS: Intervention: 18, Control: 18
Population	LOCAL TERM FOR CHW: Facilitator RECRUITMENT CRITERIA: Women speaking the local language and having the ability to travel to meetings; identified by the community. HEALTHCARE SETTING: Community PATIENT POPULATION: Women aged 15–49 years, residing in the project area, having given birth during the study. INTERVENTION SAMPLE SIZE: 8662 mothers, 9686 births CONTROL SAMPLE SIZE: 8125 mothers, 9089 births
Intervention	OBJECTIVE: Assess community mobilisation through participatory women's groups INTERVENTION DESCRIPTION: Facilitator convened 13 groups every month to support participatory action and learning for women, and facilitated the development and implementation of strategies to address maternal and newborn health problems. Every group met monthly for a total of 20 meetings. KEY CHW COMPETENCY: Use methods such as picture-card games, role play, and story-telling to help discussions about the causes and effects of typical problems in mothers and infants, and devised strategies for prevention, home- care support, and consultations TRAINING DURATION (DAYS): 7 TRAINING DESCRIPTION: Residential training course to review the cycle's contents, and to practice participatory communication techniques REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: District coordinators met with facilitators to offer support INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 31/07/2005 END DATE: 30/07/2008 CONTROL DESCRIPTION: Standard care
Outcomes	OUTCOMES: Primary: reductions in NMR and maternal depression scores. Secondary: stillbirths, maternal and perinatal deaths, uptake of antenatal and delivery services, home-care practices during and after delivery (Any antenatal care, Three or more antenatal visits, Any iron, tablets during pregnancy, Institutional delivery, Home delivery, Delivery attended by traditional birth attendant, Birth attendant washed hands, Birth attendant used safe delivery kit, Birth attendant used plastic sheet, Cord cut with new or boiled blade, Cord tied with boiled thread, Infant wiped within 30 min, Infant wrapped within 30 min, Infants alive at 1 month, Exclusive breastfeeding for 6 weeks) and health-care-seeking behaviour, equity impact in terms of NMR among lower and higher socio-economic groups. DURATION OF FOLLOW UP: 3 years
Funding	Health Foundation, UK Department for International Development, Wellcome

	Trust, and the Big Lottery Fund (UK).	
<i>Risk of Bias</i>		
Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "An external observer from a partner non-governmental organisation...drew folded papers with numbers corresponding to clusters with existing groups from a basket. The first four clusters were allocated to the intervention group, the rest to the control group."
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "Because of the nature of the intervention, neither the intervention team nor the participants were masked to group assignment during the trial...the intervention and surveillance teams were not unaware of allocation."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat...We excluded data from mothers who migrated out of the region and their infants from intention-to-treat analyses since many of these women probably came into the clusters at the time of delivery and would therefore not have been exposed to the intervention in pregnancy."
Selective reporting (reporting bias)	Low Risk	Comment: From the trial registration to the paper, maternal mortality was demoted to a secondary outcome and maternal depression was promoted to a primary outcome, but all outcomes listed were reported.
Recruitment bias	Unclear Risk	Quote: "The study population was an open cohort—i.e., women could enter the study at any time during the trial period if they had given birth." Comment: Unclear how women were approached to participate
Baseline imbalance	Low Risk	Quote: "Numbers of births were similar in intervention and control clusters, but differences were noted in household assets, maternal education, literacy, and tribal membership, with women in the intervention clusters being generally poorer and more disadvantaged than those in the control clusters" Comment: Trial adjusted for these differences
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Analysis was...at cluster and participant levels."

Tripathy 2016

Trial reg.	ISRCTN31567106
Country	India
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Clusters CLUSTERS: Intervention: 15 Control: 15
Population	LOCAL TERM FOR CHW: Accredited Social Health Activists (ASHAs) RECRUITMENT CRITERIA: In each cluster, the study team approached ASHA supervisors, village health sanitation and nutrition committees, and existing women's groups, and asked them to identify ASHAs to facilitate group meetings. The 152 ASHAs selected through this process were typical of ASHAs working in the study area; they had undergone at least three government-approved training sessions and were all village-based. About 60% of them had between 6 and 9 years of education, and 70% of them were adivasi (indigenous) HEALTHCARE SETTING: Community PATIENT POPULATION: Study participants were women of reproductive age (15–49 years) who gave birth between Sept 1, 2009, and Dec 31, 2012

	INTERVENTION SAMPLE SIZE: 82702 (3700 births) CONTROL SAMPLE SIZE: 73817 (3519 births)
Intervention	OBJECTIVE: Test the effect of participatory women's groups facilitated by ASHAs on birth outcomes, including neonatal mortality. INTERVENTION DESCRIPTION: ASHAs supported women's groups through a participatory learning and action meeting cycle. Groups discussed and prioritised maternal and newborn health problems, identified strategies to address them, implemented the strategies, and assessed their progress KEY CHW COMPETENCY: Lead women in a learning action cycle TRAINING DURATION (DAYS): 11 TRAINING DESCRIPTION: Training by Ekjut (NGO) staff in addition to previous government training. The training was divided into four phases of 3–4 days each, which was acceptable to National Health Mission staff. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Coordinators employed by Ekjut (NGO) supervised about ten ASHAs each through bimonthly meetings in the first 3 months of the intervention, and meetings every month thereafter. INCENTIVE: Monetary INCENTIVE DESCRIPTION: ASHAs were given an incentive of INR 100 (US\$2) during training and INR 200 for each woman's group meeting. RATIO CHW TO POPULATION: N.S. START DATE: September, 2009 END DATE: December, 2012 CONTROL DESCRIPTION: No women's groups
Outcomes	OUTCOMES: Primary: NMR. Secondary: stillbirths, perinatal mortality, maternal mortality, and home care and care-seeking practices. DURATION OF FOLLOW UP: 2 years
Funding	Big Lottery Fund (UK)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "public randomisation meeting"
Allocation concealment (selection bias)	Low Risk	Quote: "In each district, we invited local stakeholders, allocated a number to each cluster, wrote these numbers on small plastic balls, and placed the balls in a dark bag. We asked each participant to draw one ball from the bag and read out the cluster numbers out until all balls had been picked. We wrote the cluster numbers on a sheet of paper in order of selection. Participants then placed 20 pieces of paper numbered 1 to 20—each corresponding to a unique allocation sequence generated by an independent statistician—in the dark bag, and asked a participant to select a paper and read out the number. The corresponding sequence was then used to publicly allocate each cluster to one of two groups."
Blinding (performance and detection bias)	Low Risk	Quote: "The data collection team was masked to allocation, both at the cluster and at the individual level."
Incomplete outcome data (attrition bias)	High Risk	Comment: Analyses for the primary and secondary outcomes were by intention-to-treat but one cluster (pop 6203) was lost. Loss to follow-up as a result of migration or refusal to be interviewed was 35 (<1%) of 3735 births in the intervention clusters and 28 (<1%) of 3547 births in the control clusters.
Selective reporting (reporting bias)	Low Risk	Quote: Primary outcome listed in trial registration was reported

Recruitment bias	Low Risk	Comment: All qualifying women in included villages were asked to be part of the trial
Baseline imbalance	Low Risk	Quote: "Mothers in intervention and control groups had similar socioeconomic characteristics, although slightly more had received secondary education in the control group...than in the intervention group"
Loss of clusters	High Risk	Quote: "One intervention cluster was lost to follow up"
Unit of analysis issues	Low Risk	Quote: "We then used logistic regression with random effects to estimate the effect of the intervention on mortality outcomes, adjusting for clustering. We used generalised estimating equations to estimate effects on home care and care-seeking practices, because these had high intracluster correlation coefficients."

Tylleskar 2011 (incl. Birungi 2015, Chola 2011, Engebretsen 2014, Nankabirwa 2011, Nkonki 2014)

Trial reg.	NCT00397150
Country	Burkina Faso, Uganda, South Africa
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Communities CLUSTERS: Burkina Faso (BF): Intervention 12, Control 12 Uganda (Ug): Intervention: 12, Control: 12 South Africa (SA): Intervention: 17, Control: 17
Population	LOCAL TERM FOR CHW: Peer counsellor RECRUITMENT CRITERIA: BF: all the peer counsellors had motherhood (including breastfeeding) experience based on the community selection; their mean age was 30 years. Three quarters of them could write (French or the native Dioula language) and no one had more than 2-6 years of primary schooling. Ug: To be selected, a woman had to be between 18 and 45 years and to reside in the area with no plans of leaving the area within 2 years. She had to have a good reputation in the community. Furthermore, she had to be literate in the local language, willing to participate in the trial including a one-week residential training, and to undertake home visits in order to help women breastfeed their babies. Previous personal experience of breastfeeding was an additional advantage, although not a strict inclusion criterion. SA: Personal, formal and practical criteria were set in the selection process: 12 year of schooling, owning a bank account, good reputation in the community, previous community involvement and personal positive characteristics (empathy, patience, etc.) Formal tests and an interview were conducted prior to an observation during role plays. HEALTHCARE SETTING: Household PATIENT POPULATION: Mother-infant pairs INTERVENTION SAMPLE SIZE: BF: 392, Ug: 396, SA: 535, Total: 1323 CONTROL SAMPLE SIZE: BF: 402, Ug: 369, SA: 485, Total: 1256
Intervention	OBJECTIVE: The effect of breastfeeding counselling by peer counsellors was assessed in Africa. INTERVENTION DESCRIPTION: In the intervention group, one antenatal breastfeeding peer counselling visit and four post-delivery visits by trained peers were scheduled. The women were supposed to be visited by the peer counsellor once during the antenatal period, and at 1, 4, 7 and 10 weeks post-delivery. Mothers with breastfeeding problems were offered extra visits, as well as mothers who so KEY CHW COMPETENCY: individual home-based EBF peer counselling. TRAINING DURATION (DAYS): 7

	TRAINING DESCRIPTION: The core components of the curriculum included sessions on: Communication and counselling skills, composition of breast milk and the importance of breastfeeding, how milk is produced and released by the breast, helping a mother with positioning herself, attachment of the baby to the breast, breast conditions, expressing and storing breast milk, normal stools and urination, practicing exclusive breastfeeding, common feeding difficulties, role plays, clinical practice sessions REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Fortnightly SUPERVISION DESCRIPTION: The peer counsellors were followed up fortnightly by supervisors, and each peer counsellor was observed monthly to obtain feedback and improve counselling skills. In addition, the supervisors organised monthly group meetings with the peer counsellors. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: START DATE: 05/2006 END DATE: 07/2008 CONTROL DESCRIPTION: Mothers and infants in control clusters in Burkina Faso and Uganda were given standard health care only, and those control clusters in South Africa were visited by peer counsellors, with the same schedule as in the intervention clusters, but who assisted families in obtaining birth certificates and social welfare grants. The peer counsellors for the intervention and control clusters in South Africa were kept separate during the study.
Outcomes	OUTCOMES: EBF and diarrhoea prevalence. Other: weight-for-length (WLZ), length-for-age (LAZ) and weight-for-age (WAZ) z-scores, cost (Uganda), cost (South Africa), prevalence of early childhood caries assessed by the decayed missing filled teeth index at the five year follow-up), parasitaemia DURATION OF FOLLOW UP: 6 months
Funding	European Union Sixth Framework International Cooperation-Developing Countries, Research Council of Norway, Swedish International Development Cooperation Agency, Norwegian Programme for Development, Research and Education, South African National Research Foundation, and Rockefeller Brothers Foundation.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "24 communities in Burkina Faso, 24 in Uganda, and 34 in South Africa were assigned in a 1:1 ratio, by use of a computer-generated randomisation sequence, to the control or intervention clusters."
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with sequence generation
Blinding (performance and detection bias)	Low Risk	Quote: "The data gathering team were masked to the intervention allocation."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat."
Selective reporting (reporting bias)	Low Risk	Comment: The primary indicators listed in the trial registration were reported.
Recruitment bias	Low Risk	Quote: "In order to minimise selection bias in this cluster-randomised trial, ² all pregnant women in the randomised clusters were identified and approached. The likelihood of the study team missing a pregnant woman was very small. And in addition, only 52 of the women (2%) declined participation. This makes it unlikely that selective recruitment contributed importantly to the observed effect sizes."
Baseline imbalance	Low Risk	Quote: "Few baseline differences were noted between the intervention and control clusters at the country level"

Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Country-specific prevalence ratios were adjusted for cluster effects and sites"

UNICEF/UNDP/World Bank/WHO TDR 2000

Trial reg.	None
Country	Ghana and Kenya
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Full health districts in Ghana and half health districts in Kenya CLUSTERS: Intervention: 8, Control: 8
Population	LOCAL TERM FOR CHW: Community Drug Distributors (CDDs) RECRUITMENT CRITERIA: Ghana: The drug distributors were of varied backgrounds, including civil servants, teachers and unit committee members. Each of them had various reasons for why they thought they were selected to carry out the assignment. They included the fact that they were educated, trustworthy and dedicated, and had been involved in similar activities before. All the unit committee members mentioned that it was their duty to get involved in any activity that goes on in the community. Their special reasons for agreeing to take part in such a programme were that they wanted their community to be disease free, they all have the time and generally do a lot of things for the community. Kenya: The criteria for selection of CDDs had to be adhered to (ability to read and write; ability to keep records; trustworthy and willingness to distribute the drugs to all eligible persons in the allocated areas without expecting payment from the project). HEALTHCARE SETTING: Community PATIENT POPULATION: Males and females age 5-65+ years INTERVENTION SAMPLE SIZE: Ghana: 1436 Kenya: 1747 CONTROL SAMPLE SIZE: Ghana: 1589 Kenya: 1718
Intervention	OBJECTIVE: Determine the feasibility, effectiveness and potential sustainability of ComDT for lymphatic filariasis. INTERVENTION DESCRIPTION: A strategy of Community-Directed Treatment of filariasis, which incorporates the health services at the level of implementation (ComDT/HS). the community itself designs and implements a method for drug delivery that is most suitable to its needs. In the present study, the ComDT approach was introduced to the community by the Public Health System. KEY CHW COMPETENCY: Distribute drugs TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: Ghana: All the distributors selected in the ComDT/HS arm went for a 2-day training session in their community. All expenses for the training, transportation, feeding, etc., were borne by the communities. According to them, every community has a fund from which it takes money to support community programmes. The health worker in charge of a particular area together with the DHMT conducted the training. The topics covered were: aetiology of filariasis; treatment, its benefits and side-effects; management of side-effects; the treatment schedule using height; exclusion criteria for treatment; record keeping. Kenya: Similarly, the training of the CDDs took place at the district headquarters. Each participating village selected two CDDs (to help each other with measurements, drug administration and encourage and support each other during the long, sometimes dangerous walks. Training materials, content and language were similar to that of the community leaders. Training was conducted by the trained peripheral health staff. CDDs were taught to store the drugs away from fire, animals and children. To facilitate easy movement in the community, each pair of CDD was provided with a shoulder tote for carrying the pills, record books, a scooping spoon and the pens. As in the training of the Malindi community leaders, the issue of non-payment came up again during the training of the Malindi CDDs

	and was handled as before. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: During the training of the community leaders in Malindi, the issue of non-payment of wages to the CDDs was discussed at length as the leaders felt that the CDDs would deserve some kind of motivation for their services. Finally, it was agreed that ComDT was a community project and the community would be at liberty to decide the best way on how to remunerate their CDDs. RATIO CHW TO POPULATION: N.S. START DATE: 07/1997 END DATE: 10/ 1999 CONTROL DESCRIPTION: A delivery strategy of mass treatment by the regular health care system (HST)
Outcomes	OUTCOMES: Treatment coverage DURATION OF FOLLOW UP: Immediate
Funding	UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR)

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "In each country, four study units (full health districts in Ghana and half health districts in Kenya) were selected. Two were randomly allocated to HST and two to ComDT/HS."
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with randomisation
Blinding (performance and detection bias)	High Risk	Quote: "The survey was conducted by the research team."
Incomplete outcome data (attrition bias)	Unclear Risk	Quote: "The final evaluation included a household survey to assess the treatment coverage achieved in the different communities. A questionnaire was administered to all members of the 1,467 sample households who were present at the time of the survey."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified
Recruitment bias	Low Risk	Quote: "In each village 20 households were randomly selected. The sampling procedure involved 4 random starts with 5 households visited at intervals of 5 households each in random directions. All members of the sampled households were interviewed to determine whether they received and swallowed the drugs"
Baseline imbalance	Unclear Risk	Comment: Not specified
Loss of clusters	Unclear Risk	Comment: No cluster lost
Unit of analysis issues	Unclear Risk	Comment: Not specified

Vazir 2013

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT (3-arm) UNIT OF RANDOMISATION: Village CLUSTERS: Intervention A: 20, Intervention B: 20, Control: 20
Population	LOCAL TERM FOR CHW: trained village women (VW) RECRUITMENT CRITERIA: High school-educated VW who were themselves

	mothers. HEALTHCARE SETTING: Household PATIENT POPULATION: Mother-infants (3 months of age) dyads INTERVENTION SAMPLE SIZE: Intervention A: 210, Intervention B: 195 CONTROL SAMPLE SIZE: 202
Intervention	OBJECTIVE: This cluster-randomised trial tested the hypothesis that teaching caregivers appropriate complementary feeding and strategies for how to feed and play responsively through home-visits would increase children's dietary intake, growth and development compared with home-visit-complementary feeding education alone or routine care INTERVENTION DESCRIPTION: Intervention A: the complementary feeding group received the ICDS plus the World Health Organization recommendations on breastfeeding and complementary foods. Intervention B: the responsive complementary feeding and play group received the same intervention as the complementary feeding group plus skills for responsive feeding and psychosocial stimulation. Both intervention groups received bi-weekly visits by trained village women. KEY CHW COMPETENCY: The intervention teams (60 VW) were trained to have focused 'conversations' with mothers for the various intervention topics. The messages varied by child age. TRAINING DURATION (DAYS): TRAINING DESCRIPTION: They received supervised training on how to counsel mothers/caregivers using the pictorial flip charts. The intervention teams (60 VW) were trained to have focused 'conversations' with mothers for the various intervention topics. The messages varied by child age. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Trained graduates in nutrition supervised the VW, examined their records of visits and asked mothers independently what they were told in the VWs' last visit. They also held periodic reinforcement training sessions with VWs. The VWs visiting Intervention A and Intervention B homes were trained separately with specific flip charts. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: After 2001 END DATE: N.S. CONTROL DESCRIPTION: The control group received routine Integrated Child Development Services (ICDS). Mothers and infants in this group received only the routine ICDS1 services, which were operating across all study groups. These services consist mainly of centre-based supplemental food provided to 1–6-year-olds, pregnant and nursing mothers, home-visit counselling on breastfeeding and complementary feeding, monthly growth monitoring, and non-formal preschool education for children 3–5 years of age.
Outcomes	OUTCOMES: Primary: nutrient intake, growth and development measures. Secondary: morbidity, haemoglobin, and maternal knowledge, beliefs and behaviours. DURATION OF FOLLOW UP: 1 year
Funding	Indian Council of Medical Research, India and the NIH/NICHD (5 R01 HD042219-S1); additional funding from UNICEF, New York.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "The random allocation using a random number generator (facilitated through a tailor-made syntax programme in the Statistical Package for the Social

		Sciences (SPSS), which uses the select cases function) was undertaken by a researcher who was not familiar with the villages or their characteristics other than what could be derived from the 2001 census data"
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with sequence generation
Blinding (performance and detection bias)	Low Risk	Quote: "The assessment teams (psychologists and nutritionists) were blinded to the intervention and had no interaction with the VWs."
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Those who completed the intervention did not differ significantly from those lost to follow-up in these baseline measures"
Selective reporting (reporting bias)	Low Risk	Comment: Protocol not identified
Recruitment bias	Low Risk	Quote: "We explained the study objectives to all the pregnant women in the villages and asked if they would like to participate in the study. There were no refusals."
Baseline imbalance	Low Risk	Quote: "There were no significant differences among the three groups in any of the baseline characteristics "
Loss of clusters	Low Risk	Quote: "No clusters were lost"
Unit of analysis issues	Low Risk	Quote: "The intra-cluster correlation (ICC) of 0.05 to account for clustering"

Waiswa 2015

Trial reg.	ISRCTN50321130
Country	Uganda
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Village CLUSTERS: Intervention: 31; Control: 32
Population	LOCAL TERM FOR CHW: Village Health Team/Community health workers RECRUITMENT CRITERIA: CHWs from the intervention clusters were recruited by the community with the aim of identifying individuals with the following attributes: empathy; experience of similar problems and situations; respected in the local community; and considered to be a natural helper or someone that community members would naturally go to in the event of a problem. Women were preferred, although males were also accepted (women and men can serve as VHT members according to the national strategy). HEALTHCARE SETTING: Household PATIENT POPULATION: The trial included all consenting pregnant women and their newborns INTERVENTION SAMPLE SIZE: 894 CONTROL SAMPLE SIZE: 893
Intervention	OBJECTIVE: To assess the effect of a home visit strategy combined with health facility strengthening on uptake of newborn care seeking, practices and services, and to link the results to national policy and scale-up in Uganda. INTERVENTION DESCRIPTION: In intervention villages volunteer community health workers (CHWs) were trained to identify pregnant women and make five home visits (two during pregnancy and three in the first week after birth) to offer preventive and promotive care and counselling, with extra visits for sick and small newborns to assess and refer. Health facility strengthening was done in all facilities to improve quality of care. KEY CHW COMPETENCY: The CHW advises mothers on ideal family care practices like immediate breastfeeding, keeping baby warm by skin-to-skin care and proper wrapping, delayed bathing as well as appropriate cord care. The CHW further assesses and screens for danger signs all babies and mothers at each of the three postnatal visits and facilitates referral to health facility any baby or mother found with danger signs. TRAINING DURATION (DAYS): 5

	TRAINING DESCRIPTION: Training in essential maternal and newborn care. The training was conducted by a team of Ugandan trainers, who had attended the training of trainers course, with involvement of staff from the two districts. The training for CHWs was skills-based, and focused on promoting key selected practices as obtained from formative research and the recommendations of the workshop. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Quarterly SUPERVISION DESCRIPTION: After the initial training no additional off-site trainings were conducted, but knowledge and skills were reinforced during quarterly supervisory meetings and through directly observed supervision by health workers and health assistants. INCENTIVE: Non monetary INCENTIVE DESCRIPTION: CHWs' incentives were simple, and included a t-shirt, briefcase, certificate, and official commission following their training. The CHWs were not paid a salary by the study, but received a travel refund after supervision meetings. RATIO CHW TO POPULATION: 1:100-150 households START DATE: 03/2008 END DATE: 11/2011 CONTROL DESCRIPTION: In the comparison villages, women and their newborns had access to the standard health services, overseen by the district health team, in addition to the improved health facilities.
Outcomes	OUTCOMES: Improved coverage of services for antenatal care (One or more ANC visits, Four or more ANC visits, Prepared for birth, Knowledge of two or more pregnancy-related danger signs, One or more home visit during pregnancy, Delivered in health facility, Skilled attendant at delivery), birth preparedness, skilled attendance at delivery, and postnatal care as well as increases in healthy practices (Baby put to the breast within 1 hour of birth, Baby given colostrum, Baby exclusively breastfed in first month of life, Baby placed skin-to-skin with mother within 1 hour of birth, Baby dried immediately after birth, Baby wrapped immediately after birth, First bath delayed ≥ 6 hours after birth, First bath delayed ≥ 24 hours after birth, Hygienic care, Cord cut with clean instrument, Nothing applied to umbilical cord after cutting), DURATION OF FOLLOW UP: 3 years
Funding	Sida/SAREC-Makerere University-Karolinska Institutet Research collaboration as well as by funds provided by Save the Children through a grant from the Bill & Melinda Gates Foundation

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Computer-generated restricted randomisation was done in a one-to-one ratio by an independent epidemiologist from the London School of Hygiene and Tropical Medicine. "
Allocation concealment (selection bias)	Low Risk	Comment: Done simultaneously with sequence generation
Blinding (performance and detection bias)	Unclear Risk	Quote: "Data for the endline survey were collected "
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analyses were by intention to treat."
Selective reporting (reporting bias)	Unclear Risk	Comment: 19 outcomes listed in the protocol
Recruitment bias	Low Risk	Comment: Trial included all consenting pregnant women in study areas
Baseline imbalance	Low Risk	Quote: "Women's age, parity, and household wealth were similar at baseline across the intervention and control arms"
Loss of clusters	Low Risk	Comment: No clusters lost

Unit of analysis issues	Low Risk	Quote: "Primary sampling units and strata for the data set were defined to account for the cluster-randomised design"
-------------------------	----------	---

Wang 2007

Trial reg.	None
Country	China
Methods	STUDY DESIGN: C-RCT (3-arm) UNIT OF RANDOMISATION: Townships CLUSTERS: Intervention A: 40, Intervention B: 40, Control: 40
Population	LOCAL TERM FOR CHW: Village-based health workers RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Children under 60 months INTERVENTION SAMPLE SIZE: Intervention A: 391, Intervention B: 410 CONTROL SAMPLE SIZE: 401
Intervention	OBJECTIVE: Hepatitis B vaccination of newborn infants in rural China: evaluation of a village-based, out-of-cold-chain delivery strategy INTERVENTION DESCRIPTION: Intervention A: Vaccine was stored out of the cold chain in villages and administered by village-based health workers to infants at home. Intervention B: Same as Intervention A but vaccine was packaged in a prefilled injection device. KEY CHW COMPETENCY: Administer Hep B vaccine TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 09/2003 END DATE: 09/2004 CONTROL DESCRIPTION: Vaccine was stored within the cold chain and administered in township hospitals.
Outcomes	OUTCOMES: Improvements in timely delivery of the hepatitis B vaccine at birth, Proportion of children who received the first dose of hepatitis B vaccine within 24 hours of birth DURATION OF FOLLOW UP: 1 day
Funding	Children's Vaccine Program of the Program for Appropriate Technology in Health, with funding from the Bill & Melinda Gates Foundation.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Rural townships (those lying more than 20 km from the county capital) in the participating counties were randomly divided into three groups:" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear Risk	Comment: Not mentioned
Blinding (performance and detection bias)	Unclear Risk	Quote: "Two surveys were conducted, one before study implementation (the "baseline coverage survey") and one after (the "follow-up coverage survey")" Comment: Unclear if assessors were blinded or if they were different from those delivering the intervention.
Incomplete outcome	Unclear Risk	Comment: Attrition unclear

data (attrition bias)		
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified.
Recruitment bias	Low Risk	Quote: "At each selected village, the survey team started at the village centre and continued in a randomly chosen direction until it had identified and interviewed the parents of at least 10 children aged 9–20 months for the baseline survey and aged 1–11 months for the follow-up survey"
Baseline imbalance	Low Risk	Comment: Groups similar at baseline
Loss of clusters	Low Risk	Comment: No clusters lost
Unit of analysis issues	Low Risk	Quote: "Which take into account the clustered sampling used in these surveys."

Wright 2004

Trial reg.	None
Country	Swaziland
Methods	STUDY DESIGN: Stratified RCT UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Community health workers (rural health motivators) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: Household PATIENT POPULATION: Patients were eligible if they were adults or children diagnosed with smear positive or smear negative pulmonary tuberculosis; extra pulmonary tuberculosis or relapse of previously treated tuberculosis. INTERVENTION SAMPLE SIZE: 668 CONTROL SAMPLE SIZE: 667
Intervention	OBJECTIVE: Implement community-based DOTS, including a randomised controlled trial of direct observation either by CHWs vs. family members, under operational conditions in a region of Swaziland. INTERVENTION DESCRIPTION: Community health worker supervision. Community health workers acted as treatment supporters with the patient visiting every day for direct observation of treatment. The health worker was trained in observing the patient take his/her medication on a daily basis, reminding them if they forgot and recording adherence on a Treatment Supporter Card. Defaulters were notified to the diagnostic centre. KEY CHW COMPETENCY: The health worker was trained in observing the patient take his/her medication on a daily basis, reminding them if they forgot and recording adherence on a Treatment Supporter Card TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: Clinic-based training sessions were run to explain the study and how to support and observe treatment. The health worker was trained in observing the patient take his/her medication on a daily basis, reminding them if they forgot and recording adherence on a Treatment Supporter Card. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 04/2000 END DATE: 06/2002 CONTROL DESCRIPTION: Family member/carer supervision. A family member or carer was nominated by the patient to become the treatment supporter. This person was then trained in observing the patient take his/her medication on a daily basis, reminding them if they forgot, and recording adherence on a Treatment

	Supporter Card.
Outcomes	OUTCOMES: Cured or completed treatment DURATION OF FOLLOW UP: 6 months
Funding	The Department for International Development

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear Risk	Quote: "Adults were randomised into four groups according to diagnosis: smear positive pulmonary tuberculosis; smear negative tuberculosis; extra pulmonary and relapse/retreatment. A stratified randomisation process was used with sealed, sequentially numbered, stratum specific envelopes containing treatment assignments. " Comment: Not random.
Allocation concealment (selection bias)	Unclear Risk	Quote: "After the patient was diagnosed, an independent administrator handed the clinician the envelope." Comments: Envelopes were sequentially numbered
Blinding (performance and detection bias)	Low Risk	Quote: "Outcome assessment was undertaken by laboratory examination of sputum by technicians unaware of treatment allocation. "
Incomplete outcome data (attrition bias)	Low Risk	Comment: 1% LTFU
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified

Yeomans 2010

Trial reg.	None
Country	Burundi
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Facilitator RECRUITMENT CRITERIA: Chosen by the nonprofit organisation for their extensive experience with trauma workshop facilitation and for having demographics comparable to participants. CHWs were rural, poor, many without substantial formal education, and balanced in gender and ethnicity HEALTHCARE SETTING: Community PATIENT POPULATION: Adults among future participants of trauma workshops offered by internally displaced people camps. INTERVENTION SAMPLE SIZE: Intervention A: 41, Intervention B: 41 CONTROL SAMPLE SIZE: 42
Intervention	OBJECTIVE: Evaluate the effects of PTSD psychoeducation within a larger trauma healing and reconciliation intervention in a rural region of Burundi INTERVENTION DESCRIPTION: (A) intervention with PTSD psychoeducation: the standard intervention included two phases. Six groups of approximately 20 participants gathered for 3 days, and 1 month later each workshop group reconvened for a full-day follow-up session during which major workshop components were reinforced. The 3-day workshop used discussion, experiential exercises aimed at fostering interpersonal exchange, and games to explore themes of trauma, loss, anger, trust, and the roots of violence. Psychoeducational content on the first day of the workshop included a 90-minute presentation and discussion of the 17 specific symptoms of PTSD. (B) Intervention without PTSD education: identical to Intervention A with two exceptions. First, this condition did not include the introduction of PTSD psychoeducational content. Second, to ensure that both workshop conditions were of equal length, additional time was devoted to an exercise in which participants formed pairs and answered questions provided to

	them. KEY CHW COMPETENCY: Facilitate Healing and Reconciling Our Communities workshop and PTSD education TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Training on the modification of the standard workshop to accommodate planned differences in condition REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: Spring/07 END DATE: N.S. CONTROL DESCRIPTION: Waitlist control participants were enrolled in the standard workshop. Their participation was postponed until after the termination of the initial workshop cycle
Outcomes	OUTCOMES: Hopkins Symptom Checklist -25 combined with 10 somatic symptoms from Hopkins Symptom Checklist -58 comprised a hybrid Hopkins Symptom Checklist instrument (anxiety and depression, global measure of emotional distress); HTQ Part IV (Trauma); Harvard Trauma Questionnaire-b (guilt, loneliness, shame, betrayal and rumination). DURATION OF FOLLOW-UP: 2 weeks post intervention
Funding	N.S.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "using a computerized random-number generator, participants were assigned to condition according to stratified randomisation"
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "Participants and interviewers (at pre and post-test) were blind to condition assignment"
Incomplete outcome data (attrition bias)	Low Risk	Comment: Attrition <10%
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified

Yousafzai 2014 (incl. Yousafzai 2015)

Trial reg.	NCT007159636
Country	Pakistan
Methods	STUDY DESIGN: C-RCT (2x2 factorial) UNIT OF RANDOMISATION: LHW catchments CLUSTERS: Intervention A: 20, Intervention B: 20, Intervention C: 20, Control: 20
Population	LOCAL TERM FOR CHW: Lady health workers (LHWs) RECRUITMENT CRITERIA: 18–45 years of age, live locally, preferably married, at least 8 years of education, community acceptance. HEALTHCARE SETTING: Household PATIENT POPULATION: All children born in the study area were eligible for enrolment if they were up to 2.5 months old without signs of severe impairments. INTERVENTION SAMPLE SIZE: Intervention A: 364, Intervention B: 383, Intervention C: 374 CONTROL SAMPLE SIZE: 368
Intervention	OBJECTIVE: Investigate the feasibility and effectiveness of the integration of interventions to enhance child development and growth outcomes in the LHW programme in Sindh, Pakistan.

	INTERVENTION DESCRIPTION: Intervention A: Nutrition education and multiple micronutrient powders, Intervention B: responsive stimulation, or Intervention C: a combination of both enriched interventions. KEY CHW COMPETENCY: TRAINING DURATION (DAYS): 90 TRAINING DESCRIPTION: 3 months classroom-based training and 12 months field-based training REFRESHER OR ONGOING TRAINING: 15 days allocated to refresher training each year. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Supervision provided by Lady Health Supervisor. Supportive supervision is promoted, including use of checklists and problem solving. INCENTIVE: Monetary INCENTIVE DESCRIPTION: The average cost of each LHW is 44 000 Pakistani rupees per year (US\$570). The main cost is her monthly stipend RATIO CHW TO POPULATION: 1:1000 START DATE: 01/06/2009 END DATE: 31/03/2012 CONTROL DESCRIPTION: Routine health and nutrition services
Outcomes	OUTCOMES: Primary: development at 12 and 24 months of age (assessed with the Bayley Scales of Infant and Toddler Development, Third Edition) and growth at 24 months of age. Secondary: Child morbidity, Mother-child interaction, Caregiving environment, Maternal knowledge and practices, Maternal depressive symptoms DURATION OF FOLLOW UP: 24 months
Funding	UNICEF

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "To minimise the risk of contamination of interventions across groups, we used a two-stage stratified random sampling strategy...straightforward random selection of these clusters was undertaken independently from the study team until the necessary sample size was reached. Random assignment of the intervention group was done independently from the study team. "
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding (performance and detection bias)	Low Risk	Quote: "The data collection team were masked to the allocated intervention"
Incomplete outcome data (attrition bias)	Low Risk	Quote: "Analysis was by intention to treat. "
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol identified
Recruitment bias	Low Risk	Quote: "Every infant born in the study area between April 1, 2009, and March 31, 2010, was eligible for enrolment. "
Baseline imbalance	Low Risk	Quote: "The mean household socioeconomic status scores, the nutritional status of infants at the time of enrolment...did not differ significantly between the groups. A similar proportion of girls and boys were enrolled in each group...However, significant differences between intervention groups were noted in maternal education. Additionally, infants in the responsive stimulation and enhanced nutrition groups were significantly older at enrolment than were those in the comparative groups"

Loss of clusters	High Risk	Quote: "During the trial, 16 LHWs stopped working for a prolonged period of time. Four LHWs left their posts officially, all of whom were in the combined responsive stimulation and enhanced nutrition group."
Unit of analysis issues	Low Risk	Quote: "All analyses reported are adjusted for clustering effects, accounting for both levels of sampling, using generalised estimating equations."

Zwarenstein 2000

Trial reg.	None
Country	South Africa
Methods	STUDY DESIGN: RCT (3-arm) UNIT OF RANDOMISATION: Patient
Population	LOCAL TERM FOR CHW: Lay health worker (LHW) RECRUITMENT CRITERIA: N.S. HEALTHCARE SETTING: LHW Household PATIENT POPULATION: The study population consisted of adult (aged > 15 years) pulmonary TB patients, who started a course of TB treatment (new and retreatment). INTERVENTION SAMPLE SIZE: 54 CONTROL SAMPLE SIZE: Control A: 58, Control B: 44
Intervention	OBJECTIVE: Comparison of successful tuberculosis treatment outcome rates between self-supervision, supervision by LHW, and supervision by clinic nurse. INTERVENTION DESCRIPTION: Patients on LHW supervision took their drugs several times per week at their LHW's home and under the LHW's direct supervision. The drugs, held by the LHW, were replenished each month by the project. Patients visited the clinic only for tests and in the case of complications. If a patient missed a day's treatment the LHW visited the patient's home. KEY CHW COMPETENCY: Hold drugs, directly observe patient taking drugs, visit patient home if dose missed TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 05/1994 END DATE: 09/1995 CONTROL DESCRIPTION: Control A: DOT patients were expected to attend the clinic during working hours to take their medication under observation by clinic nurses. Control B: Self-supervised patients visited the clinic once a week, or sent a family member to collect their drugs. As responsibility for adherence lay with the patient, this was termed 'self supervision.'
Outcomes	OUTCOMES: Successful treatment (includes patients who were cured and those who completed treatment) DURATION OF FOLLOW UP: 6 months
Funding	Medical Research Council, the Department of Health and the Health Systems Trust.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low Risk	Quote: "Researchers (CV and HS) regularly checked that trial numbers were allocated in sequence, that changes in option were explained, and that questionnaires were complete."

Allocation concealment (selection bias)	Low Risk	Quote: "the nurse drew an envelope from the top of a box of consecutively numbered opaque sealed envelopes, recording the allocation in the trial book for that clinic. Researchers (CV and HS) regularly checked that trial numbers were allocated in sequence, that changes in option were explained, and that questionnaires were complete. "
Blinding (performance and detection bias)	Low Risk	Comment: Participants not blinded, personnel not blinded. Not clear if outcome assessors blinded but outcome assessment was not likely to be influenced by the lack of blinding.
Incomplete outcome data (attrition bias)	Low Risk	Quote: "intention-to-treat analysis."
Selective reporting (reporting bias)	Unclear Risk	Comment: No protocol specified

AL: artemether-lumefantrine; **ARI:** acute respiratory infection, **AQ:** amodiaquine; **BNCP:** Birth and newborn care preparedness; **CCM:** community case-management; **CDR:** Clinical Dementia Rating scale; **CHF:** chronic heart failure; **C-RCT:** cluster randomised controlled trial; **CVD:** cardiovascular disease; **DFID:** UK Department for International Development; **DOTS:** directly observed therapy short-course; **DSM:** Diagnostic and Statistical Manual of Mental Disorders; **HB-VCT:** home-based voluntary counselling and testing; **HIV/AIDS:** human immunodeficiency virus/ acquired immune deficiency syndrome; **LBW:** low birth weight; **ICD-10-DCR:** tenth edition of the International Classification of Diseases, Diagnostic Criteria for Research; **IMCI:** Integrated Management of Childhood Illness; **IPTc:** intermittent preventive treatment in children; **IPT-G:** group-based interpersonal psychotherapy; **NBW:** normal birth weight; **NMR:** neonatal mortality; **N.S.:** not stated; **NSP:** New smear-positive; **PMTCT:** Prevention of mother-to-child transmission of HIV; **PPH:** postpartum haemorrhage; **PTSD:** post-traumatic stress disorder; **RCT:** randomised controlled trial; **RDT:** Malaria rapid diagnostic test; **SP:** sulfadoxine/pyrimethamine; **TB:** tuberculosis; **USAID:** United States Agency for International Development; **WAZ:** weight for age Z-score; **WHO TDR:** WHO Special Programme for Research and Training in Tropical Diseases; **WRA:** women of reproductive age;

Characteristics of excluded studies

Record name	Year	Reason for exclusion
Aboud	2011	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Aboud	2008	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
AFRINEST	2015A	CHW contribution can't be disaggregated from larger intervention
AFRINEST	2015b	CHW contribution can't be disaggregated from larger intervention
Akogun	2001	Not a randomised controlled trial
Aksu	2011	Intervention not delivered by CHWs (hospital midwives)
Ali	2005	Not a randomised trial
Allen	2014	Not in a LMIC (USA)
Allen	2011	Not in a LMIC (USA)
Andreoni	2016	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)

Araya	2003	Health workers not primarily based in the community
Araya	2006	Intervention not delivered by CHWs
Arifeen	2009	CHW contribution can't be disaggregated from larger intervention
Arifeen	2012	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Armstrong Schellenberg	2010	Intervention not delivered by CHWs
Arrossi	2015	Not in a LMIC (Argentina)
Ashraf	2015	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)
Banerjee	2015	CHW contribution can't be disaggregated from larger intervention
Bang	1999	Not a randomised trial
Baqui	2002	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Baqui	2015	Intervention not delivered by CHWs
Baqui	2008	Not a randomised trial
Bhandari	2003	CHW contribution can't be disaggregated from larger intervention
Bhandari	2005	CHW contribution can't be disaggregated from larger intervention
Bhandari	2004	CHW contribution can't be disaggregated from larger intervention
Bhandari	2012	CHW contribution can't be disaggregated from larger intervention
Bhutta	2008	CHW contribution can't be disaggregated from larger intervention
Bhutta	2011	CHW contribution can't be disaggregated from larger intervention
Bhutta	2009	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Björkman	2009	Intervention not delivered by CHWs
Björkman	2007	Intervention not delivered by CHWs
Bolam	1998	CHW contribution can't be disaggregated from larger intervention
Bowman	2000	Intervention not delivered by CHWs
Brooker	2015	Intervention not delivered by CHWs
Brugha	1996	CHW contribution can't be disaggregated from larger intervention
Buttorff	2012	Health workers not primarily based in the community
Capuccio	2006	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Castro	2012	Intervention not delivered by CHWs

Cavalcante	2010	Health workers not primarily based in the community
Chang	2015	Health workers not primarily based in the community
Chang	2011	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)
Chaudhuri	2008	Not a randomised trial
Chwo	2002	Intervention not delivered by CHWs
Cooke	1983	Not in a LMIC (USA)
Corbett	2010	Intervention not delivered by CHWs
Das	2015	CHW contribution can't be disaggregated from larger intervention
Datta	1987	Not a randomised trial
DeRenzi	2013	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)
Derman	2006	Intervention not delivered by CHWs
Dicko	2011	Intervention not delivered by CHWs
Dinshaw	2007	Intervention not delivered by CHWs
Dorner	2013	Not in a LMIC (Austria)
Dybdahl	2001	Intervention not delivered by CHWs
El Arifeen	2004	CHW contribution can't be disaggregated from larger intervention
Eriksen	2010	CHW contribution can't be disaggregated from larger intervention
Feikin	2014	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Florez-Arango	2011	Intervention not delivered by CHWs
George	2012	Intervention not delivered by CHWs
Gill	2014	Intervention not delivered by CHWs
Gill	2011	Intervention not delivered by CHWs
Hamadani	2006	Health workers not primarily based in the community
Han	1989	Intervention not delivered by CHWs
Hargraves	2012	Not in a LMIC (USA)
Harrington	2000	Not in a LMIC (UK)
Hirani	2010	Health workers not primarily based in the community
Hopper	1984	Not in a LMIC (USA)
Hunchangsith	2012	CHW contribution can't be disaggregated from larger intervention
Iyengar	2013	Intervention not delivered by CHWs
Jahanfar	2005	CHW contribution can't be disaggregated from larger intervention
Jokhio	2005	Intervention not delivered by CHWs
Jordans	2013	Not a randomised trial
Joshi	2012	Intervention not delivered by CHWs
Kamolratanakul	1999	CHW contribution can't be disaggregated from larger intervention
Katula	2011	Not in a LMIC (USA)

Khan	2002	CHW contribution can't be disaggregated from larger intervention
Kikweewa	2013	CHW contribution can't be disaggregated from larger intervention
Kilaru	2005	CHW contribution can't be disaggregated from larger intervention
Kirkwood	2010	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Kittayapong	2012	CHW contribution can't be disaggregated from larger intervention
Korfmacher	1999	Not in a LMIC (USA)
Kouyaté	2008	CHW contribution can't be disaggregated from larger intervention
Kramer	2001	Intervention not delivered by CHWs (hospital midwives)
Labhardt	2011	Health workers not primarily based in the community
Le	2005	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Lechtig	1982	CHW contribution can't be disaggregated from larger intervention
Lee	2015	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Li	1989	CHW contribution can't be disaggregated from larger intervention
Littlejohns	1991	Not in a LMIC (UK)
Lozoff	2010	Intervention not delivered by CHWs
Luby	2006	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Luby	2004	Intervention not delivered by CHWs
Lutalo	2010	CHW contribution can't be disaggregated from larger intervention
MacPherson	2014	Intervention not delivered by CHWs
Mazumder	2014	CHW contribution can't be disaggregated from larger intervention
Menendez	1994	Intervention not delivered by CHWs
Mermin	2012	Not a randomised trial
Mermin	2012	Not in a LMIC (Taiwan)
Midhet	2011	Intervention not delivered by CHWs
Mirmolaei	2014	Intervention not delivered by CHWs
Mirzoev	2008	Not a randomised trial
	2014	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Morris		
Morris	2012	Not a randomised trial
Mosha	2005	CHW contribution can't be disaggregated from larger intervention

Mubi	2011	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Muhamadi	2011	CHW contribution can't be disaggregated from larger intervention
Mukanga	2012	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Mullany	2006	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Nahar	2012	Intervention not delivered by CHWs
Nikiema	2014	Intervention not delivered by CHWs
Nonvignon	2012	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Nyamathi	2013	CHW contribution can't be disaggregated from larger intervention
Nyamathi	2012	CHW contribution can't be disaggregated from larger intervention
Omer	2008	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)
Osrin	2003	Intervention not delivered by CHWs
Owais	2011	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Pandey	1991	Not a randomised trial
Patel	2010	CHW contribution can't be disaggregated from larger intervention
Patel	2011	CHW contribution can't be disaggregated from larger intervention
Pearson	2007	Not a randomised trial
Peltzer	2012	CHW contribution can't be disaggregated from larger intervention
Persson	2013	CHW contribution can't be disaggregated from larger intervention
Phillips	2006	Intervention not delivered by CHWs
Pisani	2006	Intervention not delivered by CHWs
Prinja	2013	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Quick	1999	Trial testing the efficacy or effectiveness or different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Reifsnider	2013	Not in a LMIC (USA)
Roberts	2011	Not in a LMIC (Australia)
Rodriguez-Garcia	1990	CHW contribution can't be disaggregated from larger intervention

Ryman	2012	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Samuels	2014	CHW contribution can't be disaggregated from larger intervention
Sazawal	1996	Health workers not primarily based in the community (Author email)
Shankar	2009	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Shankar	1999	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Shargie	2006	CHW contribution can't be disaggregated from larger intervention
Shewade	2013	Intervention not delivered by CHWs
Sikander	2013	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Sikander	2015	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Som	2014	Not a randomised trial
Soofi	2013	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Staedke	2009	Intervention not delivered by CHWs
Stinson	2014	Health workers not primarily based in the community
Sutcliffe	2009	Intervention not delivered by CHWs
Sweat	2011	Intervention not delivered by CHWs
Syed	2006	Not a randomised trial
Tagbor	2011	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Talukder	2012	Health workers not primarily based in the community
Tang	2014	Not in a LMIC (USA)
Terefe	1993	Intervention not delivered by CHWs
Tiwari	2010	Intervention not delivered by CHWs
Tofail	2013	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Townsend	1983	CHW contribution can't be disaggregated from larger intervention
Tsai	2014	Health workers not primarily based in the community
Turab	2014	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Uwimana	2012	CHW contribution can't be disaggregated from larger intervention

Uwimana	2013	CHW contribution can't be disaggregated from larger intervention
Vasan	1997	Intervention not delivered by CHWs
Waitzkin	2011	Health workers not primarily based in the community
Walley	2001	CHW contribution can't be disaggregated from larger intervention
Wamae	2006	Not a randomised trial
Wandwalo	2004	CHW contribution can't be disaggregated from larger intervention
Webster	2016	CHW contribution can't be disaggregated from larger intervention
Winch	2003	Delivery of same health intervention in both arms with changes made to type of CHW support (i.e. testing CHW performance)
Wu	2011	Intervention not delivered by CHWs
Xavier	2016	Health workers not primarily based in the community
Yeboah-Antwi	2010	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Yeomans	2010	Intervention not delivered by CHWs (author email)
Zaidi	2013	Intervention not delivered by CHWs
Zhao	2013	Trial testing the efficacy or effectiveness of different interventions, not the efficacy or effectiveness of CHW delivery versus an alternative
Zwarenstein	2011	Health workers not primarily based in the community

Characteristics of ongoing studies

Abdul Latif Jameel Poverty Action Lab 2015

Trial Registration	NCT01487707
Trial name	Community Based Strategies to Reduce Maternal Mortality in Northern Nigeria
Objective	Assess the impact of several community-based interventions that address the key factors underlying the high maternal mortality, as well as neonatal mortality and morbidity in northern Nigeria.
Methods	C-RCT
Unit of Randomisation	Village
Country	Nigeria
Participants	Females 15 to 49 years
Interventions	(A) Voluntary Health Workers will conduct door-to-door campaigns on maternal and child health, and will visit pregnant women regularly to educate and to encourage use of antenatal care and facility-based deliveries. (B) Voluntary Health Workers will conduct door-to-door

	campaigns on maternal and child health, and will visit pregnant women regularly to to educate and to encourage use of antenatal care and facility-based deliveries. A safe birth kit, containing sterile materials to diminish infection risk will be given to pregnant women. (C) Voluntary Health Workers will conduct door-to-door campaigns on maternal and child health, and will visit pregnant women regularly to to educate and to encourage use of antenatal care and facility-based deliveries. Community-wide media activities, including dramas, intended to address community norms regarding maternal and child health will be conducted.
Control	N.S.
Outcomes	Primary: Maternal Mortality Ratio
Starting Date (dd/mm/yyyy)	12/2011
Contact Information	Jessica Leight Department of Economics Williams College, USA jl14@williams.edu
Funding Source	Abdul Latif Jameel Poverty Action Lab, Planned Parenthood Federation Nigeria, and MacArthur Foundation

Bjorkman-Nykqvist 2016

Trial Registration	PACTR201308000601715
Trial name	Evaluating the impact of the Living Goods entrepreneurial model of community health delivery in Uganda: A cluster-randomized controlled trial
Objective	Assess the impact of a non-profit entrepreneurial model of community health delivery
Methods	C-RCT
Unit of Randomisation	Village
Country	Uganda
Participants	Children under 5 years
Interventions	Living Goods and BRAC Community Health Promoters conducted home visits, educated households on essential health behaviours and sold preventive and curative health products at 20-30% below prevailing retail prices
Control	N.S.
Outcomes	Under-five child mortality rate
Starting Date	2013

(dd/mm/yyyy)	
Contact Information	Jakob Svensson IIES, Stockholm University Jakob.Svensson@iies.su.se
Funding Source	Children Investment Fund Foundation; Swedish Research Council (421-2009-2209); the Program for Development Research, Swedish International Development Cooperation; William F. Milton Fund at Harvard Medical School; Harvard Center for Population and Development Studies Grant.

Gong 2014

Trial Registration	ChiCTR-TRC-13003263
Trial name	Village doctor-assisted case management of rural patients with schizophrenia: protocol for a cluster randomised control trial.
Objective	Train village ‘doctors’ (i.e., villagers with basic medical training and currently operating in villages across China delivering basic clinical and preventive care) to manage rural patients with schizophrenia with respect to compliance and monitoring symptoms.
Methods	C-RCT
Unit of Randomisation	Village
Country	China
Participants	Diagnosed by psychiatrist as having schizophrenia
Interventions	Village doctors who have received training in a schizophrenia case management protocol will manage case records, supervise drug taking, educate patients and families on schizophrenia and its treatment, and monitor patients for signs of relapse in order to arrange prompt referral. (20 villages)
Control	Hospital visit as per the centres’ standard of care
Outcomes	Drug compliance; symptoms score; social functions; quality of life score; hospitalization rate; incidence of risky behaviour
Starting Date (dd/mm/yyyy)	01/07/2013
Contact Information	Wenjie Gong School of Public Health Central South University 110 Xiangya Road, Changsha 410078 Hunan, China xiaosy@csu.edu.cn
Funding	2012 China Medical Board Open Competition Program

Source	
Innovations for Poverty Action 2016	
Trial Registration	None (See: http://www.poverty-action.org/study/evaluation-community-based-environmental-health-promotion-program-rwanda)
Trial name	Evaluation of the Community-Based Environmental Health Promotion Program in Rwanda
Objective	To evaluate the effects of Rwanda's Community Based Environmental Health Promotion Program
Methods	C-RCT
Unit of Randomisation	Village
Country	Rwanda
Participants	Households
Interventions	(A) "Classic" programme group: Communities in this study arm will receive the complete CBEHPP program. Highly-trained facilitators – one selected in each village – will lead households through 20 weekly community hygiene club sessions, use high-quality instruction materials, and issue membership cards to participants. (50 villages) (B) "Lite" programme group: Communities in this study arm will receive a two-month programme that contains eight weekly sessions covering all the WASH topics. Facilitators selected from each village will receive minimal training and will deliver the topics with black/white photocopies of instructional materials. (50 villages)
Control	Communities in this study arm will not receive any programme during the study period. They will receive the Classic programme at the conclusion of the study. (50 villages)
Outcomes	Diarrhea prevalence, especially among children under age 5; child anthropometrics; household drinking water quality; self-reported disease burden; socio-economic conditions; and cooperation within communities, measured through field-based behavioral games.
Starting Date (dd/mm/yyyy)	N.S.
Contact Information	Sophie Boisson London School of Hygiene and Tropical Medicine sophie.boisson@lshtm.ac.uk
Funding	N.S.

Source

Kamath 2014

Trial Registration	REF/2013/03/004737
Trial name	Rationale and design of a randomised controlled trial evaluating community health worker-based interventions for the secondary prevention of acute coronary syndromes in India (SPREAD)
Objective	Open randomised trial of CHW-based interventions versus standard care for prevention of acute coronary syndrome (ACS).
Methods	RCT
Unit of Randomisation	Patient
Country	India
Participants	Patients admitted with a clinical diagnosis of ACS (standard definitions used) aged > 18 years
Interventions	At hospital visits (1, 5, 7, and 12 months), the CHW assessed risk factors, reviewed key investigations, set targets for lifestyle modifications, and collected medication adherence data. At home visits (3 months, 9 months), CHWs collected basic clinical data and reviewed risk factors and medication adherence. Between visits, patients were encouraged to call the CHW for any health-related matters.
Control	Hospital visit as per the centres' standard of care
Outcomes	The primary outcome measure was medication adherence. The secondary outcomes were differences in adherence to lifestyle modification, physiological parameters (blood pressure [BP], body weight, body mass index [BMI], heart rate, lipids), and major adverse cardiovascular events
Starting Date (dd/mm/yyyy)	07/01/2011
Contact Information	Dr. Deepak Y Kamath Division of Clinical Research & Training St. John's Research Institute Bengaluru, India
Funding Source	United States NHLBI, National Institutes of Health, Department of Health and Human Services, under contract no. HHSN268200900025C. This project has also been funded in part by the UnitedHealth Group

Khan 2012

Trial Registration	ISRCTN86264432
Trial name	Making birthing safe for Pakistan women: a cluster randomised trial
Objective	Test whether structured planning for safe birthing and effective

	Emergency Obstetric & Newborn Care (EmONC) services and/or travel facilitation in Pakistan, facilitated mainly through female health workers will reduce the neonatal mortality rate
Methods	3 Arm C-RCT
Unit of Randomisation	N.S.
Country	Pakistan
Participants	Pregnant women
Interventions	Arm 1: structured planning for safe birthing/EmONC Arm 2: structured planning for safe birthing/EmONC plus transport facilitation
Control	Arm 3: no transport facilitation or structured birth planning-education for safe birthing/EmONC (current/routine practice)
Outcomes	Primary: Neonatal mortality rate (NMR) Secondary: Neonatal morbidity rates
Starting Date (dd/mm/yyyy)	1/2/11
Contact Information	Dr Muhammad Amir Khan 12 Street 48 F-7/4 Islamabad 44000 Pakistan
Funding Source	Research & Advocacy Fund (RAF) Ref: RAF/LG/DTL/18-6/20

Kimani-Murage 2013

Trial Registration	ISRCTN83692672
Trial name	Effectiveness of personalised, home-based nutritional counselling on infant feeding practices, morbidity and nutritional outcomes among infants in Nairobi slums
Objective	Test the effectiveness of a home-based intervention on infant feeding practices, nutrition and health outcomes of infants born in two slums
Methods	C-RCT
Unit of Randomisation	Community units (defined by the Government's health care system)
Country	Kenya
Participants	Mother-child dyads
Interventions	The mothers will receive regular, personalised, home-based counselling by trained Community Health Workers on maternal, infant and young child nutrition
Control	Home-based counselling by CHWs on usual care (for example, ante-natal care, delivery with skilled attendants, immunization)
Outcomes	Primary: exclusive breastfeeding. Secondary: other infant feeding

	practices, diarrhoea morbidity and nutritional outcomes)
Starting Date (dd/mm/yyyy)	1/3/12
Contact Information	Elizabeth W Kimani-Murage African Population and Health Research Center (APHRC), P.O. 10787, 00100 Nairobi, Kenya ekimani@aphrc.org
Funding Source	Wellcome Trust

Kramer 2014

Trial Registration	Not stated
Trial name	A randomised longitudinal factorial design to assess malaria vector control and disease management interventions in rural Tanzania
Objective	This project will evaluate two malaria control interventions in Mvomero District, Tanzania: (1) a disease management strategy involving early detection and treatment by community health workers using rapid diagnostic technology and (2) vector control through community-supported larviciding. Six study villages were assigned to each of four groups (control, early detection and treatment, larviciding, and early detection and treatment plus larviciding).
Methods	C-RCT
Unit of Randomisation	Village
Country	Tanzania
Participants	Any member of an enrolled study household
Interventions	(A) A disease management strategy involving early detection and treatment by community health workers using rapid diagnostic technology; (B) Early detection and treatment plus vector control through community-supported larviciding
Control	(A) Larviciding alone; (B) No intervention
Outcomes	Change in malaria infection prevalence across the intervention groups measured during annual longitudinal cross-sectional survey
Starting Date (dd/mm/yyyy)	01/03/2013
Contact Information	Marie Lynn Miranda Departments of Pediatrics and Obstetrics & Gynecology, School of Medicine University of Michigan, Ann Arbor, MI 48109, USA mlmirand@umich.edu
Funding Source	National Institute of Allergy and Infectious Diseases

Lund 2014

Trial Registration	NCT01977326
Trial name	Task sharing of a psychological intervention for maternal depression in Khayelitsha, South Africa: study protocol for a randomised controlled trial
Objective	The objective of this RCT is to determine the effectiveness and cost-effectiveness of a task sharing counselling intervention for maternal depression in South Africa.
Methods	RCT
Unit of Randomisation	Patients
Country	South Africa
Participants	Women, aged >18 years, screening positive for depression with a cut off of 13 or more on the EPDS
Interventions	The intervention group will be given six sessions of basic counselling over a period of 3 to 4 months, provided by trained community health workers (CHWs).
Control	The control group will receive three monthly phone calls from a CHW trained to conduct phone calls but not basic counselling.
Outcomes	Primary: 17-Item Hamilton Depression Rating Scale (HDRS-17) change scores. Secondary: Edinburgh Postnatal depression scale (EPDS), Mini International Psychiatric Interview 2.0.0, Multidimensional Scale of Perceived Social Support (MSPSS), WHO Disability Assessment Schedule (WHO-DAS), Health services utilisation questionnaire, Cape Town Functional Assessment Instrument for Maternal Depression, and Obstetric and Infant outcome measures.
Starting Date (dd/mm/yyyy)	01/10/2013
Contact Information	Professor Crick Lund, Associate Professor, University of Cape Town South Africa crick.lund@uct.ac.za +27216850120
Funding Source	National Institute of Mental Health (NIMH) of the National Institutes of Health (NIH) under Award Number U19MH095699.

Morrison 2011

Trial Registration	ISRCTN99834806
Trial name	Community mobilisation and health management committee strengthening to increase birth attendance by trained health workers in rural Makwanpur, Nepal: study protocol for a cluster randomised controlled trial
Objective	Test the effect of community mobilisation through women's groups, and health management committee strengthening, on institutional deliveries and home deliveries attended by trained health workers in Makwanpur District.
Methods	C-RCT

Unit of Randomisation	Village development committee
Country	Nepal
Participants	Women who deliver infants at gestation 28 complete weeks or greater; women of reproductive age (12 - 49 years old); and infants born within the study period.
Interventions	The intervention has two components, community mobilisation through women's groups, and health management committee (HMC) strengthening.
Control	Study will train all health workers in the district in Obstetric First Aid and re-supply essential newborn care equipment. This ensures a minimum level of care in the study area.
Outcomes	Primary: deliveries conducted by a trained health worker; institutional deliveries. Secondary: neonatal deaths; stillbirths; maternal morbidities; care behaviours
Starting Date (dd/mm/yyyy)	01/10/2010
Contact Information	Professor Anthony Costello Institute of Child Health University College London 30 Guilford Street London WC1N 1EH United Kingdom +44 (0)20 7242 9789 a.costello@ich.ucl.ac.uk
Funding Source	Wellcome Trust

Patel 2014

Trial Registration	ISRCTN76465238
Trial name	The effectiveness and cost-effectiveness of lay counsellor-delivered psychological treatments for harmful and dependent drinking and moderate to severe depression in primary care in India: PREMIUM study protocol for randomised controlled trials
Objective	Evaluate the effectiveness and cost-effectiveness of two psychological treatments developed for the treatment of depression and alcohol use disorders in primary care in India.
Methods	Parallel group RCT
Unit of Randomisation	Patients
Country	India
Participants	Adult primary care attendees will be screened with the Patient Health Questionnaire for depression and, in men only, the Alcohol Use Disorders Identification Test for drinking problems. Screen-positive attendees will be invited to participate
Interventions	Psychological treatments (Healthy Activity Program for moderate to severe depression, Counselling for Alcohol Problems for

	harmful and dependent drinking) will be delivered by lay counsellors, over a maximum period of three months.
Control	The enhanced usual care comprises providing primary health centre doctors with contextualized World Health Organization guidelines and screening results.
Outcomes	Primary outcomes are severity of disorder and remission rates at three months post-enrolment and, for the Counselling for Alcohol Problems trial, drinking and the impact of drinking on daily lives. Secondary outcomes include severity of disorder and remission rates at 12 months, disability scores, suicidal behaviour and economic impact, and cost-effectiveness at three and 12 months.
Starting Date (dd/mm/yyyy)	Oct-13
Contact Information	Prof Vikram Patel Department of Population Health London School of Hygiene and Tropical Medicine Keppel Street London WC1E 7HT United Kingdom +91 (0)98221 32038 vikram.patel@lshtm.ac.uk
Funding Source	The Wellcome Trust (UK)

Ramsey 2013

Trial Registration	ISRCTN96819844
Trial name	The Tanzania Connect Project: a cluster-randomised trial of the child survival impact of adding paid community health workers to an existing facility-focused health system
Objective	Connect aims to test whether introducing a community health worker into a general programme of health systems strengthening and referral improvement will reduce child mortality, improve access to services, expand utilization, and alter reproductive, maternal, newborn and child health seeking behaviour; thereby accelerating progress towards Millennium Development Goals 4 and 5.
Methods	C-RCT
Unit of Randomisation	Village
Country	Tanzania
Participants	Women of reproductive age or the primary care takers of at least one under-five year old child
Interventions	The community health agent intervention integrates various primary health care functions into a single community-based work package that includes all stages of the continuum of care.
Control	Not specified
Outcomes	Mortality rates

Starting Date (dd/mm/yyyy)	07/01/2010
Contact Information	Dr James Phillips 60 Haven Ave B-2 New York 10032 USA
Funding Source	African Health Initiative of the Doris Duke Charitable Foundation. Comic Relief UK also supports Connect interventions and research. Additional support was provided by the Bill and Melinda Gates Foundation for the formative research on emergency referral and the Maternal Health Task Force to initiate the mobile phone component of the project.

Sando 2014

Trial Registration	EJF22802
Trial name	Trial to Improve Access to PMTCT Services and Reduce HIV Transmission From Mother to Child (FS)
Objective	The purpose of this study is to determine the effectiveness, cost-effectiveness, feasibility and acceptability of an enhanced community health worker (CHW) intervention and outreach system to improve antenatal care and PMTCT uptake and retention, and to decrease mother-to-child HIV transmission.
Methods	C-RCT
Unit of Randomisation	Administrative wards
Country	Tanzania
Participants	All pregnant women who are identified by the CHW during the routine household visits
Interventions	In the wards receiving the CHW intervention, the CHWs: (1) identify pregnant women through home visits and refer them to antenatal care; (2) provide education to pregnant women on antenatal care, PMTCT, birth, and postnatal care; (3) routinely follow up on all pregnant women to ascertain whether they have attended ANC; and (4) follow up on women who have missed ANC or PMTCT appointments.
Control	Standard of care
Outcomes	ANC and PMTCT outcomes
Starting Date (dd/mm/yyyy)	01/01/2013
Contact Information	Till Barnighausen, Associate Professor of Global Health Harvard School of Public Health
Funding Source	Comic Relief - UK and the Elton John AIDS Foundation.

Sikander 2015

Trial	NCT02111915 (Pakistan) & NCT02104232 (India)
--------------	--

Registration	
Trial name	The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials
Objective	Evaluate the effectiveness and cost-effectiveness of the Thinking Healthy Programme delivered by peers (Thinking Healthy Programme Peer-delivered; THPP), for women with moderate to severe perinatal depression in rural and urban settings in Pakistan and India.
Methods	C-RCT (Pakistan); RCT (India)
Unit of Randomisation	Villages (Pakistan); Patient (India)
Country	Pakistan and India
Participants	Pregnant women who are registered with the lady health workers in the study area in Pakistan and pregnant women attending outpatient antenatal clinics in India. They will be screened using the patient health questionnaire-9 (PHQ-9) for depression symptoms and will be eligible if their PHQ-9 is equal to or greater than 10 (PHQ-9 $\geq$ 10).
Interventions	Women in the intervention arm (THPP) will be offered ten individual and four group sessions (Pakistan) or 6–14 individual sessions (India) delivered by a peer (defined as a mother from the same community who is trained and supervised in delivering the intervention).
Control	Enhanced usual care: will receive health care as usual, enhanced by providing the gynaecologist or primary-health facilities with adapted WHO mhGAP guidelines for depression treatment, and providing the woman with her diagnosis and information on how to seek help for herself.
Outcomes	Primary: remission and severity of depression symptoms at the 6-month postnatal follow-up. Secondary: remission and severity of depression symptoms at the 3-month postnatal follow-up, functional disability, perceived social support, breastfeeding rates, infant height and weight, and costs of health care at the 3- and 6-month postnatal follow-ups.
Starting Date (dd/mm/yyyy)	10/2014 (Both)
Contact Information	Prof Vikram Patel (Pakistan) Department of Population Health London School of Hygiene and Tropical Medicine Keppel Street London WC1E 7HT United Kingdom +91 (0)98221 32038 vikram.patel@lshtm.ac.uk

	Siham Sikander, PhD (India) siham.sikander@hdrfoundation.org
Funding Source	NIMH, USA, under award number U19MH95687

Scott 2014

Trial Registration	ISRCTN37259296
Trial name	Community-based scheduled screening and treatment of malaria in pregnancy for improved maternal and infant health in The Gambia, Burkina Faso and Benin: study protocol for a randomised controlled trial
Objective	The primary objective is to determine the protective efficacy of community-scheduled screening and treatment (CSST) using community health workers (CHW) against the primary outcome of prevalence of placental malaria.
Methods	C-RCT
Unit of Randomisation	Villages
Country	The Gambia, Burkina Faso and Benin
Participants	Pregnant women aged 16 years or older
Interventions	CHW and their related catchment population who are randomised into the intervention arm will receive specific training on community-based case management of malaria in pregnancy.
Control	Current treatment
Outcomes	The primary outcome is prevalence of placental malaria. The secondary outcomes are maternal anaemia, maternal peripheral infection, low birth weight, selection of sulfadoxine-pyrimethamine (SP) resistance markers, and on antenatal clinic (ANC) attendance and coverage of intermittent preventive treatment during pregnancy (IPTp-SP).
Starting Date (dd/mm/yyyy)	01/01/2013
Contact Information	Dr Henk Schallig Royal Tropical Institute (KIT) Meibergdreef 39 Amsterdam 1105 AZ Netherlands +31 20 5665447 h.schallig@kit.nl
Funding Source	European Union Seventh Framework Programme

Shrestha 2011

Trial Registration	ISRCTN87820538
Trial name	Community interventions to reduce child mortality in Dhanusha, Nepal: study protocol for a cluster randomised controlled trial

Objective	The objective of this trial is to evaluate the effects on newborn health of two community-based interventions involving Female Community Health Volunteers.
Methods	Factorial C-RCT
Unit of Randomisation	Village development committee clusters
Country	India
Participants	Women who either join community groups or have a pregnancy.
Interventions	Female Community Health Volunteers will be supported in convening monthly women's groups and will be trained to care for vulnerable newborn infants.
Control	Not specified
Outcomes	Primary: Neonatal mortality rates Secondary: Care practices and health care-seeking behaviour, Stillbirth rates, Maternal mortality ratios, Sepsis management, Identification and treatment of neonatal sepsis by community health volunteers, Causes of death assessed by community verbal autopsy
Starting Date (dd/mm/yyyy)	01/01/2013
Contact Information	Professor Anthony Costello UCL Centre for International Health and Development 30 Guilford Street London WC1N 1EH United Kingdom
Funding Source	An anonymous Swiss Foundation for covering most of the costs involved since 2004; USAID for assistance 2005-8; DFID Towards 4 + 5 Research Programme Consortium for contributing to the costs of capacity building and staffing of the sepsis management intervention during 2008-10; "Population Science of Maternal and Child Survival" Wellcome Trust Strategic Award for covering technical advisor and some principal investigator staff costs since Nov 2009

Siabani 2014

Trial Registration	ACTRN12614000788673
Trial name	A randomised controlled trial evaluating self-care among three groups of patients with heart failure educated by health educators, by volunteers and a control group
Objective	Assess a new educational strategy involving community health volunteers (CHVs) that could reduce the cost and, hypothetically, increase the effectiveness of self-care education in patients with chronic heart failure (CHF).
Methods	RCT
Unit of Randomisation	Patient
Country	Iran

Participants	Aged > 18 years, diagnosed with heart failure
Interventions	Patients at a referral cardiovascular hospital in Iran were randomly allocated into three groups (A) trained by community health volunteers at patients' homes.
Control	(B) Trained by formal health professionals at hospital and (C) control group with no formal educational exposure.
Outcomes	The self-care components including self-care maintenance, self-care management and self-care confidence would be the main outcome variables of the study
Starting Date (dd/mm/yyyy)	09/01/2013
Contact Information	Dr Soraya Siabani Victor Coppleson Building - D02 Menzies Center for Health Policy Medical School The University of Sydney NSW 2006 Australia +61478262727 ssia5034@uni.sydney.edu.au
Funding Source	Kermanshah University of Medical Sciences in Iran and the University of Sydney in Australia

Souza 2010

Trial Registration	None
Trial name	Community health workers effectiveness in diabetes education
Objective	Evaluate the effectiveness of a diabetes education programme delivered to CHW in improving the metabolic control of patients with type 2 diabetes.
Methods	RCT
Unit of Randomisation	Patient
Country	Brazil
Participants	Diabetes patients
Interventions	One-month (four sessions, two hours each) diabetes education programme delivered by CHWs
Control	An education course in other health issues delivered by CHWs
Outcomes	Primary: Change in A1C. Secondary: effect in weight, blood pressure, lipids and acquired learning 4-month after the intervention.
Starting Date (dd/mm/yyyy)	N.S.
Contact Information	C. F. Souza Endocrinology, Hospital de Clínicas de Porto Alegre Brazil
Funding Source	N.S.

Tripathy 2011

Trial Registration	ISRCTN31567106
Trial name	Community mobilisation with women's groups facilitated by Accredited Social Health Activists (ASHAs) to improve maternal and newborn health in underserved areas of Jharkhand and Orissa: study protocol for a cluster-randomised controlled trial
Objective	The study aims to assess the impact of a community mobilisation intervention with women's groups facilitated by ASHAs to improve maternal and newborn health outcomes among rural tribal communities of Jharkhand and Orissa.
Methods	C-RCT
Unit of Randomisation	A rural cluster of approximately 5000 population
Country	India
Participants	Women who give birth during the study period
Interventions	Trained and incentivised Accredited Social Health Activists and convene women's groups
Control	In all clusters, control and intervention, activities are implemented to strengthen Village Health Committees
Outcomes	(a) Birth outcomes, including neonatal mortality and stillbirths (b) Care practices and health care-seeking behaviour for mothers and newborns
Starting Date (dd/mm/yyyy)	01/01/2013
Contact Information	Institute of Child Health Centre for International Health and Development 30 Guilford Street London WC1N 1EH United Kingdom +44 (0)20 7905 2261 g.eroglu@ich.ucl.ac.uk
Funding Source	The Big Lottery Fund (UK) and the Wellcome Trust

Vedanthan 2014

Trial Registration	NCT01844596
Trial name	Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study)
Objective	This study investigates whether community health workers, equipped with a tailored behavioural communication strategy and smartphone technology, can increase linkage and retention of hypertensive individuals to a hypertension care programme and significantly reduce blood pressure among them
Methods	3 Arm RCT
Unit of Randomisation	Community unit

Country	Kenya
Participants	Those >18 years with elevated blood pressure (SBP > 140 or DBP >90)
Interventions	A: Community Health Workers with an additional tailored behavioural communication strategy. B: Community Health Workers with a tailored behavioural communication strategy, also equipped with smartphone-based tool linked to the AMPATH Medical Record System (AMRS).
Control	In the usual care arm, the CHW will first measure the individual's BP. If it is elevated (SBP >140 or DBP >90), the CHW will refer the individual to the dispensary for further evaluation and management, as per the usual care protocol.
Outcomes	The co-primary outcome measures are: 1) linkage to hypertension care, and 2) one-year change in systolic blood pressure among hypertensive individuals. Cost-effectiveness analysis will be conducted.
Starting Date (dd/mm/yyyy)	01/04/2014
Contact Information	Rajesh Vedanthan, MD, MPH 212-659-9180 rajesh.vedanthan@mssm.edu
Funding Source	National Heart, Lung, and Blood Institute of the National Institutes of Health

Yan 2014

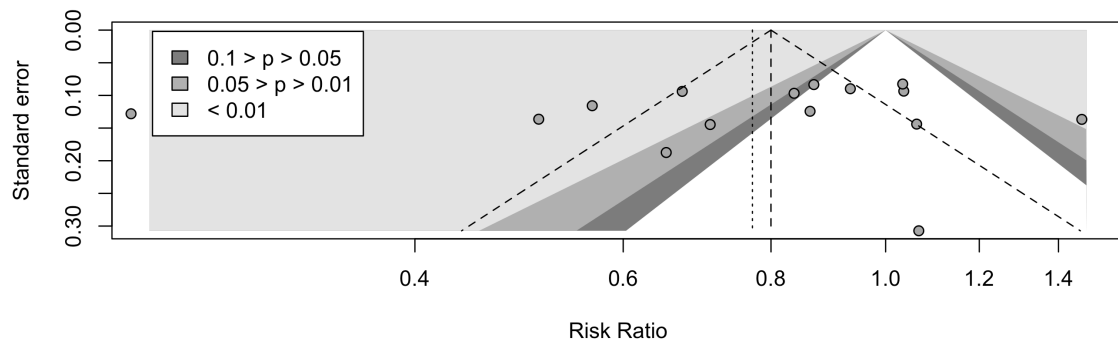
Trial Registration	NCT01259700
Trial name	Population impact of a high cardiovascular risk management programme delivered by village doctors in rural China: design and rationale of a large, cluster-randomised controlled trial
Objective	This study aims to assess, at the population level, the effects of an evidence-based high cardiovascular risk management programme delivered by village doctors in rural China.
Methods	C-RCT
Unit of Randomisation	Villages
Country	China
Participants	Participants had physician-diagnosed history of coronary heart disease, ischemic stroke, or haemorrhagic stroke or having physician-diagnosed Type I or Type II diabetes or systolic blood pressure 160 mmHg
Interventions	Village doctors in intervention villages will be trained to implement a simple evidence-based management programme designed to identify, treat and follow-up as many as possible individuals at high-risk of cardiovascular disease in the village. The intervention will also include performance feedback as well as a performance-based incentive payment scheme.
Control	Usual care
Outcomes	Primary: Mean systematic blood pressure level, 24 hour urinary

	sodium
Starting Date (dd/mm/yyyy)	01/12/2010
Contact Information	Yangfeng Wu, PhD The George Institute for Global Health China
Funding Source	United Health Group Chronic Disease Initiative & United States National Heart, Lung, and Blood Institute, National Institutes of Health

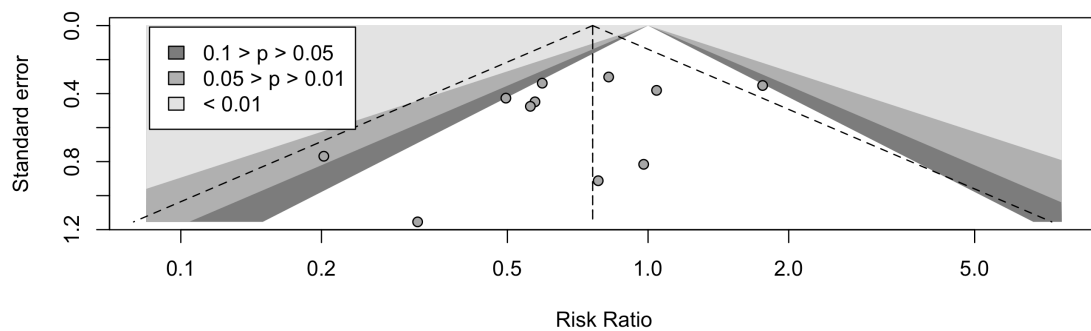
List of trials awaiting classification

Author	Year	Title
Chao	2012	The effect of community-based health management on the health of the elderly: a randomised controlled trial from China
Hughes	2009	Randomised, controlled trial of a perinatal psycho-social intervention for postnatal depression in Goa, India
Kaona	2003	Improving ability to identify malaria and correctly use chloroquine in children at household level in Nakonde District, Northern Province of Zambia.
Leslie	2004	Compliance with 14-day primaquine therapy for radical cure of vivax malaria – a randomised placebo-controlled trial comparing unsupervised with supervised treatment
Malwenna	2012	Effectiveness of a community based health educational intervention in reducing unmet need for modern methods of family planning among ever married reproductive age women in the Kalutara district, Sri Lanka.
McKenzie	2011	Community-based interventions to improve neonatal survival in low-resource settings.
Pant	1998	A randomised controlled trial of effectiveness of school children versus health workers in imparting health education about cancer in hills of Uttar Pradesh, India
Qureshi	2011	Using community volunteers to promote exclusive breastfeeding in Sokoto State, Nigeria
Rabadi	2013	Community based intervention improves infant feeding practices in Bethlehem villages
Rasaily	2013	Effectiveness of Home Based Care of Young Infants by Community Based Workers in Reducing Mortality Rates.
Sharma	1987	Utilising dais to impart health and nutrition education: impact on growth.

Funnel plots



Funnel plot for outcome 3.2 Neonatal Mortality



Funnel plot for outcome 3.6 Maternal Mortality

Quality of included economic evaluations

Item		Aung 2014	Baqui 2008	Bari 2011	Bojang 2011	CDI Study Group 2010	Chang 2010	Clarke 2005	Colbourn 2013	Datiko 2009	Fottrell 2013	Jafar 2009	Jaffar 2009	Kirkwood 2013	Kweku 2009	Lewycka 2013	Manandhar 2004	Sankaranarayanan 2005	Tylleskar 2011 (UG)	Tylleskar 2011 (SA)
1	Is the study population clearly described?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	Are competing alternatives clearly described?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	Is a well-defined research question posed in answerable form?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4	Is the economic study design appropriate to the stated objective?	+	+		+			+	+	+	+	+		+		+	+	+		
5	Is the chosen time horizon appropriate to include relevant costs and consequences?	+	+		-			-	+	-	-	+		+		+	-	+		

6	Is the actual perspective chosen appropriate?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
7	Are all important and relevant costs for each alternative identified?	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
8	Are all costs measured appropriately in physical units?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
9	Are costs valued appropriately?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
10	Are all important and relevant outcomes for each alternative identified?	-	+		+			+	+	+	+	+		+		+	+	+	
11	Are all outcomes measured appropriately?	+	+		+			+	+	+	+	+		+		+	+	+	
12	Are outcomes valued appropriately?	+	+		+			+	+	+	+	+		+		+	+	+	
13	Is an incremental analysis of costs and outcomes of alternatives performed?	+	+		+			-	+	+	-	+		+		-	+	+	
14	Are all future costs and outcomes discounted appropriately?	+	+	-	-	-	-	+	+	-	+	-	+	+	-	+	-	+	+
15	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	+	+	-	-	-	+	-	+	+	-	+	-	+	+	-	+	+	-
16	Do the conclusions follow from the data reported?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
17	Does the study discuss the generalizability of the results to other settings and patient/client groups?	+	+		-			-	-	+	-	+		+		-	-	+	
18	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+
19	Are ethical and distributional issues discussed appropriately?	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-

Appendix 4d: Data And Analyses

The R markdown file used to produce these calculations can be found and

downloaded at: <http://dx.doi.org/10.5281/zenodo.47953>

MENTAL HEALTH OUTCOMES**Comparison 1: CHW interventions for treatment of common mental disorders**

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
1.1 Prevalence of depression – adjusted for clustering	2	1244	Risk Ratio (Random, 95% CI)	0.46 [0.38, 0.56]
Prevalence of depression – unadjusted	2	1244	Risk Ratio (Random, 95% CI)	0.46 [0.39, 0.54]
1.2 Severity of common mental disorder symptoms (includes anxiety and depression) – adjusted for clustering	7	2268	Standardised Mean Difference (Random, 95% CI)	-0.64 [-0.95, -0.32]
Severity of common mental disorder symptoms (includes anxiety and depression) – unadjusted	7	2268	Standardised Mean Difference (Random, 95% CI)	-0.62 [-0.94, -0.31]
1.3 Functional impairment of adults with common mental disorders – adjusted for clustering	3	1392	Standardised Mean Difference (Random, 95% CI)	-0.44 [-0.80, -0.09]
Functional impairment of adults with common mental disorders – unadjusted	3	1392	Standardised Mean Difference (Random, 95% CI)	-0.45 [-0.80, -0.10]
1.4 Prevalence of PTSD – unadjusted	2	390	Standardised Mean Difference (Random, 95% CI)	-0.54 [-0.81, -0.28]

1.5 Severity of depression symptoms in adolescents – unadjusted	2	390	Standardised Mean Difference (Random, 95% CI)	-0.41 [-1.01, 0.20]
---	---	-----	--	------------------------

MATERNAL AND PAEDIATRIC OUTCOMES

Comparison 2: CHW interventions to reduce mortality and morbidity in children under five

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
2.1 Mortality among children under five years – adjusted for clustering	3	106378	Risk Ratio (Random, 95% CI)	0.65 [0.38, 1.10]
Mortality among children under five years – unadjusted	3	106378	Risk Ratio (Random, 95% CI)	0.76 [0.60, 0.98]
2.2 Child morbidity – adjusted for clustering	4	9453	Risk Ratio (Random, 95% CI)	0.70 [0.51 to 0.96]
Child morbidity – unadjusted	4	9453	Risk Ratio (Random, 95% CI)	0.70 [0.52 to 0.94]
2.3 Infant mortality – adjusted for clustering	2	9459	Risk Ratio (Random, 95% CI)	0.99 [0.80 to 1.23]
Infant mortality – unadjusted	2	9459	Risk Ratio (Random, 95% CI)	0.99 [0.83 to 1.19]
2.4 Infant morbidity – adjusted for clustering	7	48768	Risk Ratio (Random, 95% CI)	0.73 [0.56, 0.96]
Infant morbidity – unadjusted	7	48768	Risk Ratio (Random, 95% CI)	0.90 [0.87, 0.92]

Comparison 3: CHW interventions to improve maternal and newborn health

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
3.1 Stillbirths – adjusted for clustering	7	84328	Risk Ratio (Random, 95% CI)	0.89 [0.69 to 1.13]
Stillbirths – unadjusted	7	84328	Risk Ratio (Random, 95% CI)	0.93 [0.81, 1.07]
3.2 Neonatal mortality – adjusted for clustering	15	157161	Risk Ratio (Random, 95% CI)	0.76 [0.61, 0.96]
Neonatal mortality – unadjusted	15	157161	Risk Ratio (Random, 95% CI)	0.78 [0.62, 0.97]
3.3 Early neonatal mortality – adjusted for clustering	7	70346	Risk Ratio (Random, 95% CI)	0.69 [0.57, 0.85]
Early neonatal mortality – unadjusted	7	70346	Risk Ratio (Random, 95% CI)	0.72 [0.59, 0.87]
3.4 Late neonatal mortality – adjusted for clustering	7	70346	Risk Ratio (Random, 95% CI)	0.70 [0.49, 1.01]
Late neonatal mortality – unadjusted	7	70346	Risk Ratio (Random, 95% CI)	0.71 [0.50, 1.02]
3.5 Perinatal mortality – adjusted for clustering	7	72646	Risk Ratio (Random, 95% CI)	0.83 [0.72, 0.95]
Perinatal mortality – unadjusted	7	72646	Risk Ratio (Random, 95% CI)	0.88 [0.82, 0.93]
3.6 Maternal mortality– adjusted for clustering	11	102775	Risk Ratio (Random, 95% CI)	0.66 [0.35, 1.25]
Maternal mortality– unadjusted	11	102775	Risk Ratio (Random, 95% CI)	0.75 [0.58, 0.97]

3.7 Maternal morbidity–adjusted for clustering	2	13238	Risk Ratio (Random, 95% CI)	1.04 [0.63, 1.73]
Maternal morbidity–unadjusted	2	13238	Risk Ratio (Random, 95% CI)	0.92 [0.69, 1.23]
3.8 Facility-based deliveries –adjusted for clustering	13	163257	Risk Ratio (Random, 95% CI)	0.94 [0.86, 1.02]
Facility-based deliveries –unadjusted	13	163257	Risk Ratio (Random, 95% CI)	0.93 [0.85, 1.02]
3.9 Use of clean delivery kits –adjusted for clustering	4	55078	Risk Ratio (Random, 95% CI)	0.89 [0.83, 0.96]
Use of clean delivery kits–unadjusted	4	55078	Risk Ratio (Random, 95% CI)	0.89 [0.83, 0.96]
3.10 Exclusive breastfeeding (6 months) –adjusted for clustering	10	9382	Risk Ratio (Random, 95% CI)	0.66 [0.52, 0.83]
Exclusive breastfeeding (6 months) –unadjusted	10	9382	Risk Ratio (Random, 95% CI)	0.65 [0.52, 0.83]
3.11 Initiation of breastfeeding within 1 h of birth– adjusted for clustering	11	75740	Risk Ratio (Random, 95% CI)	0.61 [0.45, 0.82]
Initiation of breastfeeding within 1 h of birth – unadjusted	11	75740	Risk Ratio (Random, 95% CI)	0.60 [0.45, 0.82]

Comparison 4: CHW interventions to promote cognitive development in infants

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
---------------------------	----------------	---------------------	--------------------	-------------

4.1 Maternal knowledge – adjusted for clustering	2	1699	Standardised Mean Difference (Random, 95% CI)	-1.37 [-3.03, 0.29]
Maternal knowledge – unadjusted	2	1699	Standardised Mean Difference (Random, 95% CI)	-1.37 [-3.03, 0.29]
4.2 Mental development – adjusted for clustering	7	2676	Standardised Mean Difference (Random, 95% CI)	-0.56 [-0.86, -0.27]
Mental development – unadjusted	7	2676	Standardised Mean Difference (Random, 95% CI)	-0.57 [-0.86, -0.27]

Comparison 5: CHW interventions to promote healthy feeding practices

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
5.1 Weight-for-age Z score – adjusted for clustering	3	2328	Mean Difference (Random, 95% CI)	-0.12 [-0.21, -0.02]
Weight-for-age Z score – unadjusted	3	2328	Mean Difference (Random, 95% CI)	-0.12 [-0.21, -0.02]
5.2 Height-for-age Z score– adjusted for clustering	3	2328	Mean Difference (Random, 95% CI)	-0.10 [-0.29, 0.09]
Height-for-age Z score – unadjusted	3	2328	Mean Difference (Random, 95% CI)	-0.09 [-0.28, 0.09]
5.3 Weight-for-height Z score – adjusted for clustering	2	1623	Mean Difference (Random, 95% CI)	-0.14 [-0.27, -0.01]
Weight-for-height Z score – unadjusted	2	1623	Mean Difference (Random, 95% CI)	-0.14 [-0.26, -0.02]

INFECTIOUS DISEASE OUTCOMES

Comparison 6: CHW interventions to improve TB treatment outcomes compared with other forms of adherence support

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
---------------------------	----------------	---------------------	--------------------	-------------

6.1 Cure rate – adjusted for clustering	5	1565	Risk Ratio (Random, 95% CI)	0.66 [0.42, 1.04]
Cure rate – unadjusted	5	1565	Risk Ratio (Random, 95% CI)	0.67 [0.42, 1.04]
6.2 Completion rate (vs. facility-based care) – adjusted for clustering	4	655	Risk Ratio (Random, 95% CI)	0.63 [0.45, 0.88]
Completion rate (vs. facility-based care) – unadjusted	4	655	Risk Ratio (Random, 95% CI)	0.63 [0.45, 0.88]
6.3 Completion rate (vs. family DOTS) – adjusted for clustering	2	2242	Risk Ratio (Random, 95% CI)	0.97 [0.84, 1.12]
Completion rate (vs. facility-based care) – unadjusted	2	2242	Risk Ratio (Random, 95% CI)	1.11 [0.76, 1.63]

Comparison 7: CHW interventions for maintenance of antiretroviral therapy

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
7.1 Death (12-24 months) – adjusted for clustering	2	2789	Risk Ratio (Random, 95% CI)	1.03 [0.76, 1.39]
Death (12-24 months) – unadjusted	2	2789	Risk Ratio (Random, 95% CI)	1.04 [0.83, 1.29]
7.2 Lost to follow-up (12-24 months) – adjusted for clustering	3	2997	Risk Ratio (Random, 95% CI)	0.85 [0.46, 1.54]
Lost to follow-up (12-24 months) – unadjusted	3	2997	Risk Ratio (Random, 95% CI)	0.76 [0.49, 1.18]

PUBLIC AND PREVENTATIVE HEALTH OUTCOMES

Comparison 8: CHW interventions to promote injury prevention in children

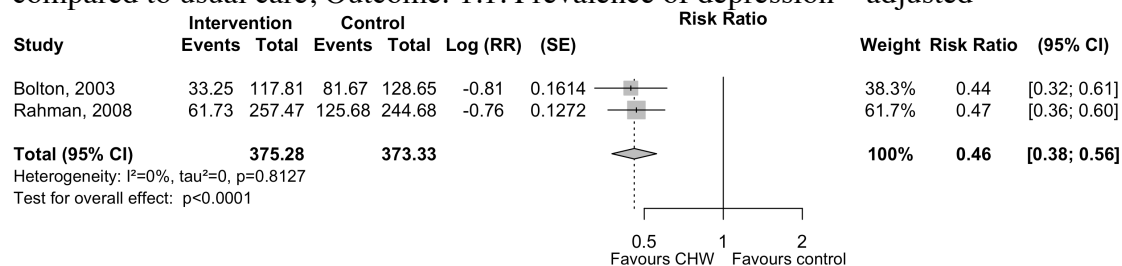
Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
8.1 Hazard score –	2	586	Mean Difference	-1.8

adjusted for clustering			(Random, 95% CI)	[-5.02, 1.42]
Hazard score – unadjusted	2	586	Mean Difference (Random, 95% CI)	-1.79 [-5.01; 1.43]

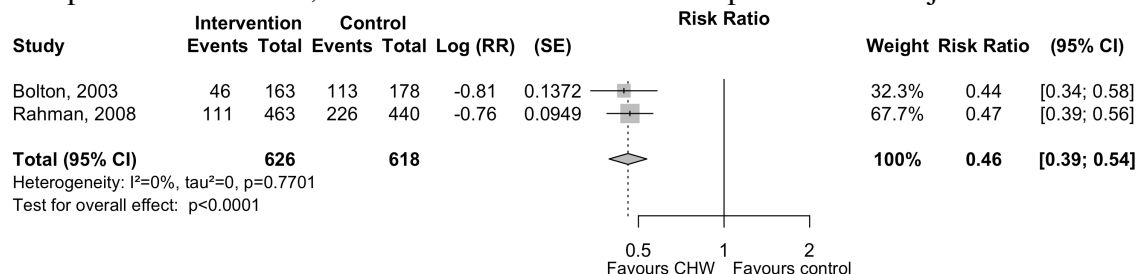
Comparison 9: CHW interventions to promote immunization uptake and mass drug administration

Outcome or subgroup title	No. of studies	No. of participants	Statistical Method	Effect Size
9.2 Mass drug administration coverage – adjusted for clustering	5	12382	Risk Ratio (Random, 95% CI)	0.56 [0.31; 1.00]
Mass drug administration coverage – unadjusted	5	12382	Risk Ratio (Random, 95% CI)	0.56 [0.31; 1.00]

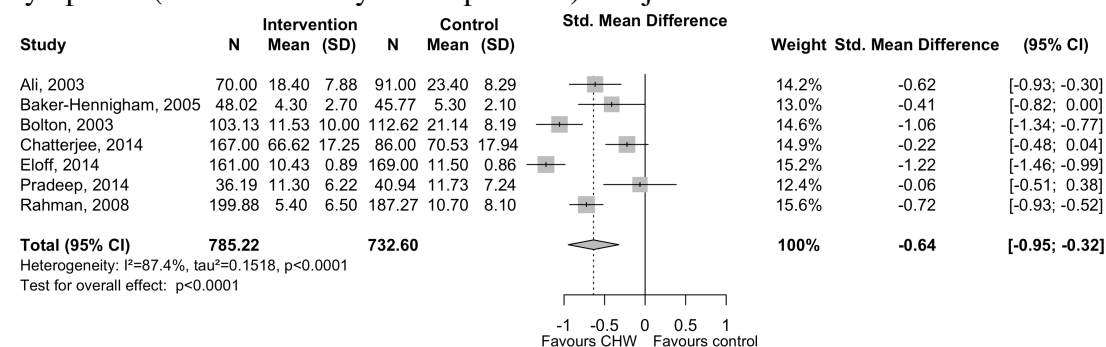
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.1. Prevalence of depression – adjusted



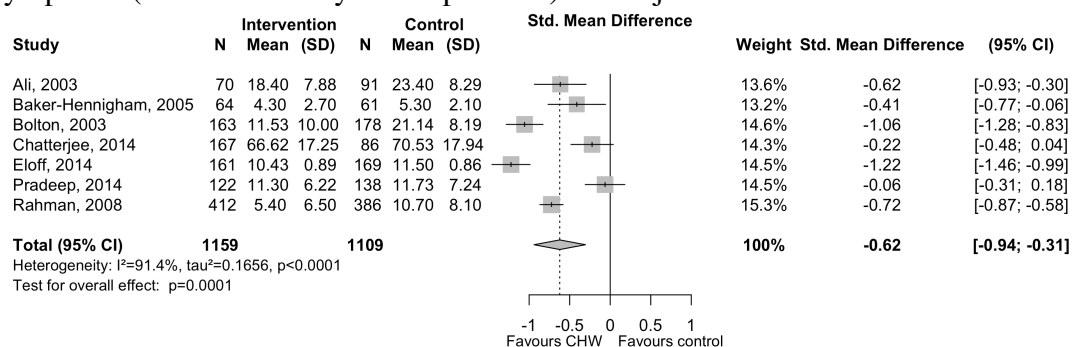
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.1. Prevalence of depression – unadjusted



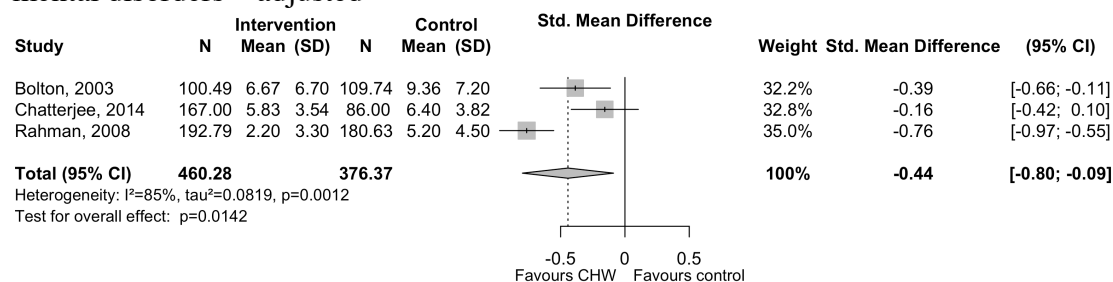
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.2. Severity of common mental disorder symptoms (includes anxiety and depression) – adjusted



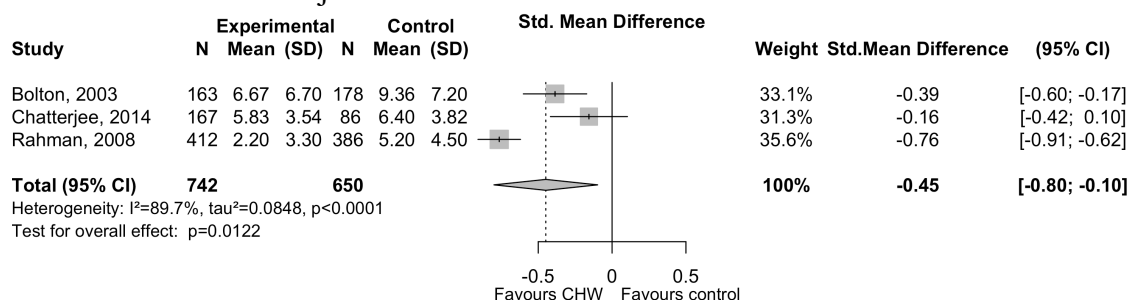
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.2. Severity of common mental disorder symptoms (includes anxiety and depression) – unadjusted



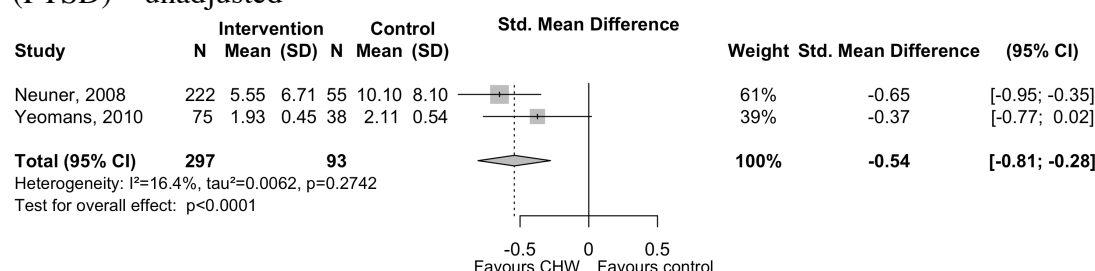
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.3. Functional impairment of adults with common mental disorders – adjusted



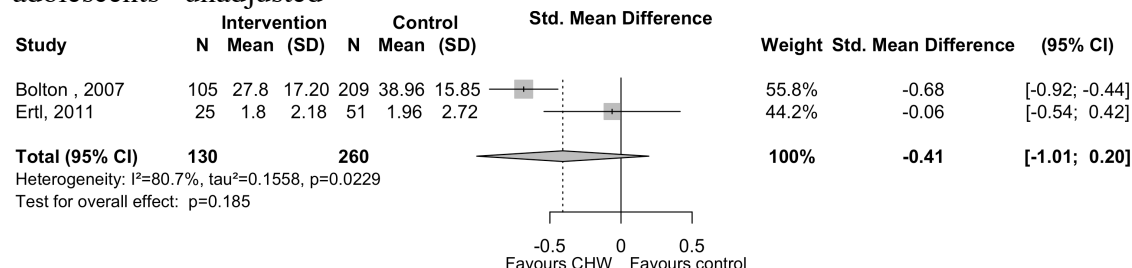
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.3. Functional impairment of adults with common mental disorders – unadjusted



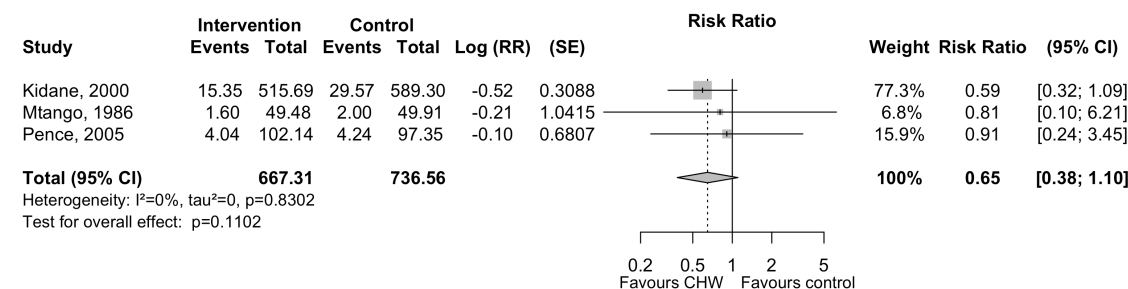
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.4. Severity of post-traumatic stress disorder (PTSD) – unadjusted



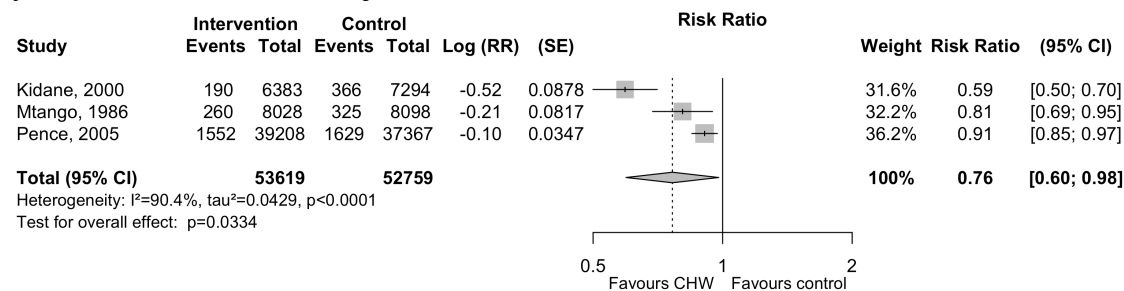
Comparison 1: CHW interventions for treatment of common mental disorders compared to usual care, Outcome: 1.5. Severity of depression symptoms in adolescents– unadjusted



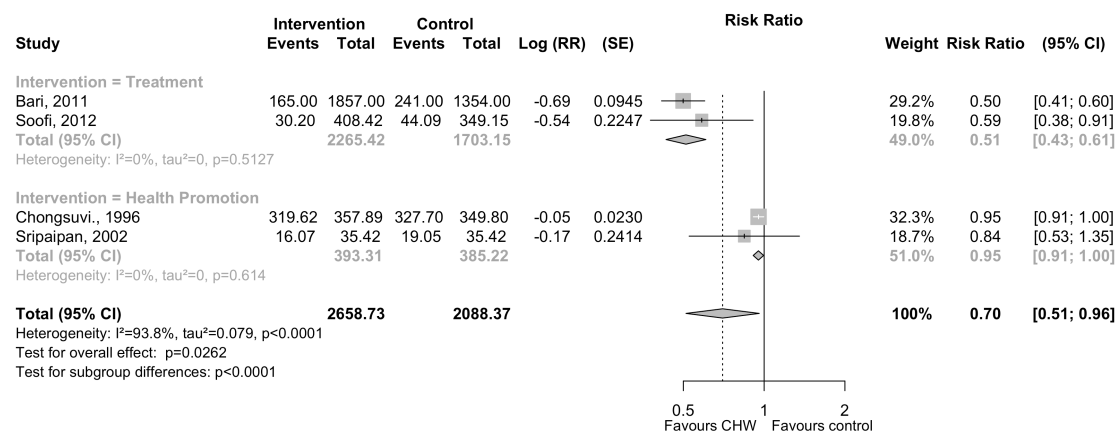
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.1 Mortality among children under five years (0-59 months) – adjusted



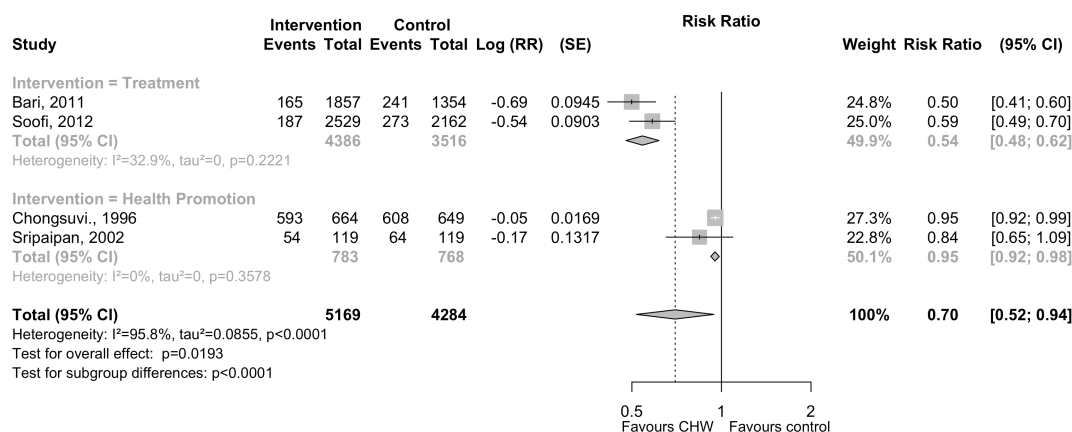
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.1 Mortality among children under five years (0-59 months) – unadjusted



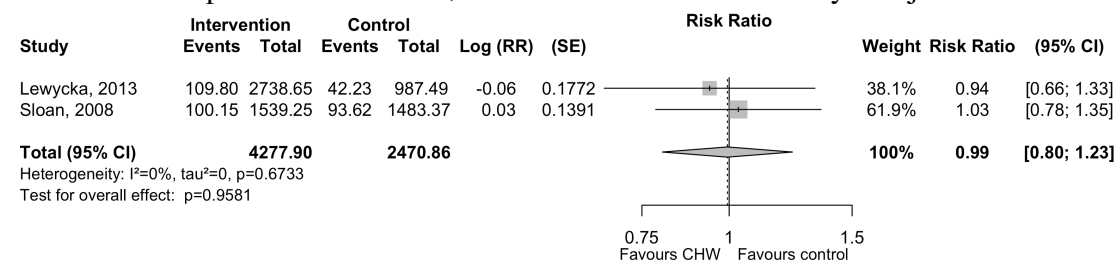
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.2 Child morbidity, reported illness in children – adjusted



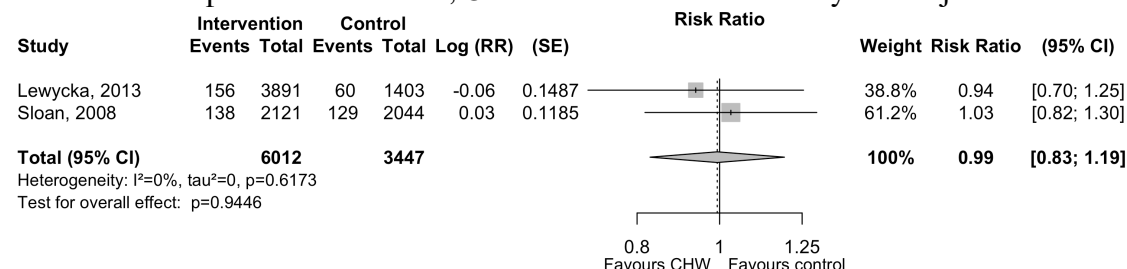
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.2 Child morbidity, reported illness in children – unadjusted



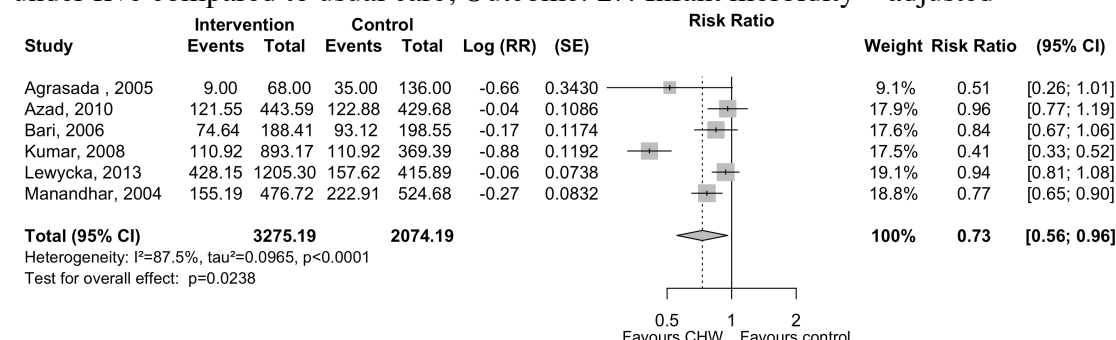
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.3 Infant mortality – adjusted



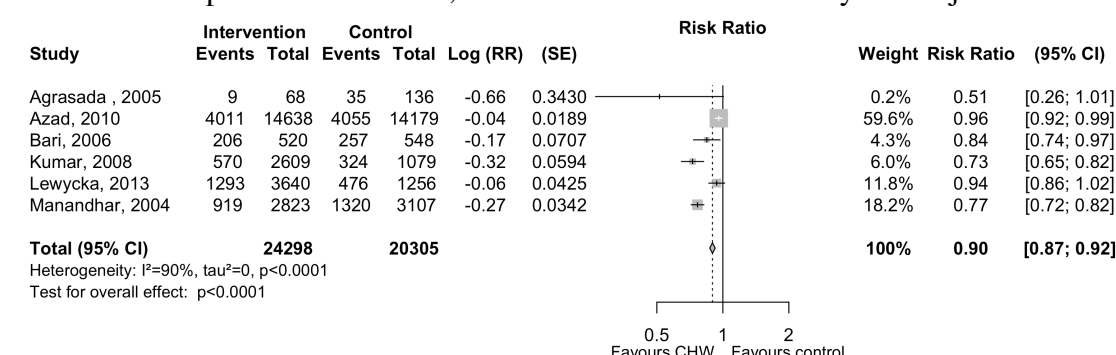
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.3 Infant mortality – unadjusted



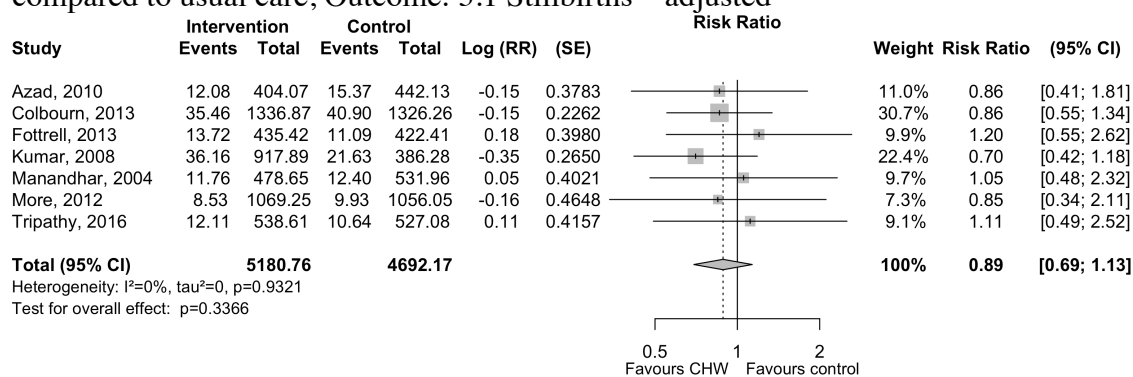
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.4 Infant morbidity – adjusted



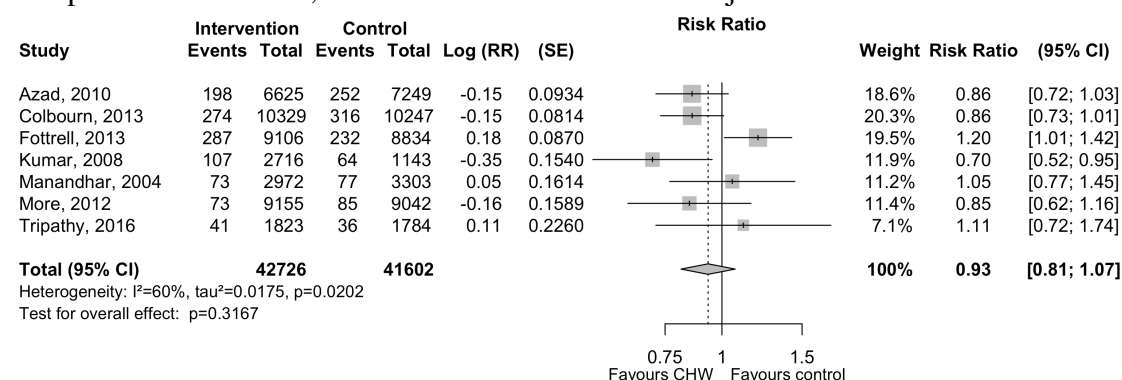
Comparison 2: CHW interventions to reduce mortality and morbidity in children under five compared to usual care, Outcome: 2.4 Infant morbidity – unadjusted



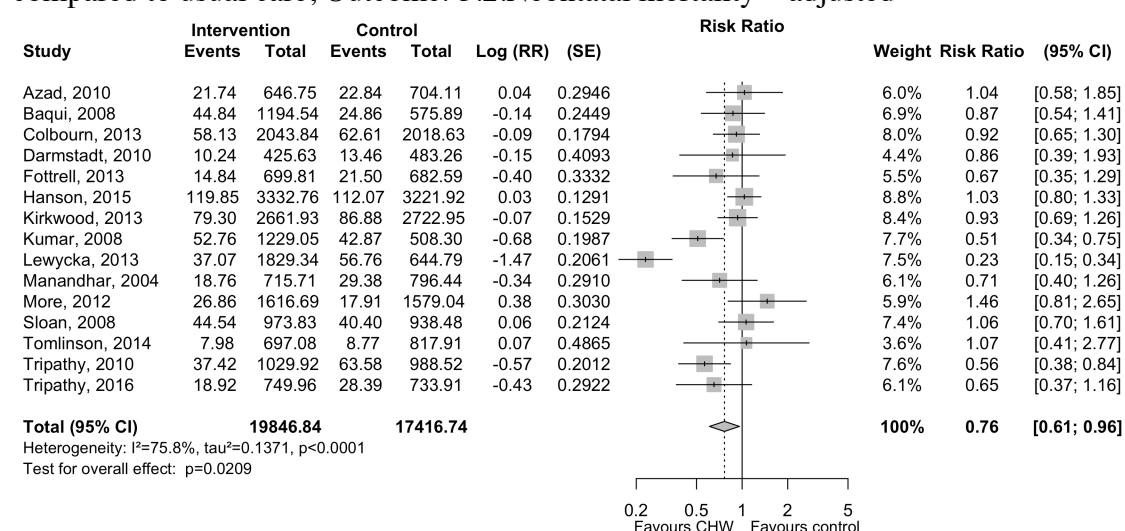
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.1 Stillbirths – adjusted



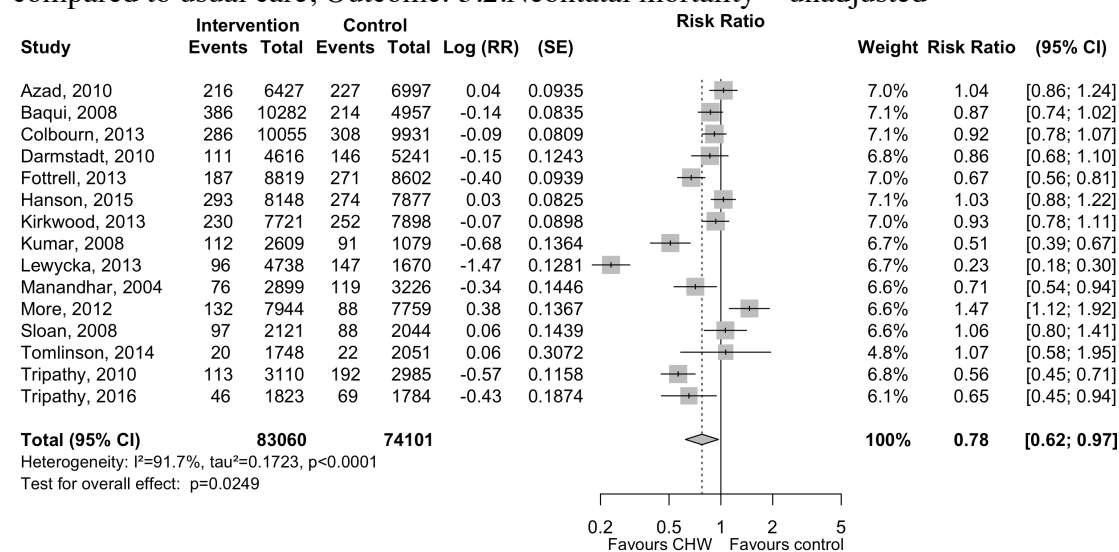
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.1 Stillbirths – unadjusted



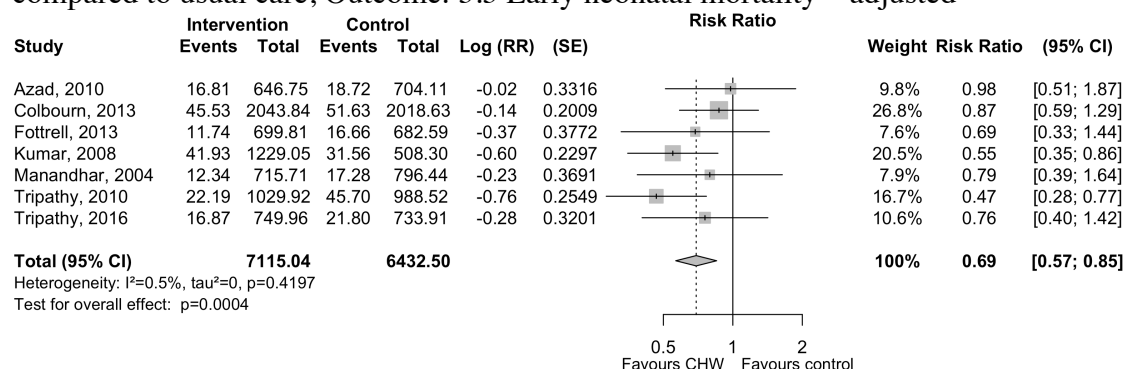
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.2 Neonatal mortality – adjusted



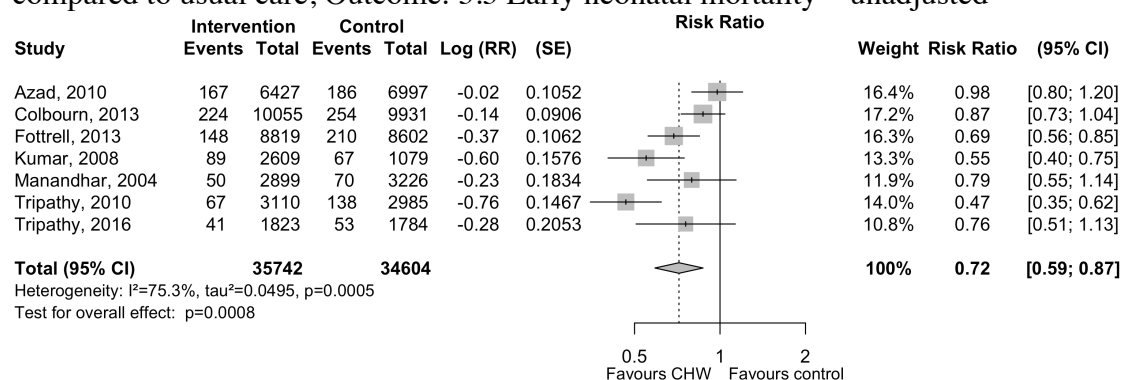
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.2. Neonatal mortality – unadjusted



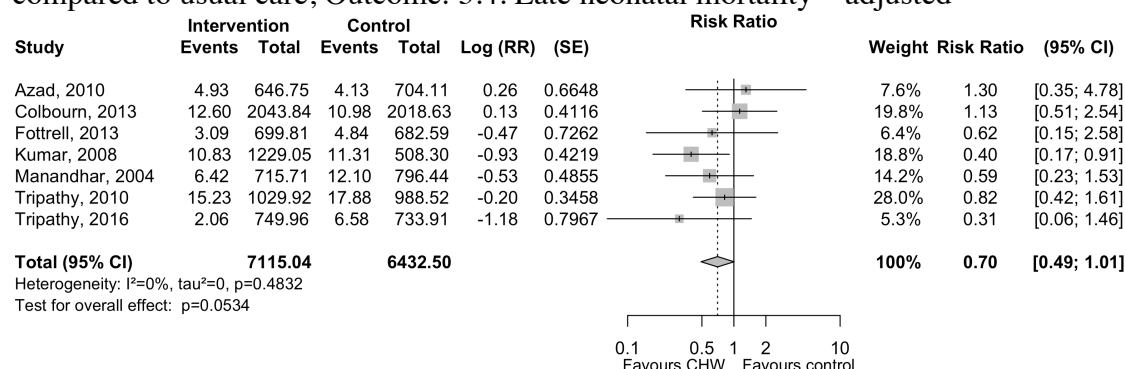
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.3 Early neonatal mortality – adjusted



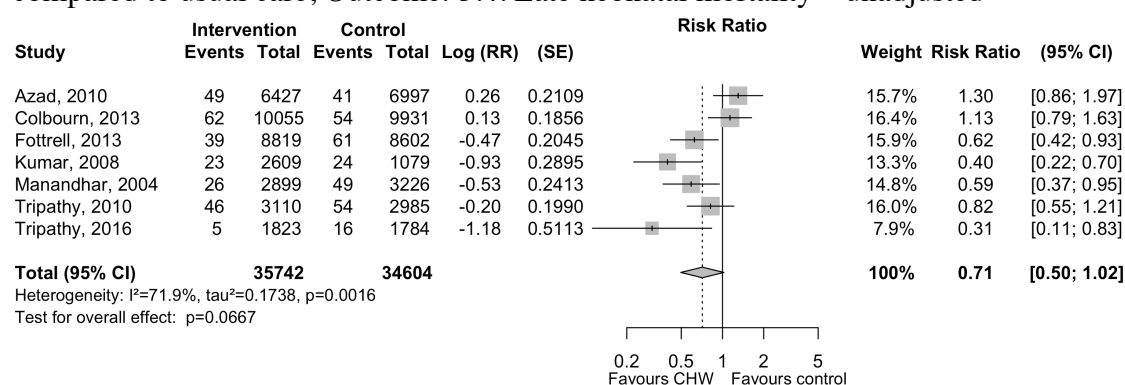
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.3 Early neonatal mortality – unadjusted



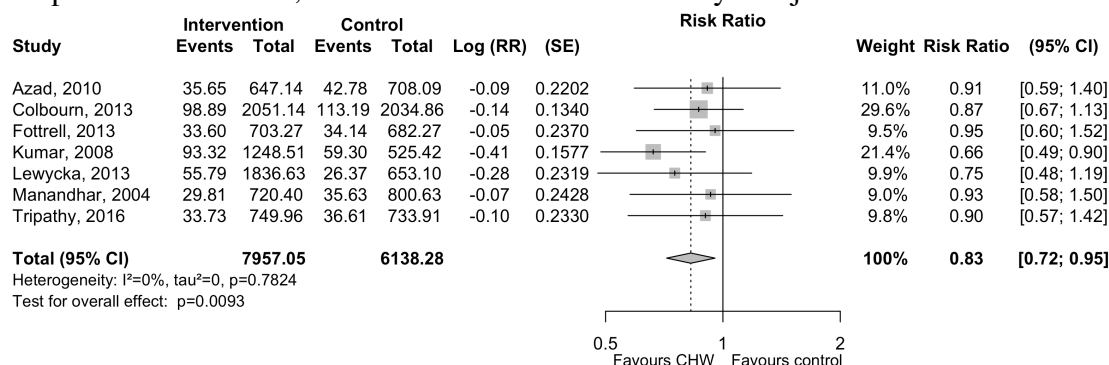
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.4. Late neonatal mortality – adjusted



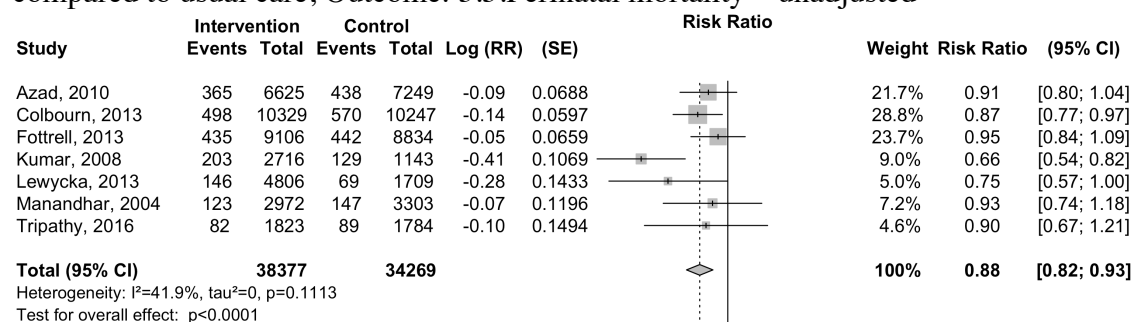
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.4. Late neonatal mortality – unadjusted



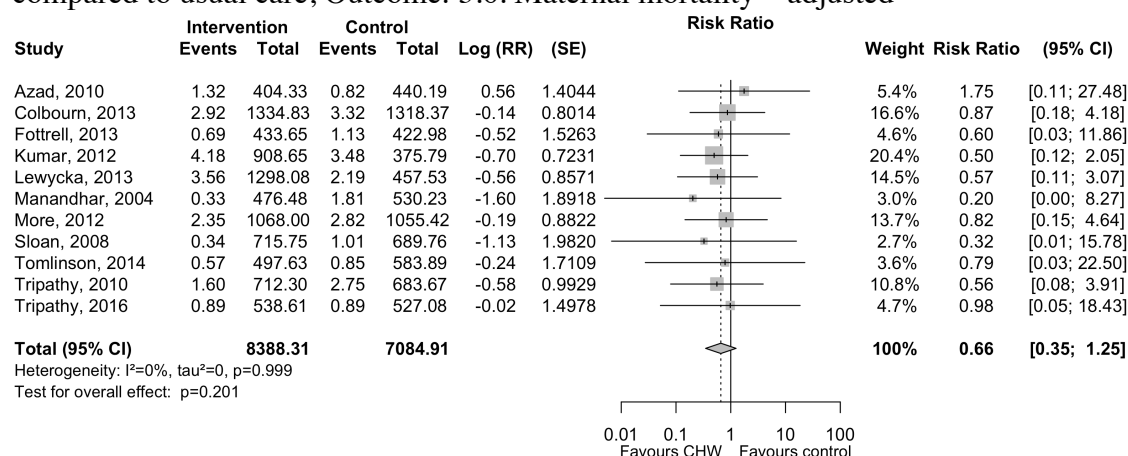
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.5. Perinatal mortality – adjusted



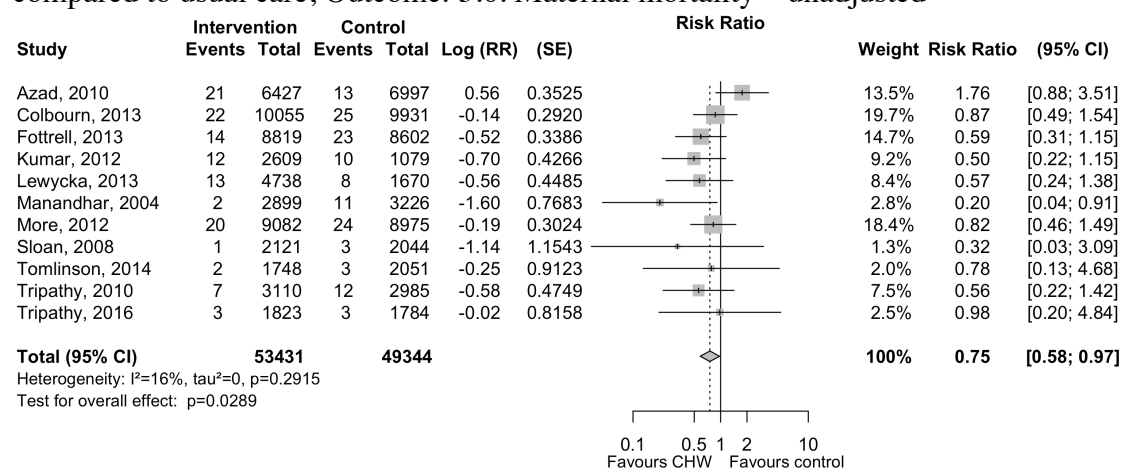
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.5.Perinatal mortality – unadjusted



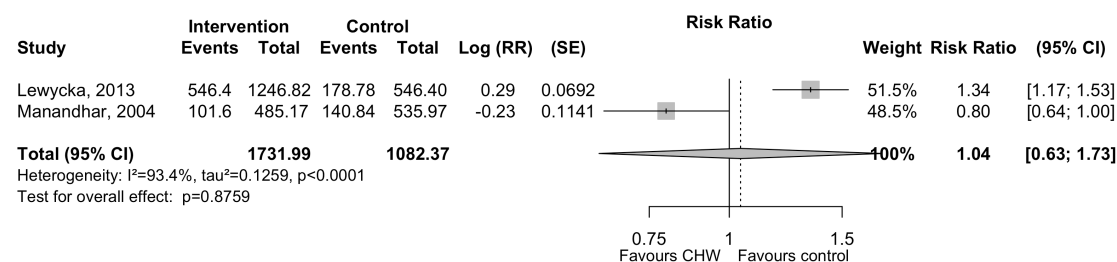
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.6. Maternal mortality – adjusted



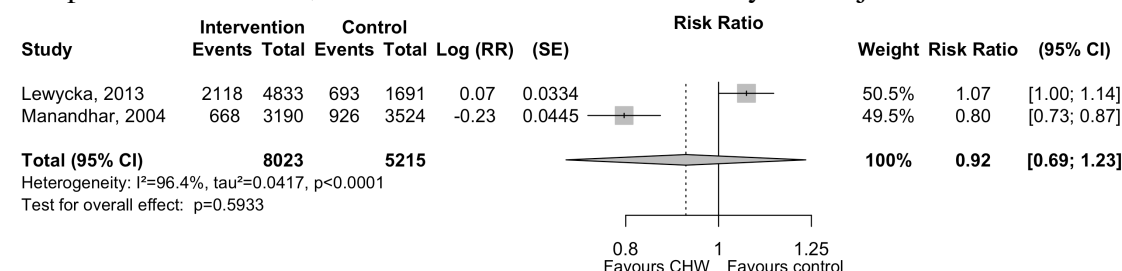
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.6. Maternal mortality – unadjusted



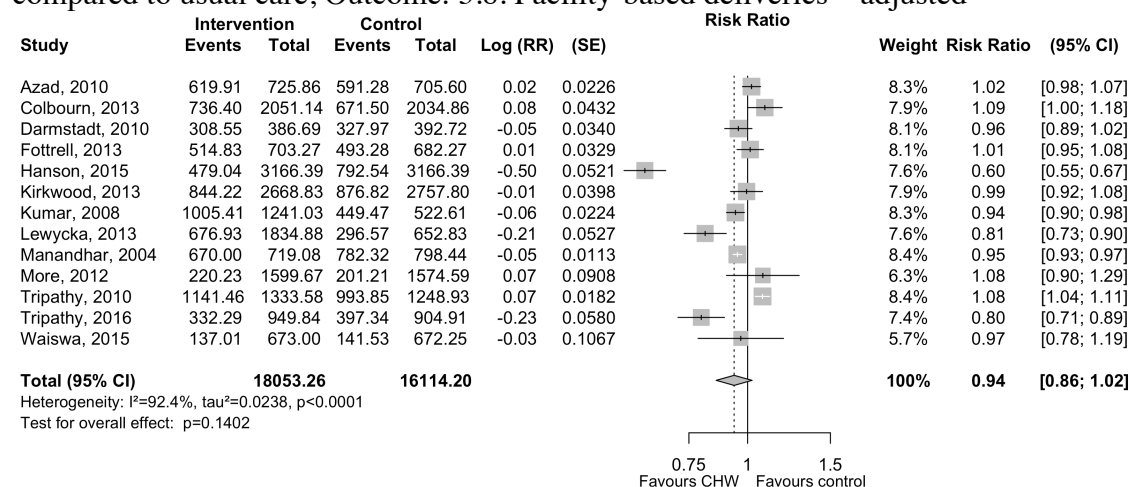
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.7. Maternal morbidity – adjusted



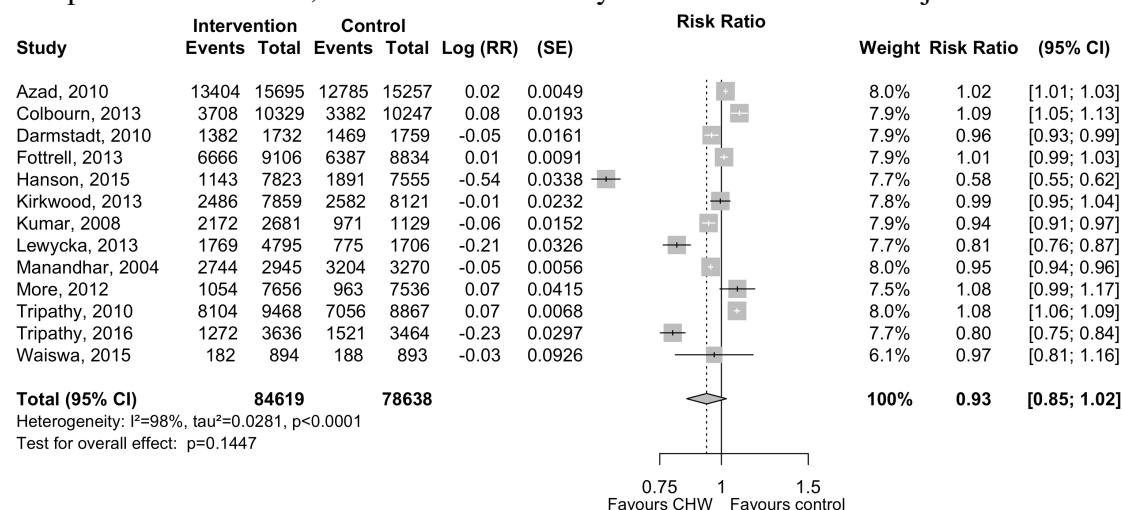
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.7. Maternal morbidity – unadjusted



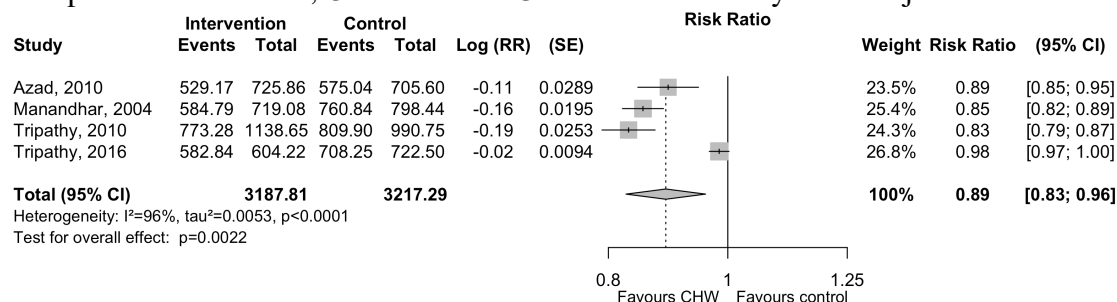
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.8. Facility-based deliveries – adjusted



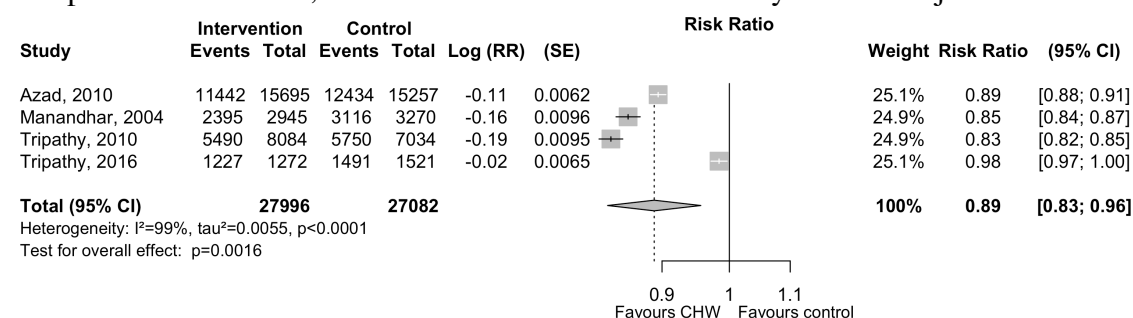
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.8. Facility-based deliveries – unadjusted



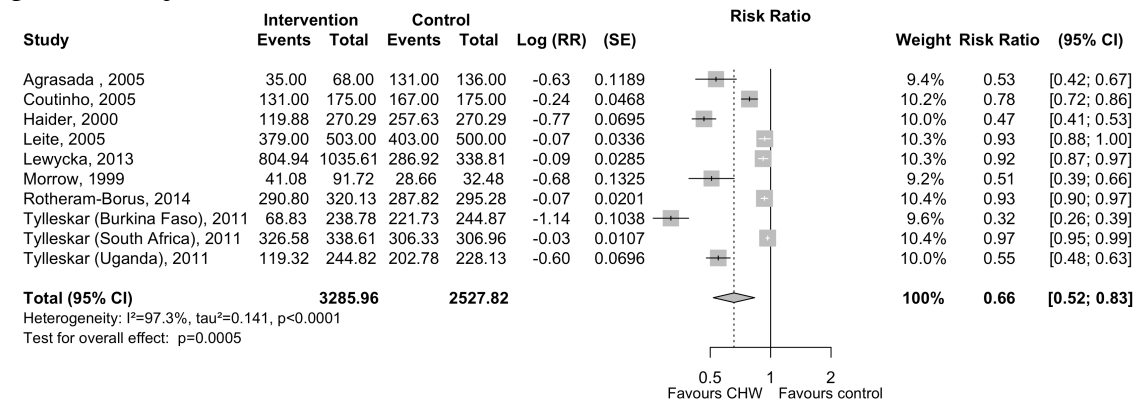
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – adjusted



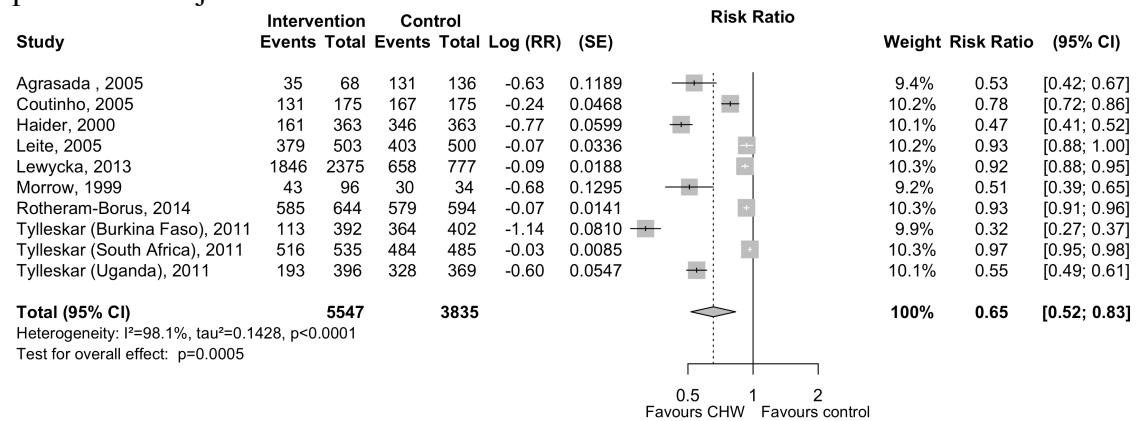
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.9. Use of clean delivery kits – unadjusted



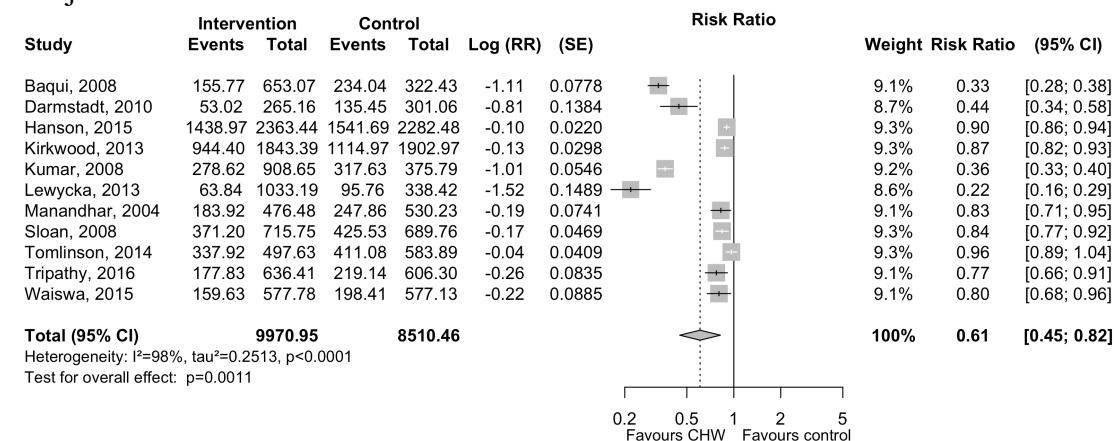
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.10. Exclusive breastfeeding up to six months post partum – adjusted



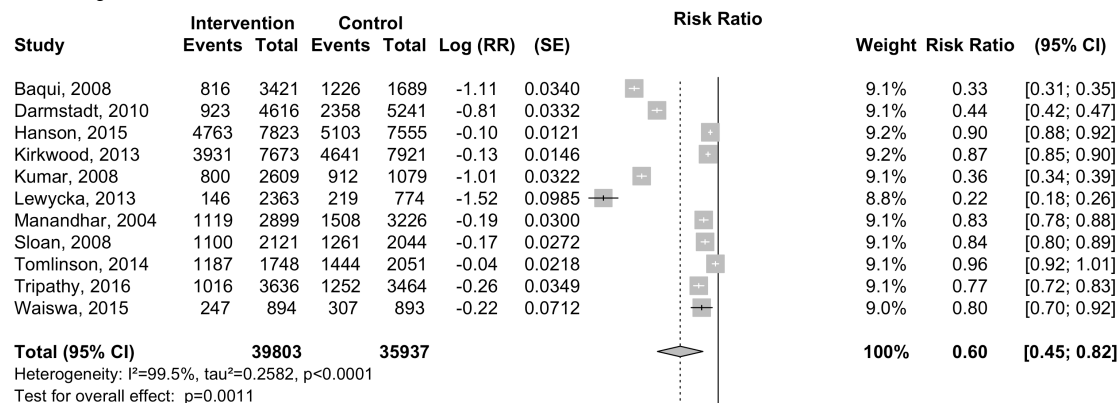
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.10. Exclusive breastfeeding up to six months post partum – unadjusted



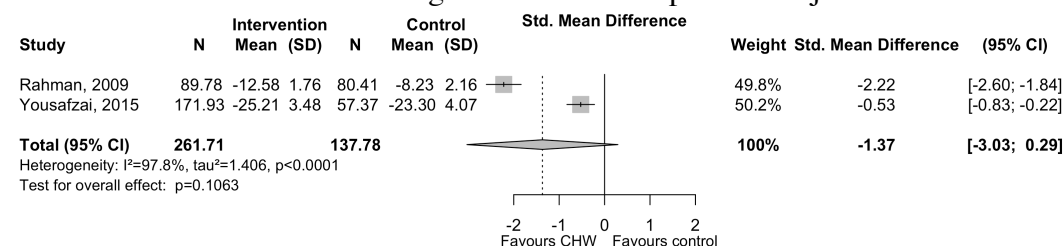
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.11. Initiation of breastfeeding within 1 h of birth – adjusted



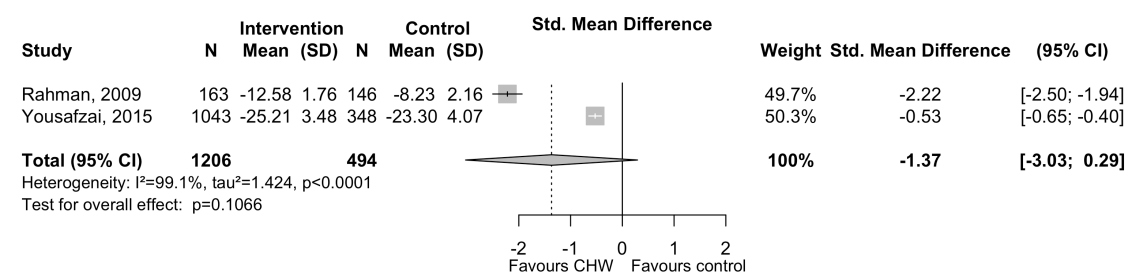
Comparison 3: CHW interventions to improve maternal and newborn health compared to usual care, Outcome: 3.11. Initiation of breastfeeding within 1 h of birth – unadjusted



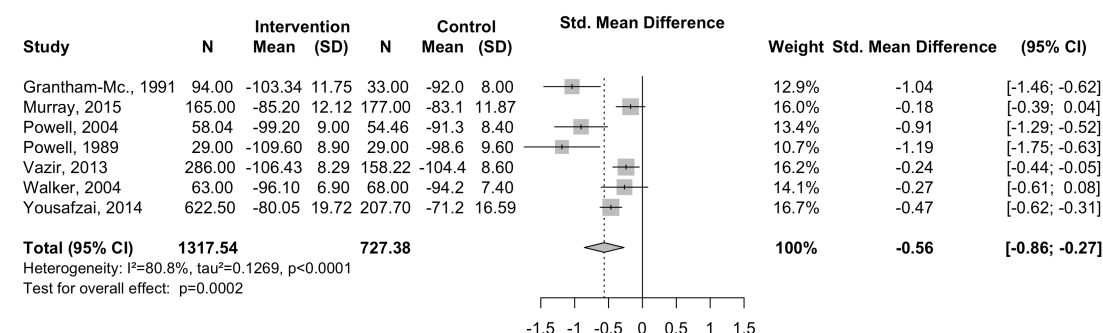
Comparison 4: CHW interventions to promote cognitive development in infants, Outcome: 4.1 Maternal knowledge of infant development – adjusted



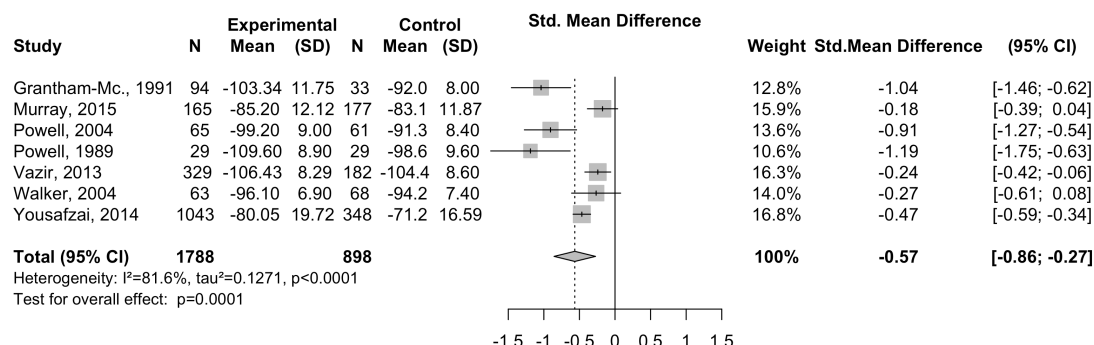
Comparison 4: CHW interventions to promote cognitive development in infants, Outcome: 4.1 Maternal knowledge of infant development – unadjusted



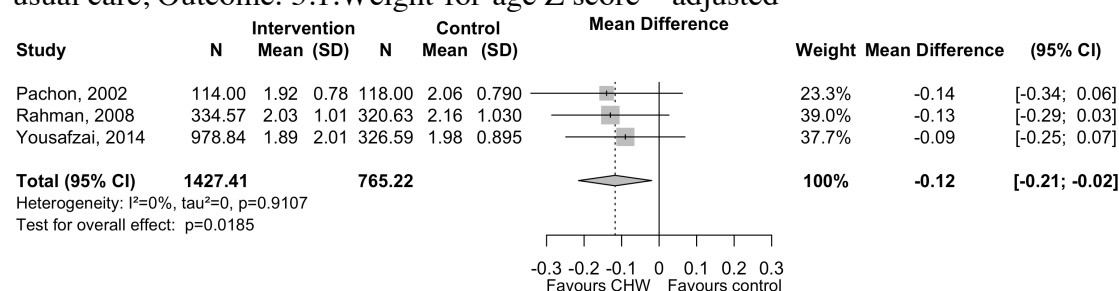
Comparison 4: CHW interventions to promote cognitive development in infants, Outcome: 4.2 Mental development – adjusted



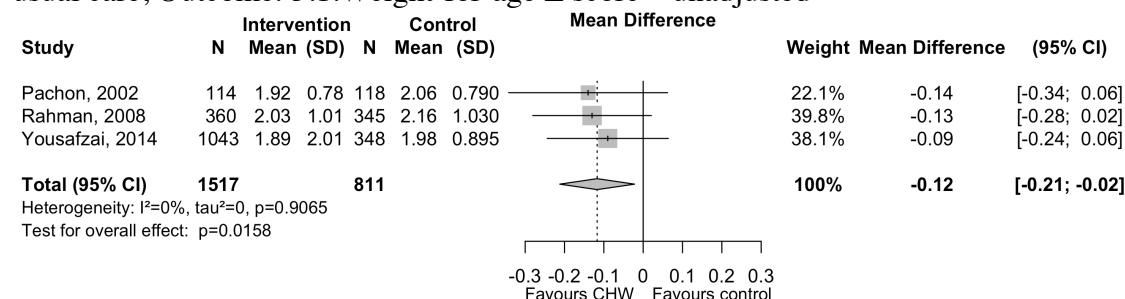
Comparison 4: CHW interventions to promote cognitive development in infants, Outcome: 4.2 Mental development – unadjusted



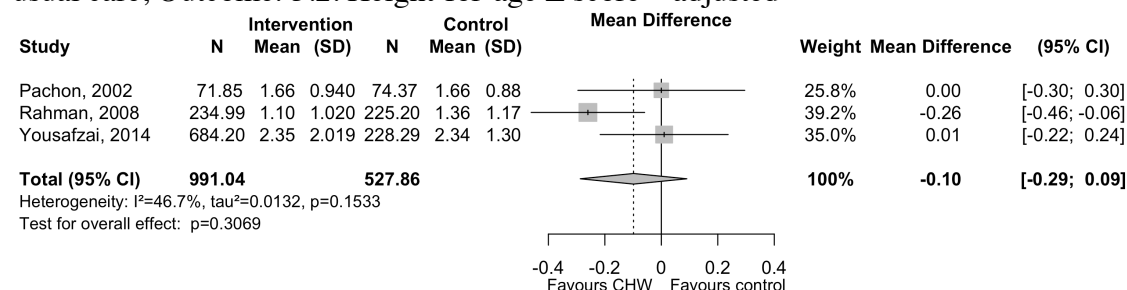
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.1.Weight-for-age Z score – adjusted



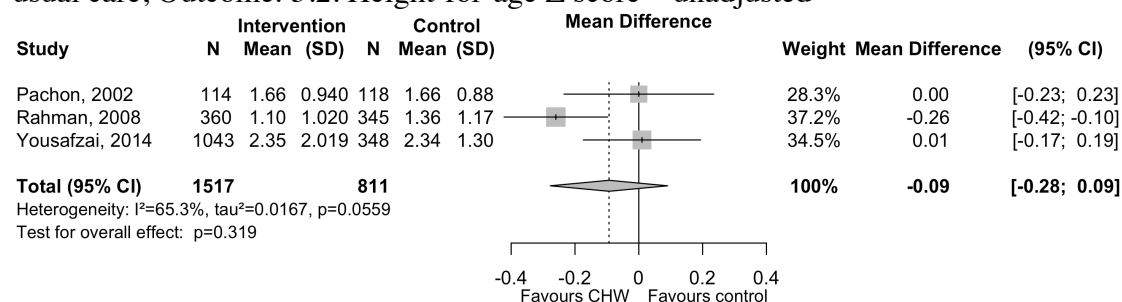
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.1.Weight-for-age Z score – unadjusted



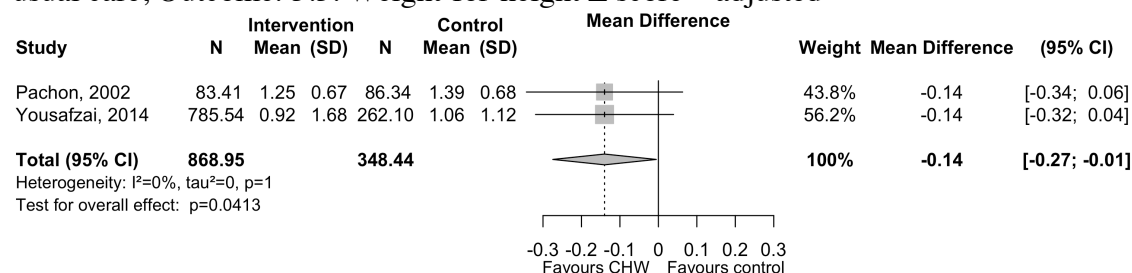
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.2. Height-for-age Z score – adjusted



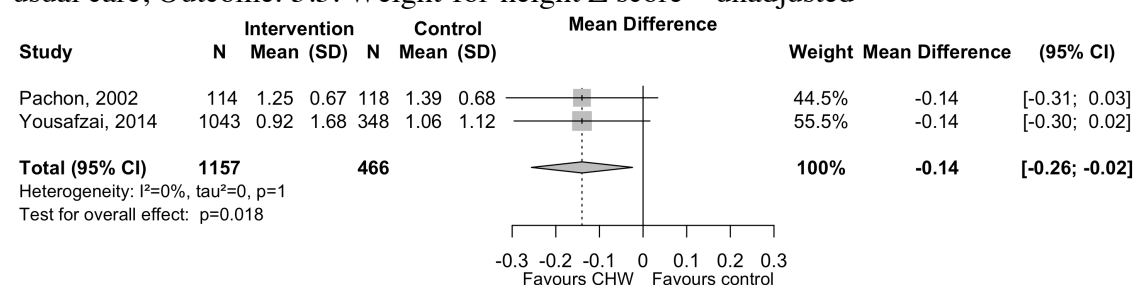
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.2. Height-for-age Z score – unadjusted



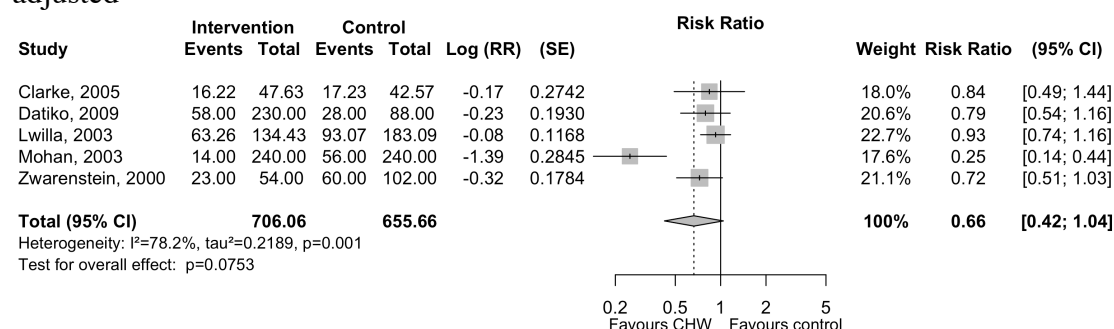
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.3. Weight-for-height Z score – adjusted



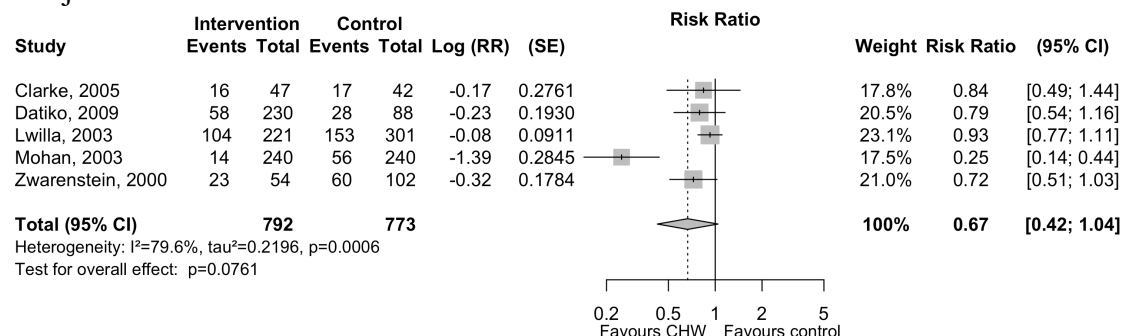
Comparison 5: CHW interventions to promote healthy feeding practices compared to usual care, Outcome: 5.3. Weight-for-height Z score – unadjusted



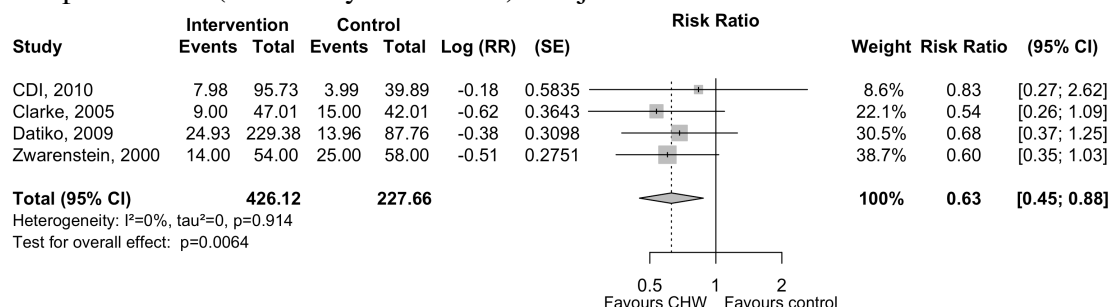
Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.1 TB Cure rate – adjusted



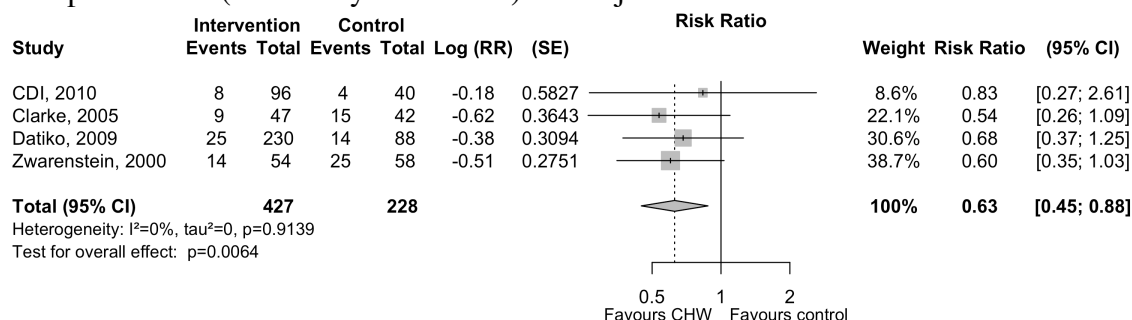
Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.1 TB Cure rate – unadjusted



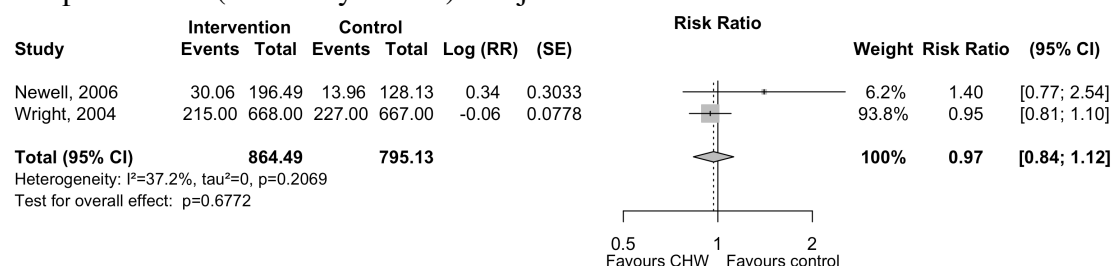
Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.2 TB Cure or completion rate (vs. facility-based care) – adjusted



Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.2 TB Cure or completion rate (vs. facility-based care) – unadjusted



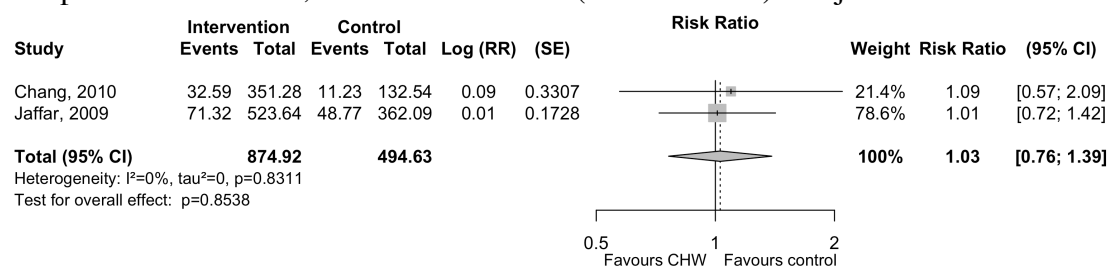
Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.3 TB Cure or completion rate (vs. family DOTS) – adjusted



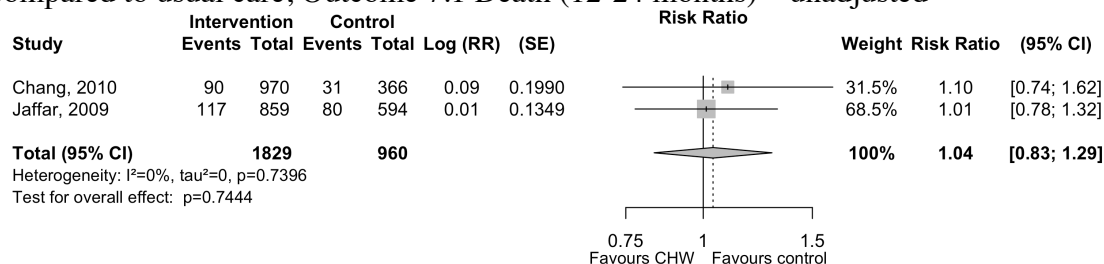
Comparison 6: CHW interventions to improve TB or pneumonia treatment outcomes compared with other forms of adherence support, Outcome: 6.3 TB Cure or completion rate (vs. family DOTS) – unadjusted



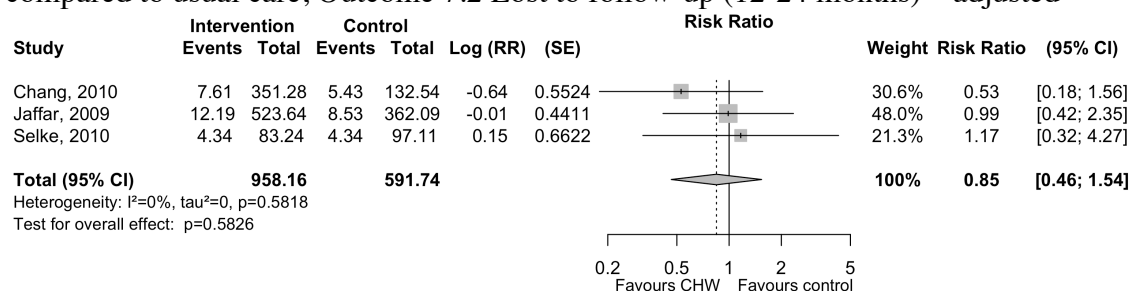
Comparison 7: CHW interventions for maintenance of antiretroviral therapy compared to usual care, Outcome 7.1 Death (12-24 months) – adjusted



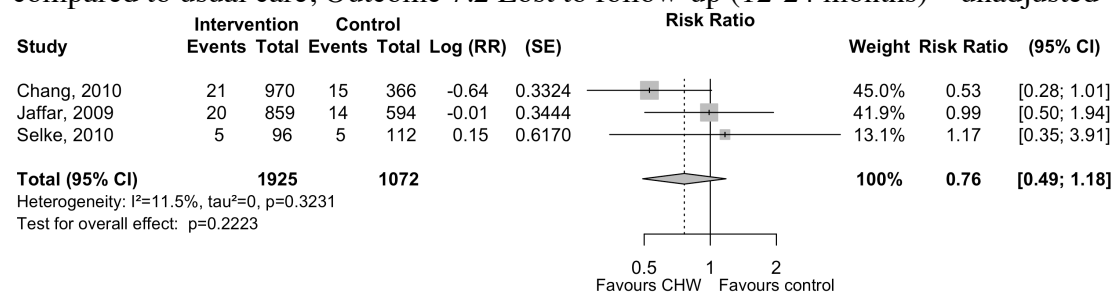
Comparison 7: CHW interventions for maintenance of antiretroviral therapy compared to usual care, Outcome 7.1 Death (12-24 months) – unadjusted



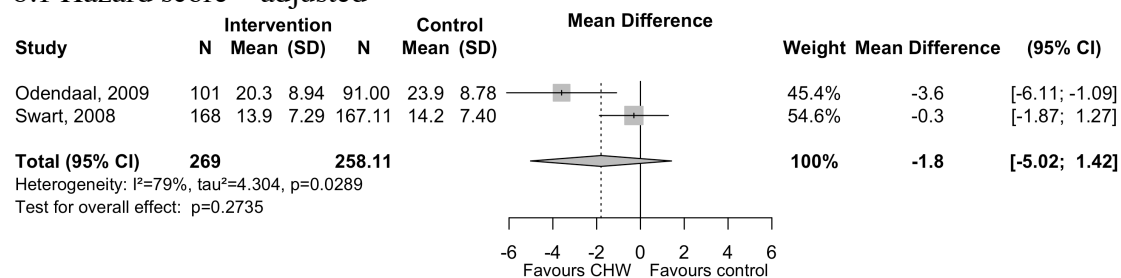
Comparison 7: CHW interventions for maintenance of antiretroviral therapy compared to usual care, Outcome 7.2 Lost to follow-up (12-24 months) – adjusted



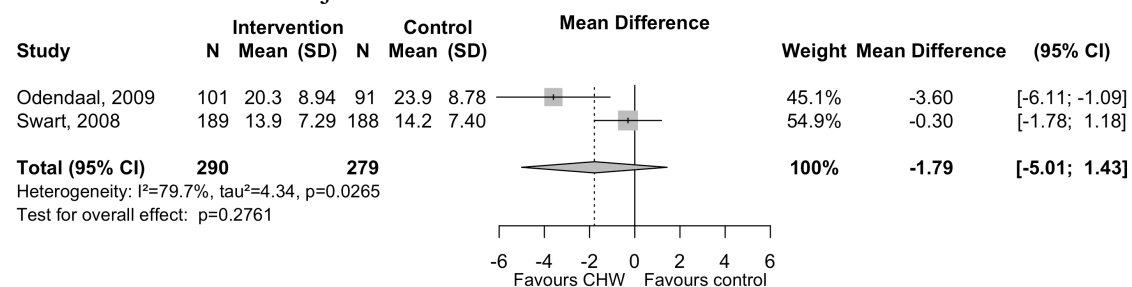
Comparison 7: CHW interventions for maintenance of antiretroviral therapy compared to usual care, Outcome 7.2 Lost to follow-up (12-24 months) – unadjusted



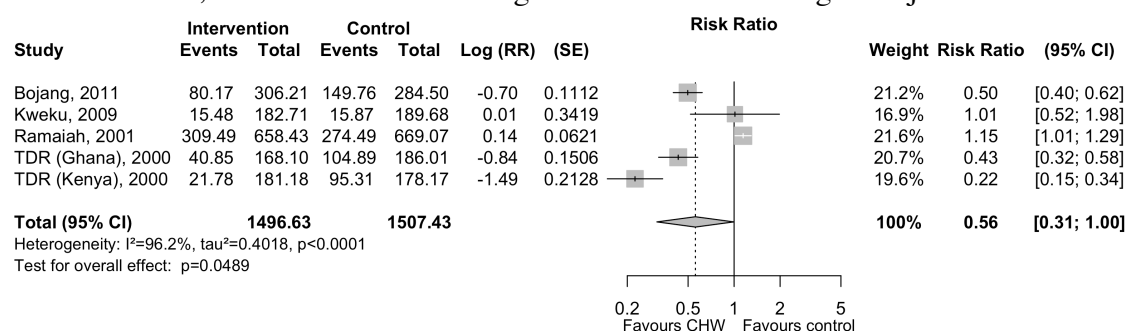
Comparison 8: CHW interventions to promote injury prevention in children, Outcome: 8.1 Hazard score – adjusted



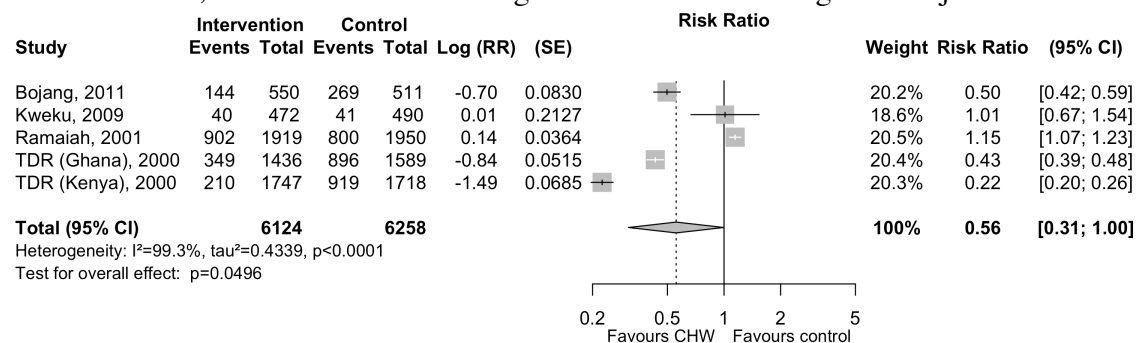
Comparison 8: CHW interventions to promote injury prevention in children, Outcome: 8.1 Hazard score – unadjusted



Comparison 9: CHW interventions to promote immunization uptake and mass drug administration, Outcome: 9.2 Mass drug administration coverage – adjusted



Comparison 9: CHW interventions to promote immunization uptake and mass drug administration, Outcome: 9.2 Mass drug administration coverage – unadjusted



Outcomes of trials not assigned to the above comparisons

Study and primary outcome measured	Intervention (N)	Control (N)	Measure of effect (95% CI)	P value	Author's conclusions
Aung 2014	CHWs supply ORS plus zinc via social franchising (Total N =3605)	Usual care (Total N =3605)			
• Per cent of cases of diarrhoea treated with ORS plus zinc	13.70%	1.80%		<0.001	The intervention increased optimal treatment of childhood diarrhoea
Dias 2008	CHW provide info on dementia, behaviour mgmt., a psychiatric assess. & psychotropic medication if needed (N=41)	Received only education & dementia information & placed on a waiting list to receive intervention in 6 months (N=40)			
• Caregiver mental health (GHQ score)	-1.12 [-2.07 to -0.17]	0.03 [-0.14 to 0.19]			Home based support for caregivers of persons with dementia leads to significant improvement

					in caregiver mental health
Fylkesnes 2013	CHWs offer HB-VCT (N=836)	Standard HIV testing services (N=858)	Measure of effect (95% CI)	P value	Author's conclusions
• Acceptance of counseling and testing	82%	52%	RR 1.7 [1.4-2.0]		At baseline the likelihood of being tested was higher for women vs. men and for more educated people. At follow-up these differences were found only in the control communities. Measured negative life events following HIV testing were similar in both groups. This HB- VCT model was found to be feasible, with a very high acceptance and to have important equity effects
• Equity in uptake (by educational attainment and sex)	Women (%): 86.1 Men (%): 76.1 years in school (%): 79.4 ≥ 10 years in school: 78.2	Women (%): 60.4 Men (%): 42.2 years in school: 27.2 ≥ 10 years in school: 56.2	Interaction between educational attainment & intervention arm was significant when excluding PMTCT-related tests	0.036	
• Negative life events following VCT	Break-up of marriage (%): 4.4 Break-up of sexual relationship (%): 7.7 Physical violence (%): 6.2 Neglected or disowned by family (%): 6.1 Loss of close family (%): 57.3	Break-up of marriage (%): 4.5 Break-up of sexual relationship (%): 6.4 Physical violence (%): 6.1 Neglected or disowned by family (%): 6.0 Loss of close family (%): 51.9	There were no significant differences between the arms in the proportions reporting negative life events	0.987 0.393 0.825 0.852 0.135	
Ohnmar 2012	CHWs provided RDT and ACT (N=268)	Usual care (N=287)	Measure of effect (95% CI)	P value	Author's conclusions
• Whether villagers had received blood test-based treatment for fever	Mean (SE): 14.9% (2.4%)	Mean (SE): 5.7% (1.7%)	OR 3.2, [1.5 to 6.7]	0.003	The results show that implementing volunteer programmes to improve the coverage of

• Interval between onset of fever and the blood test (days)	Mean (SE): 3.6 (0.3)	Mean (SE): 4.8 (1.3)		0.37	accurate and timely diagnosis with RDT and early treatment may be beneficial but the timeliness of detection and sustainability must be improved.
• Mortality (rate/100 000)	Mean (SE): 130 (37)	Mean (SE): 119 (34)	OR 1.09 (0.45 to 2.66)		
Sankaranarayanan 2005	CHW oral visual inspection (N=96517)	Standard care (N=95356)	Measure of effect (95% CI)	P value	Author's conclusions
• Oral cancer deaths	77	87	Mortality rate ratio All subjects: 0.79 (0.51 to 1.22) Tobacco or alcohol users: 0.66 [0.45–0.95]		Oral visual screening can reduce mortality in high-risk individuals
Siabani 2015	CHWs provide home-based education for patients with chronic heart failure (N=76)	A: Education programme at hospital (N = 78) B: Usual care (N = 77)	Measure of effect (95% CI)	P value	Author's conclusions
• Self-care scores (incl. maint., mgmt., and confidence, max 100)	Mean (SD) Maintenance: 26.2±12.7 Management: 29.4±11 Confidence: 9.5±17	For the health professional group were 29.5±12, 31.3±12, 18.1±17; and for the control group were 2.7±9, 10.2±10, –0.30±11 respectively	t-test	vs. control A: 0.12 0.41 0.004 vs. control B: <0.001 <0.001 <0.001	The home-based face-to-face education by CHVs improved self-care maintenance and self-care management in patients with CHF as effectively as the education provided by health professionals in a formal health education program, and much better

than the usual
care.

Appendix 4e: PRISMA Checklist

SECTION/TOPIC	#	CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	102
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	N/A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	102-103
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	103
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	104- 110
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	105
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	106-110
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	106
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic	104

		review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	110-120
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix 4b
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	110-120
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	110-120
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	110-120
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	110-120
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	110-120
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	121-123
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	131-199
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	124-131
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	131-199
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	131-199
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	124

Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	131-199
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	199-206
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	216-218
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	220-226
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	N/A

Appendices to Chapter 5

Appendix 5a: Characteristics of Studies

Characteristics of CHW trials are provided in Appendix 4c and Appendix 6d.
Characteristics of influential papers are provided here.

Author	Year	Objective
Berman	1987	Review experiences of six large-scale community-based worker programs
Bhandari	2003	Assess the feasibility, effectiveness, and safety of an educational intervention to promote exclusive breastfeeding in India
Bhutta	2005	Review community-based antenatal, intrapartum, and postnatal intervention trials in developing countries aimed to identify (1) key behaviours and interventions for which the weight of evidence is sufficient to recommend their inclusion in community-based neonatal care programs and (2) key gaps in knowledge and priority areas for future research and programme learning
Bhutta	2008	Assess the feasibility of delivering a package of community-based interventions for improving perinatal care using lady

		health workers and traditional birth attendants in rural Pakistan
Bhutta	2010	Identify CHWs programs with positive impact on Millennium Development Goals (MDGs) related to health or otherwise via a global systematic review of such interventions, as well as eight in-depth country case studies in Sub-Saharan Africa (Ethiopia Mozambique and Uganda), South East Asia (Bangladesh, Pakistan and Thailand) and Latin America (Brazil and Haiti)
Bhutta	2011	Evaluate the effectiveness of a community-based intervention package, principally delivered through CHWs working with traditional birth attendants and community health committees, for reduction of perinatal and neonatal mortality,
Callaghan	2010	Conduct a systematic review of task-shifting for HIV treatment and care in Africa
Chopra	2008	Identify all available policy options to address human resources for health in LMICs, and to assess the effectiveness of these policy options
Chowdhury	1997	Assess a CHW programme for the control of tuberculosis in Bangladesh
Corbett	2010	Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Zimbabwe
Darmstadt	2005	Identify 16 interventions with proven efficacy (implementation under ideal conditions) for neonatal survival and combine them into packages for scaling up in health systems, according to three service delivery modes (outreach, family-community, and facility-based clinical care)
Darmstadt	2009	Conduct a systematic review of evidence for the effect of community-based cadres—community-based skilled birth attendants (SBAs), trained traditional birth attendants (TBAs), and CHWs—in improving perinatal and intrapartum-related outcomes.
Douthwaite	2005	Present findings from the first national evaluation of Pakistan's Lady Health Worker Programme, initiated under the Ministry of Health
Farmer	2001a	Re-examine arguments against the use of HAART in developing countries and describe a small community-based treatment programme based on lessons gained in TB control
Farmer	2001b	Discuss objections to the widespread use of HAART, and suggest that directly-observed therapy of chronic infectious disease with multidrug regimens can be highly effective in

		settings of great privation as long as there is sustained commitment to uninterrupted care that is free to the patient
Floyd	1997	Conduct an economic evaluation of directly observed treatment (DOT) and conventionally delivered treatment for the management of new cases of tuberculosis in adults.
Fulton	2011	Use an economics perspective to review the skill mix literature to determine its strength of the evidence, identify gaps in the evidence, and to propose a research agenda.
Glenton	2010	Explore the views of stakeholders who have participated in the design and implementation of the Female Community Health Volunteer regarding Volunteer motivation and appropriate incentives, and to compare these views with the views and expectations of Volunteers.
Haines	2007	Assess the potential contribution of CHWs to achieving child survival goals
Harvey	2008	Assess (i) whether Zambian CHWs could prepare and interpret RDTs accurately and safely using manufacturer's instructions alone; (ii) whether simple, mostly pictorial instructions (a "job aid") could raise performance to adequate levels; and (iii) whether a brief training programme would produce further improvement.
Hermann	2009	Assess whether present community health worker programmes for antiretroviral treatment are taking into account the lessons learnt from past experiences with community health worker programmes in primary health care and to what extent they are seizing the new antiretroviral treatment-specific opportunities.
Kakuma	2011	Review the current state of human resources for mental health, needs, and strategies for action.
Kallander	2013	Conduct a thematic review of how mHealth projects have approached the intersection of cellular technology and public health in low- and middle-income countries and identify the promising practices and experiences learned, as well as novel and innovative approaches of how mHealth can support community health workers.
Khan	2002	Establish the costs and effectiveness of different strategies for implementing directly observed treatment (DOT) for tuberculosis.
Koenig	2004	Describe scale-up a successful HIV/AIDS treatment project in rural Haiti
Lawn	2008	Revisit the 1978 Alma-Ata Declaration, examine 30 years of changes in global policy to identify the lessons learned that are of relevance today, particularly for accelerated scale-up of

		primary health-care services necessary to achieve the Millennium Development Goals
Lehmann	2007	Review questions regarding the feasibility and effectiveness of community health worker programmes
Lehmann	2009	Consider the promise of task shifting for the human resource crisis in Africa
Lewin	2005	Assess the effects of CHW interventions in primary and community health care
Lewin	2010	Assess the effects of CHW interventions in primary and community health care on maternal and child health and the management of infectious diseases.
Magnussen	2004	Consider lessons from primary health care for global health policy
Mitnick	2003	Evaluate the results of community-based therapy for multidrug-resistant tuberculosis in a poor section of Lima, Peru
Morely	1983	Present a wide range of case studies depicting typical problems that arise in the implementation of primary health care - and how these can be overcome
Morris	2009	Assess a comprehensive task-shifting programme among existing health providers and community-based workers in Zambia
Mukherjee	2006	Describes the interventions to promote access to care and adherence to ART in rural Haiti
Mullan	2002	Provide a global overview of community-oriented primary care, tracing its conceptual roots, reviewing its manifestations, and exploring its future prospects as an organizational paradigm for the democratic organization of community health services.
Narasimhan	2004	Consider strategies for responding to the global human resources crisis
Noyes	2007	Report findings from a qualitative meta-synthesis concerning people with, or at risk of, tuberculosis, service providers and policymakers and their experiences and perceptions of tuberculosis and treatment.
Pagnoni	1997	Describe a community-based programme to ensure prompt and adequate treatment of presumptive episodes of clinical malaria in children has been established in a rural province of Burkina Faso.

Pandey	1991	Report a controlled intervention trial of indigenous community health workers to detect and treat pneumonia in Nepal
Patel	2010	Test the effectiveness of an intervention led by lay health counsellors in primary care settings to improve outcomes of people with depression and anxiety disorders.
Perez	2008	Consider the role of CHWs as social justice and policy advocates for community health and well-being
Philips	2008	Consider the potential of task shifting for antiretroviral treatment delivery in sub-Saharan Africa
Rowe	2005	Provide an overview of issues and evidence about the determinants of health worker performance and strategies for improving it.
Sauerborn	1989	Study the utilization of CHWs in relation to other sources of health care
Schneider	2008	Examine the case of South Africa, where there has been rapid growth of a range of lay workers (home-based carers, lay counsellors, DOT supporters etc.) principally in response to an expansion in budgets and programmes for HIV.
Selke	2010	Assess whether community-based care delivered by people living with HIV/AIDS (PLWAs) could replace clinic-based HIV care.
Standing	2008	Considers how poor populations can obtain access to trusted, competent knowledge and services in increasingly pluralistic health systems where unregulated markets for health knowledge and services dominate, using the example of CHWs
Stanton	1987	Assess an educational intervention was designed to improve three water-sanitation behaviours empirically shown to be associated with high rates of childhood diarrhoea in Dhaka, Bangladesh
Thiam	2007	Assess the effectiveness of a contextualized intervention strategy aimed at improving patients' adherence to treatment and to evaluate its impact on TB control in a resource-poor country in Africa with prevalent TB infection.
Van Damme	2008	Consider constraints to and strategies for scaling-up antiretroviral treatment in Southern African countries with human resource shortage
Victora	2004	Consider how to achieve universal coverage with health interventions
Volmink	1997	Determine the effectiveness of strategies to promote adherence

		to treatment for tuberculosis.
Volmink	2007	To compare DOT with self-administration of treatment or different DOT options for people requiring treatment for clinically active tuberculosis or prevention of active disease.
Walley	2001	Assess the effectiveness of different packages for tuberculosis treatment under operational conditions in Pakistan.
Walt	1990	Assess what has happened to CHWs in the national programmes that expanded so rapidly throughout the 1970s
Werner	1981	Considers whether CHWs should act as agents of political reform
Winch	2005	Categorize and describe Intervention Models involving CHWs that aim to improve case management of sick children at the household and community levels.
Zachariah	2009	Highlights the main opportunities and challenges posed by task shifting and proposes specific actions to tackle the challenges

Appendix 5b: Taxonomy

The taxonomy can be downloaded in FileMaker Pro format here:

https://github.com/madeleine25/DPhil_Chapter_6.git

Appendix 5c: Citations for illustrative examples

Item	Excerpt	Citation
3	“BRAC [decided]...	Standing, H., & Chowdhury, A. M. (2008). Producing effective knowledge agents in a pluralistic environment: what future for community health workers? <i>Social Science & Medicine</i> , 66(10), 2096-2107.
4	“mother coordinators [were]...”	Kidane, G., & Morrow, R. H. (2000). Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial. <i>Lancet</i> , 356(9229), 550-555.
5	“[CHWs] were expected to work...”	le Roux, I. M., le Roux, K., Comulada, W. S., Greco, E. M., Desmond, K. A., Mbewu, N., & Rotheram-Borus, M. J. (2010). Home visits by neighbourhood Mentor Mothers provide timely recovery from childhood malnutrition in South Africa: results from a randomized controlled trial.

		Nutrition Journal, 9, 56.
6	“The intervention was...”	Cooper, P. J., Tomlinson, M., Swartz, L., Landman, M., Molteno, C., Stein, A., . . . Murray, L. (2009). Improving quality of mother-infant relationship and infant attachment in socioeconomically deprived community in South Africa: randomised controlled trial. <i>BMJ</i> , 338, b974.
9	“[CHWs] are employees...”	Stanton, C. K., Newton, S., Mullany, L. C., Cofie, P., Tawiah Agyemang, C., Adiibokah, E., . . . Owusu-Agyei, S. (2013). Effect on Postpartum Hemorrhage of Prophylactic Oxytocin (10 IU) by Injection by Community Health Officers in Ghana: A Community-Based, Cluster-Randomized Trial. <i>Plos Medicine</i> , 10(10). doi: http://dx.doi.org/10.1371/journal.pmed.1001524
10	“The interaction between...”	Jafar, T. H., Hatcher, J., Poulter, N., Islam, M., Hashmi, S., Qadri, Z., . . . Hypertension Research, G. (2009). Community-based interventions to promote blood pressure control in a developing country: a cluster randomized trial. <i>Annals of Internal Medicine</i> , 151(9), 593-601.
11	“active involvement of...”	Haines, A., Sanders, D., Lehmann, U., Rowe, A. K., Lawn, J. E., Jan, S., . . . Bhutta, Z. (2007). Achieving child survival goals: potential contribution of community health workers. <i>Lancet</i> , 369(9579), 2121-2131. doi: 10.1016/S0140-6736(07)60325-0
12	“Women were informed...”	Ali, B. S., Rahbar, M., Naeem, S., Gul, A., Mubeen, S., & Iqbal, A. (2003). The effectiveness of counselling on anxiety and depression by minimally trained counsellors: a randomised controlled trial. <i>American Journal of Psychotherapy</i> , 57(3).
13	“one female community health...”	Baqui, A. H., El-Arifeen, S., Darmstadt, G. L., Ahmed, S., Williams, E. K., Seraji, H. R., . . . Projahnmo Study, G. (2008). Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. <i>Lancet</i> , 371(9628), 1936-1944.
14	“Formal tests and ...”	Tylleskar, T., Jackson, D., Meda, N., Engebretsen, I. M., Chopra, M., Diallo, A. H., . . . Group, P.-E. S. (2011). Exclusive breastfeeding promotion by peer counsellors in sub-Saharan Africa (PROMISE-EBF): a cluster-randomised trial. <i>Lancet</i> , 378(9789), 420-427. doi: 10.1016/S0140-6736(11)60738-1
15	“[CHWs] were nominated...”	Chang, L. W., Kagaayi, J., Nakigozi, G., Ssempijja, V., Packer, A. H., Serwadda, D., . . . Reynolds, S. J. (2010). Effect of peer health workers on AIDS care in Rakai, Uganda: a cluster-randomized trial. <i>PLoS ONE [Electronic</i>

		Resource], 5(6), e10923.
16	“the minimum...”	Dias, A., Dewey, M. E., D'Souza, J., Dhume, R., Motghare, D. D., Shaji, K. S., . . . Patel, V. (2008). The effectiveness of a home care program for supporting caregivers of persons with dementia in developing countries: a randomised controlled trial from Goa, India. <i>PLoS ONE</i> , 3(6), e2333. doi: 10.1371/journal.pone.0002333
17	“All study CHWs had previously...”	Yeboah-Antwi, K., Pilingana, P., Macleod, W. B., Semrau, K., Siazele, K., Kalesha, P., . . . Hamer, D. H. (2010). Community case management of fever due to malaria and pneumonia in children under five in Zambia: a cluster randomized controlled trial. <i>PLoS Medicine / Public Library of Science</i> , 7(9), e1000340.
18	“[CHWs] received a...”	Kumar, V., Mohanty, S., Kumar, A., Misra, R. P., & Santosham, M. (2008). Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. <i>Lancet</i> , 372(9644), 1151-1162. doi: http://dx.doi.org/10.1016/S0140-6736(08)61483-X
19	“Training on basic HIV...”	Chang, L. W., Kagaayi, J., Nakigozi, G., Ssempijja, V., Packer, A. H., Serwadda, D., . . . Reynolds, S. J. (2010). Effect of peer health workers on AIDS care in Rakai, Uganda: a cluster-randomized trial. <i>PLoS ONE [Electronic Resource]</i> , 5(6), e10923.
21	“we integrated a...”	Rahman, A., Malik, A., Sikander, S., Roberts, C., & Creed, F. (2008). Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. <i>Lancet</i> , 372(9642), 902-909.
22	“received refresher training...”	Tian, M., Ajay, V., Dunzhu, D., Hameed, S., Li, X., Liu, Z., . . . Yan, L. (2015). A Cluster-Randomized Controlled Trial of a Simplified Multifaceted Management Program for Individuals at High Cardiovascular Risk (SimCard Trial) in Rural Tibet, China, and Haryana, India. <i>Circulation</i> .
23	“an assistant quizzed...”	Harvey, S. A., Jennings, L., Chinyama, M., Masaninga, F., Mulholland, K., & Bell, D. R. (2008). Improving community health worker use of malaria rapid diagnostic tests in Zambia: package instructions, job aid and job aid-plus-training. <i>Malaria Journal</i> , 7. doi: 10.1186/1475-2875-7-160
24	“drugs...must [typically]...”	Bojang, K. A., Akor, F., Conteh, L., Webb, E., Bittaye, O., Conway, D. J., . . . Greenwood, B. (2011). Two strategies for the delivery of IPTc in an area of seasonal malaria

		transmission in the Gambia: a randomised controlled trial. PLoS Medicine / Public Library of Science, 8(2), e1000409.
25	"[CHWs] frequently..."	Bojang, K. A., Akor, F., Conteh, L., Webb, E., Bittaye, O., Conway, D. J., . . . Greenwood, B. (2011). Two strategies for the delivery of IPTc in an area of seasonal malaria transmission in the Gambia: a randomised controlled trial. PLoS Medicine / Public Library of Science, 8(2), e1000409.
27	"during home visits..."	Kumar, V., Mohanty, S., Kumar, A., Misra, R. P., & Santosham, M. (2008). Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. <i>Lancet</i> , 372(9644), 1151-1162. doi: http://dx.doi.org/10.1016/S0140-6736(08)61483-X
28	"World Vision Uganda..."	Bolton, P., Bass, J., Betancourt, T., Speelman, L., Onyango, G., Clougherty, K., . . . Verdeli, H. (2007). Interventions for Depression Symptoms Among Adolescent Survivors of War and Displacement in Northern Uganda. <i>Jama</i> , 298(5).
29	"supervisors... visited..."	Lewycka, S., Mwansambo, C., Rosato, M., Kazembe, P., Phiri, T., Mganga, A., . . . Costello, A. (2013). Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. <i>Lancet</i> , 381(9879), 1721-1735.
30	"Peer support, through..."	Haines, A., Sanders, D., Lehmann, U., Rowe, A. K., Lawn, J. E., Jan, S., . . . Bhutta, Z. (2007). Achieving child survival goals: potential contribution of community health workers. <i>Lancet</i> , 369(9579), 2121-2131. doi: 10.1016/S0140-6736(07)60325-0
31	"Performance assessment..."	Kumar, V., Mohanty, S., Kumar, A., Misra, R. P., & Santosham, M. (2008). Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. <i>Lancet</i> , 372(9644), 1151-1162. doi: http://dx.doi.org/10.1016/S0140-6736(08)61483-X
32	"Each supervising mother..."	Leite, A. J., Puccini, R. F., Atalah, A. N., Alves Da Cunha, A. L., & Machado, M. T. (2005). Effectiveness of home-based peer counselling to promote breastfeeding in the northeast of Brazil: a randomized clinical trial. <i>Acta paediatrica</i> , 94(6), 741-746.
32	"Provided	Yeboah-Antwi, K., Pilingana, P., Macleod, W. B., Semrau, K., Siazeele, K., Kalesha, P., . . . Hamer, D. H. (2010).

	occasional..."	Community case management of fever due to malaria and pneumonia in children under five in Zambia: a cluster randomized controlled trial. PLoS Medicine / Public Library of Science, 7(9), e1000340.
32	"A monthly salary..."	Soofi, S., Ahmed, S., Fox, M. P., MacLeod, W. B., Thea, D. M., Qazi, S. A., & Bhutta, Z. A. (2012). Effectiveness of community case management of severe pneumonia with oral amoxicillin in children aged 2-59 months in Matiari district, rural Pakistan: a cluster-randomised controlled trial. <i>Lancet</i> , 379(9817), 729-737.
32	"Peer counsellors were..."	Waiswa, P., Pariyo, G., Kallander, K., Akuze, J., Namazzi, G., Ekirapa-Kiracho, E., . . . Uganda Newborn Study, T. (2015). Effect of the Uganda Newborn Study on care-seeking and care practices: a cluster-randomised controlled trial. <i>Glob Health Action</i> , 8, 24584. doi: 10.3402/gha.v8.24584
32	"No financial or other..."	Jaffar, S., Amuron, B., Foster, S., Birungi, J., Levin, J., Namara, G., . . . Jinja Trial, T. (2009). Rates of virological failure in patients treated in a home-based versus a facility-based HIV-care model in Jinja, southeast Uganda: a cluster-randomised equivalence trial. <i>Lancet</i> , 374(9707), 2080-2089. doi: 10.1016/S0140-6736(09)61674-3
33	"Supported by their..."	Bojang, K. A., Akor, F., Conteh, L., Webb, E., Bittaye, O., Conway, D. J., . . . Greenwood, B. (2011). Two strategies for the delivery of IPTc in an area of seasonal malaria transmission in the Gambia: a randomised controlled trial. <i>PLoS Medicine / Public Library of Science</i> , 8(2), e1000409.
34	"the equivalent of..."	Leite, A. J., Puccini, R. F., Atalah, A. N., Alves Da Cunha, A. L., & Machado, M. T. (2005). Effectiveness of home-based peer counselling to promote breastfeeding in the northeast of Brazil: a randomized clinical trial. <i>Acta paediatrica</i> , 94(6), 741-746.
34	"All CHWs were..."	Tomlinson, M., Doherty, T., Ijumba, P., Jackson, D., Lawn, J., Persson, L. A., . . . Chopra, M. (2014). Goodstart: a cluster randomised effectiveness trial of an integrated, community-based package for maternal and newborn care, with prevention of mother-to-child transmission of HIV in a South African township. <i>Tropical Medicine & International Health</i> , 19(3), 256-266.

Appendices to Chapter 6

Appendix 6a: Search Strategy

AMED, Ovid

AMED (1985 to April May 2014) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 53 hits. The search was updated May 11, 2016 and produced 4 hits.

BiblioMap: EPPI-Centre database of health promotion research

The search strategy used for BiblioMap was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. Nutrition worker
12. Frontline health worker
13. Barangay health worker
14. Basic health worker
15. Auxiliary health worker
16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier
20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika

34. shasto karmis
35. anganwadi
36. barefoot doctor
37. agente comunitario de salud
38. agente communitario de saude
39. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR
22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31
OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

British Library for Development Studies at IDS

The search strategy used for IDS was as follows: “community health workers”

The search was conducted on June 30, 2014 and produced 5 hits. The search was updated May 11, 2016 and produced 4 hits.

CINAHL, EBSCO

CINAHL, Cumulative Index to Nursing and Allied Health (1960 to May 2014) was searched via Ebsco using the following terms:

(TI ((community OR village) N2 (agent* OR aid* OR promot* OR mobili?er* OR distribut?r* OR worker)) OR TI ((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) N2 (worker* OR volunteer*)) OR TI ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR monitora* OR promotora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude"))) AND AB ((performance or effective* or skill*))

The search was conducted on May 4, 2014 and produced 189 hits. The search was updated May 11, 2016 and produced 31 hits.

CENTRAL

The search strategy used for the Cochrane Central Register of Controlled Trials (Issue 5 of 12, May 2014) was as follows:

1. ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)):ti,ab
3. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente comunitario de saude"):ti,ab
4. (performance or effective* or skill*):ti,ab
5. #1 or #2 or #3
6. #4 and #5

(In: Trials)

The search was conducted on May 4, 2014 and produced 256 hits. The search was updated May 11, 2016 and produced 129 hits.

Database of Promoting Health Effectiveness Reviews (DoPHER) search strategy

The search strategy used for DoPHER was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. nutrition worker
12. frontline health worker
13. Barangay health worker
14. basic health worker
15. Auxiliary health worker
16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier

20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika
34. shasto karmis
35. anganwadi
36. barefoot doctor
37. agente comunitario de salud
38. agente comunitario de saude
39. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR
22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31
OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

EMBASE search strategy

EMBASE (1974 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2526 hits.

The search was updated May 11, 2016 and produced 796 hits.

Global Health search strategy

Global Health (1973 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 1341 hits. The search was updated May 11, 2016 and produced 450 hits.

IDEAS database of unpublished working papers

The search strategy used for the IDEAS database was as follows:

"community health workers" performance
Match: all; Words forms: exact; In: abstract

The search was conducted on June 30, 2014 and produced 4 hits. The search was updated May 11, 2016 and produced 4 hits.

Impact Evaluation Database (3ie)

Keyword: community health worker

Evaluation method: Randomised Control Trials (RCT)

The search was conducted on July 12, 2016 and produced 62 hits.

IPA (Innovations for Poverty Action) Database

Keyword: community

Program area: Health

The search was conducted on July 12, 2016 and produced 6 hits.

Databases within ISI Web of Science

Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index - Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities) (1945-May 2014) were searched using the following strategy:

TITLE: (((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*))) OR TITLE: (((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) NEAR/2 (worker* OR volunteer*))) OR TITLE: ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente comunitario de saude")) AND TITLE: ((performance or effective* or skill*))

Timespan: All years. Indexes: SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH,

The search was conducted on May 4, 2014 and produced 3047 hits. The search was updated May 11, 2016 and produced 657 hits.

J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database

Keyword: Community worker

Sector: Health

The search was conducted on July 12, 2016 and produced 23 hits.

MEDLINE search strategy

MEDLINE (1946 to May 2014 Week 19) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2143 hits. The search was updated May 11, 2016 and produced 405 hits.

MEDLINE In-Process In-Process & Other Non-Indexed Citations search strategy

MEDLINE In-Process (1946 to May 2014 Week 19) was searched via OVID using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 251 hits. The search was updated May 11, 2016 and produced 734 hits.

POPLINE, K4Health

The search strategy used for POPLINE (1965-May 2014) was as follows:

All fields: effective* OR performance OR skill AND

Title: ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)) OR ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)) OR (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente comunitario de saude")

The search was conducted on May 4, 2014 and produced 20 hits. The search was updated May 11, 2016 and produced 168 hits.

Pubmed

The search strategy used for Pubmed (1960-May 2014) was as follows:

(((((("Community health agent" or "Community Health Aides" or "Community health promoter" or "Community mobilizer" or "Community drug distributor" or "community health worker" or "Village health worker"[Title/Abstract])) OR ("Rural Health Worker" or "Lay Health Worker" or "Lady health worker" or "nutrition worker" or "frontline health worker" or "Barangay health worker" or "basic health worker" or "Auxiliary health worker" or "health extension worker" or "community health volunteer" or "village health volunteer"[Title/Abstract])) OR (accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente comunitario de saude"[Title/Abstract])) AND (performance or effective* or skill*[Title/Abstract]))NOT MEDLINE[sb])

The search was conducted on May 4, 2014 and produced 195 hits. The search was updated May 11, 2016 and produced 233 hits.

WHOLIS

WHOLIS (1948-May 2014) was searched via the Global Health Library the following terms: "community health workers" and performance.

The search was conducted on June 30, 2014 and produced 3 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank's JOLIS

JOLIS was searched with the following terms:

"community health workers" and performance

The search was conducted on June 30, 2014 and produced 6 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank Independent Evaluation Group

Advanced Search>>[all items]>>Systematic Reviews and Impact Evaluations>>Health, Nutrition and Population

The search was conducted on July 12, 2016 and produced 7 hits.

List of individuals and organizations contacted

1. Dr. Oladele Akogun, Federal University of Technology, Yola, Nigeria
2. Fekadu Ayele, Department of Community Health, Addis Ababa University, Ethiopia*
3. Dr. J. M. Carrón, Centro Paraguayo de Estudios de Población[†]
4. Dr. Larry Chang, Johns Hopkins School of Medicine
5. Dr. John Elder, Graduate School of Public Health, San Diego State University
6. Dr. Karin Källander, Malaria Consortium[†]
7. Dr. Charles Larson, School of Population and Public Health, University of British Columbia
8. Dr. Simon Lewin, Norwegian Knowledge Centre for the Health Services[†]
9. Dr. Khalid Omer, Community Information Empowerment and Training (CIET), Pakistan[†]
10. Dr. Nathan Miller, Bloomberg School of Public Health, Johns Hopkins University
11. Dr. Henry B. Perry, Bloomberg School of Public Health, Johns Hopkins University
12. Dr. Samantha Rowe, Center for Global Health, U.S. Centers for Disease Control and Prevention
13. Kuntal Kumar Saha, icddr,b
14. Associate Professor Pieter Serneels, University of East Anglia
15. Dr. Mitchell G. Weiss, Universität Basel
16. Dr. Peter Winch, Bloomberg School of Public Health, Johns Hopkins University[†]

[†]No response received.

*Deceased.

Appendix 6b: Data Extraction Form

STUDY ID (Author name, Year):	
--	--

VARIABLE	VALUE/ JUDGE- MENT	SOURCE	OPTIONS	VARIABLE TYPE
• INCLUSION CRITERIA				
<i>Types of Studies</i>				
Study design			• RCT • C-RCT • Factorial • ITS • Cohort • Case Control • CBA • Other (specify)	Nominal
<i>Types of Participants</i>				
Live and work in the catchment they serve.			• Yes • No • Unclear	Nominal
Are primarily based in said community			• Yes • No • Unclear	Nominal
Belong to of the formal health system in a low-income country			• Yes • No • Unclear	Nominal
Perform interventions related to healthcare delivery.			• Yes • No • Unclear	Nominal
Have received organized training in said interventions but have no formal or paraprofessional certification or tertiary education degree			• Yes • No • Unclear	Nominal
<i>Types of Interventions</i>				

Primary objective to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs			• Yes • No • Unclear	Nominal
Primary objective to examine the efficacy of a particular CHW activity as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity			• Yes • No • Unclear	Nominal
Type of intervention			• Recruitment • Training • Supervision • Incentives • Other (describe)	Nominal
<i>Types of Outcomes</i>				
List				Narrative
<i>Decision</i>				
Include or exclude			• Include • Exclude	Dichotomous
Reason for exclusion:				Narrative
•METHODS				
Study duration			• # months	Continuous
Date of study			• MM/YYYY-MM/YYYY	
Unit of allocation			• Patient • Provider • Community • Other (specify) • Unclear	Nominal
•PARTICIPANTS				
Total number			• # intervention, # control	Continuous

Setting			• Rural • Urban • Mixed (specify)	Nominal
Country				Nominal
Local name for CHWs				Nominal
Definition of CHW				Narrative
Age			• Mean (years)	Continuous
Sex			• % Male and % Female	Continuous
Education level			• Some primary • Primary • Some secondary • Secondary • Tertiary	Ordinal
Services provided				Narrative
Patients served/CHW			• Mean	Continuous
Reimbursement			• Patient pays fee for service • Patient receives care for free	Nominal
•INTERVENTION DESCRIPTION				
Number of intervention groups			•#	Continuous
Description of Intervention (who delivered, format and timing of delivery)				Narrative
Description of control				Narrative
Integrity of intervention				
•OUTCOMES				
List outcomes (definition, unit of measurement)				Narrative
For each outcome, time points (i) collected (ii) reported				

Unit of analysis			• Patient • Provider • Community • Other (specify) • Not clear	Nominal
Adverse or unintended effects?				Narrative
•RESULTS				
For each intervention group:				
Number of participants allocated			• #	Continuous
Sample size			• #	Continuous
Missing participants			• #	Continuous
Summary data			• 2x2 OR Mean, SD	
•PROGRAMME CHARACTERISTICS				
<i>Recruitment</i>				
Selected by			• Community • Health system • Both • Other (specify)	Nominal
Process			• Describe	Narrative
<i>Training</i>				
Amount			• <1 day • 1 day-1 week • 1 week - 3 months • > 6 months	Ordinal
Frequency			• Once • Weekly • Monthly • Quarterly • Yearly	Ordinal
Assessment			• Required (describe) • Not required	Dichotomous
<i>Supervision</i>				
Type			• Clinical • Peer • None	Nominal

Frequency			• None • Daily • Weekly • Monthly • Quarterly	Ordinal
<i>Incentivisation</i>				
Type			• Monetary (amount) • Non-monetary • None	Nominal
•MISCELLANEOUS				
Funding source				
Power calculation			• Done • Not clear • Not done	Nominal
Conclusions of study authors				Narrative
Equity considerations: evidence of differential impact on different parts of the population				Narrative
•Notes				

Appendix 6c: PRISMA Checklist

SECTION/TOPIC	#	CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	272
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	N/A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	272-279

Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	279
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	279-289
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	279-282
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	282-285
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	282
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	279-282
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	285-289
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix 6b
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	285-289
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	285-289
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if	285-289

		done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	285-289
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	285-289
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	289-304
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	289-304
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	304-309
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	309-323
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	304-209
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	323-324
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and	331-333

		at review-level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	335-340
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	N/A

Appendix 6d: Characteristics of studies

Characteristics of included studies (includes 'Risk of bias' tables)

Andreoni 2016

Trial reg.	AEARCTR-0000417
Country	Pakistan
Methods	STUDY DESIGN: RCT (2x2 Factorial) UNIT OF RANDOMISATION: LHW
Population	LOCAL TERM FOR CHW: Lady Health Workers (LHWs) CHWs: 100% female, mean 10 years of education, age not specified PATIENTS: Children under five HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: (A): 85 CHWs, (B): 84 CHWs CONTROL SAMPLE SIZE: (A): 88 CHWs, (B): 80 CHWs
Interventions	OBJECTIVE: Use structural estimates of time preferences to customise incentives for a sample of polio vaccinators during a series of door-to-door immunization drives in Pakistan and experimentally evaluate the effect of matching contract terms to individual discounting patterns. The study investigates the effect of tailoring, but also whether tailoring is more or less effective depending on when allocations are made. INTERVENTION DESCRIPTION: The function of LHWs is to provide oral polio vaccine to children during government organized vaccination drives; LHWs are given a supply of oral vaccine and a neighbourhood map and are asked to travel door-to-door vaccinating children with a suggested target for vaccination. LHWs are offered a fixed bonus of 1000 rupees (around \$10) for completing a total of $V = 300$ vaccinations over a two-day drive. Vaccinators set daily targets v_1 and v_2 corresponding to vaccinations on day 1 and day 2 of the drive, respectively. If either of the vaccination targets, v_1 or v_2 , are not met, the bonus is not received. Vaccinators are randomly assigned an interest rate, R , such that a single vaccination that is allocated to day 2 reduces by R the number of vaccinations required on day 1. That is, v_1 and v_2 satisfy the intertemporal budget constraint. There were two intervention groups: (A) Contracts tailored to individual discounting patterns with advance allocation choice and (B) Contracts tailored to individual discounting patterns with immediate allocation choice. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: No description of baseline training. All CHWs received a two hour training on the smartphone vaccination application and its correct use. Independent of treatment status, LHWs were also trained on the intertemporal bonus contracts and the process by which allocations were made and submitted.

	REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Daily SUPERVISION DESCRIPTION: At the end of each day of the drive, LHWs in each neighbourhood convened at a local clinic with their supervisor, and self-reported their vaccinations for the day. Though, in principle, a monitor could investigate the neighbourhood chalk markings, in practice, this system provided for virtually no monitoring. INCENTIVE: Monetary INCENTIVE DESCRIPTION: No explicit incentives for completing vaccinations were provided and LHWs received a fixed daily wage of 100 rupees (around \$1). As part of the trial, LHWs were offered fixed bonus of 1000 rupees (around \$10) for completing a total of $V = 300$ vaccinations over a two-day drive. RATIO CHW TO POPULATION: N.S. START DATE: September 26, 2014 END DATE: December 9, 2014 CONTROL DESCRIPTION: Prior to our project, LHWs self-reported achievement and no technology existed for monitoring vaccinations. Each vaccinator in the sample was provided a smart-phone, equipped with a reporting application that permits precise observation of when and where vaccinations are conducted. There were two control groups: (A) Untailored contracts and advance allocation choice and (B) Untailored contracts and immediate allocation choice.
Outcomes	PATIENTS: None CHWs: Closeness to the policy objective DURATION OF FOLLOW UP: Immediate
Funding source	International Growth Center (IGC), National Science Foundation, grant SES-1427355, and from the Department for International Development (DFID) Building Capacity for Research Evidence (BCURE) pilot programme initiative
Notes	BASELINE INPUT: Smartphones, as described above

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "We randomly provide half of LHWs with a tailored contract and half with a random contract" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	Comment: Not specified.
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	Low risk	Comment: None reported
Selective reporting (reporting bias)	Unclear risk	Comment: Outcomes listed in the protocol were embargoed so it is not possible to make a judgement

Ashraf 2015

Trial reg.	None
Country	Zambia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: District CLUSTERS: 24 Intervention, 24 Control
Population	LOCAL TERM FOR CHW: Community Health Agent (CHA) CHWs: 30% female, 18-45 years (mean 27 years), minimum 12 years of education PATIENTS: Women and children under 5 years, mean 4.2 years education

	HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 149 CHWs CONTROL SAMPLE SIZE: 149 CHW
Interventions	OBJECTIVE: Identify the effect of advertising career incentives at the recruitment stage. INTERVENTION DESCRIPTION: Recruitment materials make career possibilities salient by highlighting that CHAs are part of the Ministry of Health's hierarchy and that this gives them access to a career path leading to higher-ranked positions such as nurse, clinical officer, and doctor. (Being part of the civil service, CHAs are eligible for "in- service training," meaning that they attend school as a serving officer and the government pays their tuition for all of their training.) Example line from the poster: "Become a community health worker to gain skills and boost your career!" TRAINING DURATION (DAYS): 365 TRAINING DESCRIPTION: Learn basic medical practices at a residential training. REFRESHER OR ONGOING TRAINING: Refresher training referenced but not described. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: 5% of home visits are audited; CHWs meet with their supervisors at an unspecified frequency. INCENTIVE: Monetary INCENTIVE DESCRIPTION: USD 290/month RATIO CHW TO POPULATION: N.S. START DATE: August, 2010 END DATE: January, 2014 CONTROL DESCRIPTION: Recruitment materials make salient benefits to the community, thus making the CHA position look similar to existing informal positions (e.g., village health workers, traditional birth attendants, barefoot doctors) that are common in these areas. Example line from the poster: "Want to serve your community? Become a community health worker!"
Outcomes	PATIENTS: Use of health services CHWs: Number of household visits completed over the study period, number of patients seen at health post, number of community meetings organised DURATION OF FOLLOW UP: 18 months
Funding source	IGC, JPAL Governance Initiative and HBS DFRD
Notes	BASELINE INPUT: CHAs, after a year of training, are required to devote 80% of their time (4 out of 5 working days per week) to household visits. The visits' main goals are to provide advice on women's health—including family planning, pregnancy, and postpartum care—and child health—including nutrition and immunizations. In addition, CHAs are expected to inspect the household and provide advice on health-related practices such as safe water practices, household waste management, sanitation, hygiene and ventilation. During visits, CHAs are also tasked with providing basic care to any sick persons and referring them to the health post as needed. In the remaining time, CHAs are expected to assist staff at the health post (the first-level health facility in rural Zambia) by seeing patients, assisting with antenatal care, and maintaining the facility. They are also supposed to organize community meetings such as health education talks at the health post and in schools.

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Random assignment of the 48 districts is stratified by province and average district-level educational attainment" Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of	Unclear risk	Comment: Not specified.

participants and personnel (performance bias)		
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	High risk	Quote: "18% of CHAs drop out" after one year. "This attrition rate is balanced across treatments"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified
Recruitment bias	Low risk	Comment: CHWs were recruited after the study, but by district officials exposed to one treatment and unaware of the other
Baseline imbalance	Unclear risk	Comment: Not specified
Loss of clusters	Unclear risk	Comment: Not specified
Unit of analysis issues	Low risk	Quote: "standard errors clustered at the district level"

Ayele 1993

Trial reg.	None.
Country	Ethiopia
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: District CLUSTERS: Intervention: 1, Control: 1
Population	LOCAL TERM FOR CHW: Community health agents (CHAs) CHWs: 100% Male. Mean age 31.0 years. Mean schooling 5.0 years. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 50 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 50 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Determine the effect of a refresher course and monthly supervision on the health service activities of CHWs INTERVENTION DESCRIPTION: Intervention group (N= 50 CHWs) received (1) Training: refresher course (2) Supervision: monthly INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: CHWs in the intervention district received (1) a five-day refresher course based on the "expected functions of the CHWs and other relevant responsibilities which they are asked to carry out in the provision of the health care to their communities" REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Educative supervision by health unit staff. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 06/1990 END DATE: 12/1990 CONTROL DESCRIPTION: CHWs (N = 52 CHWs) received nothing after baseline input (see notes). CHWs in control group had gone an average of 1.9 years since last refresher training and 30.8% had never received supervision.
Outcomes	PATIENTS: None. CHWs: (1) Composite functional status score (FSS) computed by summing the scores for the 13 CHW activities (2) 13 activity scores: (i) Outreach (days/mo.), (ii) Health education (sessions/mo.), (iii) Environmental health (activities/mo.), (iv) Maternal and child health (activities/mo.), (v) Expanded programme on immunization (activities/mo.), (vi) Births (# registered/mo.), (vii) Deaths (# registered/mo.), (viii) Curative (pts./mo.), (ix) Reports (#/mo.), (x) School health (activities/mo.), (xi) Home visits (#/mo.), (xii) Referrals (#/3mo), (xiii) Epidemic

	control measures (#/3mo). DURATION OF FOLLOW UP: Collected monthly for 6 months and reported at 3 months and 6 months.
Funding source	Not specified.
Notes	BASELINE INPUT: (1) CHWs were provided with registration books, standard Ministry of Health of Ethiopia monthly report forms, and stationery sufficient to last a minimum of one year. (2) General meetings were held in each peasant association with the community leaders and the community at large to discuss issues related to the CHW's job description, roles, and responsibilities. (3) Communities were encouraged and helped to establish or strengthen health committees in their respective villages.

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "districts were randomly assigned" (p. 380) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra supervision and refresher training).
Blinding of outcome assessment (detection bias)	Low risk	Quote: "These supervisors were aware of the study but not the district comparison or study hypothesis."
Incomplete outcome data (attrition bias)	Low risk	Comment: 0% CHW attrition (personal communication with author)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to the health zones post randomisation.
Baseline imbalance	High risk	Comment: baseline comparability of clusters was reported, not adjusted for.
Loss of clusters	Low risk	Comment: no clusters were lost
Unit of analysis issues	High risk	Clustering not taken into account in the analysis

Bailey 1996

Trial reg.	None
Country	Guatemala
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Promotore
Population	LOCAL TERM FOR CHW: Promotores (Village health promoters) CHWs: N.S. PATIENTS: N.S. HEALTHCARE SETTING: Community INTERVENTION SAMPLE SIZE: 22 CHWs CONTROL SAMPLE SIZE: 27 CHWs
Interventions	OBJECTIVE: Assess the effectiveness of a newly developed method for rapid tiered training of health workers in improving community health worker knowledge and case management skills. INTERVENTION DESCRIPTION: Train CHWs using the interactive "Kader" method, developed in West Java, Indonesia. The Kader method employs tiered training whereby trainers teach trainees exactly as the trainers themselves have been taught. The training relies on role-playing by trainees using "Counselling Cards" to practice counselling or teaching, precisely as the trained village health promoters will

	counsel mothers in the community. The counselling cards include a diagnostic algorithm and four treatment cards outlining recommendations for five type of diarrhoea. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Promotores were trained by rural health workers ("Técnicos") in either interactively or didactically depending on trial arm REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Promotores are supervised by técnicos INCENTIVE: None INCENTIVE DESCRIPTION: Promotores are volunteers RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Train CHWs using traditional didactic training
Outcomes	PATIENTS: None. CHWs: Knowledge score. (The tests consisted of five case descriptions which were used to assess respondents' abilities to identify five diarrhoea types and the appropriate referral and treatment recommendations for each type. General knowledge regarding signs of dehydration, diarrhoea prevention, and methods for making oral rehydration solution was assessed through independent questions.) DURATION OF FOLLOW UP: Immediate
Funding source	Birmingham, Alabama chapter of The Partners of the Americas, the Cadeceus Club of the University of Alabama School of Medicine and the Larry Mayes Memorial Scholarship
Notes	BASELINE INPUT: The responsibilities of the técnicos include: overseeing vaccination campaigns, education and supervision of promotores, and coordination of health education, promotion, sanitation and water purification efforts in their districts. The promotores are responsible for promoting Department of Public Health programs at the village level and for providing a health advice and referral network for the district health clinics.

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "promotores were randomly assigned to receive Kader (experimental) or didactic (control) training" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified
Blinding of participants and personnel (performance bias)	High risk	Quote: "There is also possible bias secondary to the lack of blinding of participants and trainers"
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified
Incomplete outcome data (attrition bias)	Low risk	Quote: "three promotores had to be excluded from analysis due to incomplete data." Comment: <6% attrition
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified for this review

Bossuroy 2016

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT (4 arm)

	UNIT OF RANDOMISATION: Health worker CLUSTERS: 78 (Intervention/Control breakdown not specified)
Population	LOCAL TERM FOR CHW: ASHA CHWs: 72% male. 50% have some university education. Recruited by Delhi-based NGO Operation ASHA PATIENTS: 57.7% male, 57.6% no school or some primary HEALTHCARE SETTING: Community/Operation ASHA centre SAMPLE SIZE: 78 (Intervention/Control breakdown not specified)
Interventions	OBJECTIVE: Investigates whether financial incentives provided to health workers may encourage them to detect new tuberculosis cases and improve treatment adherence. INTERVENTION DESCRIPTION: For the first six months, CHWs were randomly assigned to receive either a fixed salary or a salary dependent on the number of patients they had detected. In the following six months, they were randomly re-assigned to either a fixed or an incentivized salary scheme, based on the number of defaults. CHWs were randomly assigned to one of four treatment arms: they either received (A) financial incentives based on patient detection for six months and incentives based on treatment adherence subsequently, (B) financial incentives based on patient detection for six months and a fixed salary subsequently or (C) a fixed salary for six months and incentives based on treatment adherence subsequently TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Part of intervention being tested, see above. INCENTIVE DESCRIPTION: Part of intervention being tested, see above. RATIO CHW TO POPULATION: 1:33 START DATE: December 2009 END DATE: April 2013 CONTROL DESCRIPTION: A fixed salary for the whole duration of the experiment.
Outcomes	PATIENTS: TB defaults (number of patients leaving the DOTS system during the course of their treatment) CHWs: Case detections (number of patients enrolled in the DOTS system), Health workers' effort/motivation DURATION OF FOLLOW UP: 6 months for each outcome
Funding source	3ie's donors, which include UK aid, the Bill & Melinda Gates Foundation and the Hewlett Foundation.
Notes	Note that the number of months during which each phase was carried out differs across cities, and may not be 6 months in all cases. Also during the second phase, some new CHWs who were not a part of the first phase were randomized into the experiment. BASELINE INPUT: Operation ASHA hires community-based health workers, who are each responsible for operating two DOTS centres. During an initial period of about 3 months, CHWs work on setting up the centres, getting to know the community and surroundings, and making the centre known to the local population. During the first 3 months, CHWs all receive a fixed salary. The experiment starts after the initial three months of a centre lifespan.

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "CHWs are randomly assigned" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Unclear how this was done
Blinding of	High risk	Comment: not possible to blind receiving incentives

participants and personnel (performance bias)		
Blinding of outcome assessment (detection bias)	Low risk	Comment: Two primary outcomes (detection and default) were assessed using TB registers and lab registers kept by public health TB officers
Incomplete outcome data (attrition bias)	Low risk	Quote: "Even though 78 CHWs were initially randomized, 3 CHWs left Op ASHA before their baseline survey...we use all 78 CHWs for impact on attrition"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified
Recruitment bias	Low risk	Comment: CHWs were working before randomisation
Baseline imbalance	Low risk	Comment: Only 5 differences out of 44 are significant at the 10 per cent level and 2 are significant at the 5 per cent level, in line with what should be expected.
Loss of clusters	Unclear risk	Comment: Not adequately described
Unit of analysis issues	Low risk	Quote: "In...all...regressions, we adjust standard errors for clustering at the health worker level since the randomization was conducted at this level."

Carrón 1994

Trial reg.	None
Country	Paraguay
Methods	STUDY DESIGN: C-RCT (2x2 Factorial) UNIT OF RANDOMISATION: Groups (each encompassing 3 districts) CLUSTERS: (A) 3 districts, (B) 3 districts, (C), 3 districts, (D) 3 districts
Population	LOCAL TERM FOR CHW: Community-based distribution (CBD) providers CHWs: 100% female. Mean age and mean education level not reported. PATIENTS: Sex, mean age and mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: I _a =45 CHWs, I _b =45 CHWs, I _{a+b} =45 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: C=45 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Develop tools of low-cost use to improve the quality of care provided by rural CHWs in Paraguay. Specifically, to determine the relative effects of (1) an interactive counselling guide based on a service algorithm and (2) a supervisory behavioural feedback instrument on the quality of CHW provided care. INTERVENTION DESCRIPTION: Group A (N = 45 CHWs) received (1) Training: use of interactive service algorithm [ABC of the Distributor]. Group B (N = 45 CHWs) received (2) Supervision: use of role-play based instrument to diagnose and provide feedback on CHW service behaviour [Behavioural Feedback Instrument]. Group C (N = 45 CHWs) received (3) Training + Supervision: Use of ABC of the Distributor service algorithm and Behavioural Feedback Instrument-based supervision. INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): <1 TRAINING DESCRIPTION: (1) ABC of the Distributor: The tool was taught by nurse midwives to CHWs in 30 min individual meetings. (2) Behavioural Feedback Instrument (BFI): CHW were trained by nurse midwives in four 30 min individual meetings. REFRESHER OR ONGOING TRAINING: A two-day course in family planning service provision is imparted once a year, serving as a basic course for untrained distributors and as a refresher course for trained ones. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S.

	START DATE: 07/1992 END DATE: 07/1993 KEY CHW COMPETENCY: CONTROL DESCRIPTION: Group D (N = 45 CHWs) CHWs received the usual two-day annual family planning course and yearly refresher training.
Outcomes	PATIENTS: None. CHWs: Pretest-posttest gains in quality of care provided. Assessed using (1) an observational service test in which a trained simulated client asks services of a CHW and enacts a learnt client profile as she interacts with the provider. When the consultation ends, the simulated client fills out a checklist (range: 41-53 items, depending on client profile) indicating whether the promoter emitted each of the expected provider behaviours; the quality of care score is the sum of the items checked. The service test was administered to all CHWs. (2) A client questionnaire examining 11 indicators across 5 domains: (i) Method choice (user chooses), (ii) Information given to the client (user knowledge of other methods, user knowledge of action upon forgetting, provider advised side effects/method shift), (iii) Interpersonal relations (positive user perceptions), (iv) Provider Technical Competence (physical complaints, contraindications, problems with method, satisfaction with method, unwanted pregnancies), (v) Follow-up mechanisms (provider focus on revisit). The client questionnaire was administered to one patient per CHW randomly selected from the client register. Both the service test and the client questionnaire were conducted by the same trained simulated client. DURATION OF FOLLOW UP: All outcomes collected and reported at baseline and after 12 months.
Funding source	US Agency for International Development (USAID), Office of Population

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: groups were "randomly assigned" (p. 10) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training or supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement
Incomplete outcome data (attrition bias)	High risk	Comment: 156/180 CBD found, 127/180 clients found. Quote: "Provider mortality affected unevenly the distribution of the six client profiles of the Service Test across provider groups. To avoid possible distortions in group comparisons, the equal distribution of client profiles across groups was re-established through a process whereby, for each client profile, (a) the lowest number of cases was determined and (b) the number of cases in the other groups was reduced to this level through random elimination of cases. 124 cases, 31 per group remained in the analysis" (p. 13) Comment: Randomly deleting data is not the same as never having collected the data in the first place; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups.
Selective reporting	Unclear risk	Comment: Insufficient information to permit judgement

(reporting bias)		
Recruitment bias	Low risk	Comment: No individuals recruited to groups post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	High risk	Comment: Clustering not taken into account in the analysis.

Chang 2011

Trial reg.	NCT00675389
Country	Uganda
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Clinic catchment area CLUSTERS:
Population	LOCAL TERM FOR CHW: Community-based peer health workers (PHW) CHWs: 100% female. Mean age and mean education level not reported. PATIENTS: HIV+ patients 66% female, 34% male. Median age 34 years (range: 15-67) in intervention group and 36 (16-76) in control group. Mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 13 CHWs, 446 patients CONTROL SAMPLE SIZE: 16 CHWs, 524 patients
Interventions	OBJECTIVE: Determine the effect of an mHealth (mobile phone) CHW support intervention on AIDS care in rural Uganda. INTERVENTION DESCRIPTION: CHWs randomized to the intervention were each given a mobile phone, a one-day residential training, and an hour-long field-based practicum training on the mHealth intervention. These CHWs were asked to send a text message reporting adherence and clinical data back to a centralized database immediately after or during each home visit. These texts were numeric codes produced by converting all quantitative data on the home visit form to simple integers arranged in a predetermined order separated by asterisks, (e.g. pills given, taken, defaulted might convert to *282828*). CHWs in the intervention arm were also encouraged to call a Rakai Health Sciences Program mobile phone hotline or toll-free warmline (staffed only during clinic hours) with questions or concerns. Clinic staff receiving CHW texts and calls could opt to provide care instructions to CHWs, send a higher-level care provider to the patient, or arrange transport to health care facilities. INTERVENTION TYPE: Reporting-Supervision TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: PHWs received a residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. In addition to their baseline training, PHWs randomised to the mHealth Arm were each given a mobile phone, a one day residential training, and an hour long field-based practicum training on the mHealth intervention. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: PHW day-to-day supervision was largely performed by a single staff member working part-time. INCENTIVE: Financial INCENTIVE DESCRIPTION: Monthly allowance of about 12.5 USD RATIO CHW TO POPULATION: 1:34 START DATE: 04/2006 END DATE: 06/2008 CONTROL DESCRIPTION: CHWs who did not receive the mHealth intervention continued providing regular care (see baseline input).
Outcomes	PATIENTS: Primary: cumulative risk of virologic failure (any failure during follow-up period equalling failure with viral loads conducted every 24 weeks; failure defined

	as .400 copies/mL). Secondary: (i) patient adherence (pill counts and 3 day self-report), (ii) virologic failure at 24 and 48 weeks of ART (failure defined as 400 copies/ml), (iii) lost to follow- up (lack of a pharmacy refill visit in over 90 days), and (iv) mortality (ascertained through verbal autopsies) DURATION OF FOLLOW UP: A summary clinic pill count was calculated by dividing the number of pills taken over the study period by the sum of pills expected to be taken over the study period. Viral loads were measured at 24 to 192 week time points from antiretroviral therapy initiation by 24-week intervals. CHWs: None.
Funding source	Doris Duke Charitable Foundation, The Division of Intramural Research, The National Institute for Allergy and Infectious Diseases, National Institutes of Health, and National Institutes of Health Training (2T32- AI07291) and Career Development (1K23MH086338-01A2) Grants.
Notes	BASELINE INPUT: CHWs received a two-day residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. CHWs provided clinical and adherence monitoring and psychosocial support to fellow patients. At home visits, CHWs recorded a review of symptoms, an adherence self-report, and pill count results using a standardized form. CHWs were each given a bicycle, t-shirts, basic supplies, and an initial monthly allowance of about 12.5 USD. CHW day-to-day supervision was largely performed by a single RHSP staff member working part-time.

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "29 PHWs at 10 clinics were randomized by clinic to receive the intervention or not." Comment: Unclear how done.
Allocation concealment (selection bias)	Low risk	Quote (parent trial): We assigned clusters using unmatched, unrestricted random allocation by a drawing of lots. Comment: Author provided the following information via email: "allocation was done simultaneously with the randomisation process"
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training and a mobile phone).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Author provided the following information via email: "this was a biological sample...no special measures taken"
Incomplete outcome data (attrition bias)	Low risk	Quote (parent trial): 9.3% died; 2.2 % LTFU. "Efficacy analyses for cumulative risk of virologic failure and for virologic outcomes at specific time points from ART initiation were by intention to treat using log-binomial regression with generalized estimating equations (GEE)"
Selective reporting (reporting bias)	Low risk	Comment: Parent trial registered on ClinicalTrials.gov (NCT00675389). All outcomes reported.
Recruitment bias	Low risk	Comment: No individuals recruited to clinic catchment area post randomisation.
Baseline imbalance	Low risk	Quote: Sociodemographic characteristics, immunologic and clinical stage of disease, the proportion of patients on ART, the median duration of time patients were on ART prior to the start of the trial, and the pre-trial 24 week and 48 week virologic

		failure” were well balanced between arms.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Quote: “...robust variance estimates appropriate for cluster-randomised trials”

DeRenzi 2012A

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT (Pilot study) UNIT OF RANDOMISATION: CHW
Participants	LOCAL TERM FOR CHW: community health worker CHWs: 100% Female. Age range 23-55 years. Mean schooling unreported. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 8 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 7 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: CHWs received an SMS message two days after they had reported a referral with CommCare, reminding them of the follow-up visit. Daily SMS messages were sent until the follow-up was recorded. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: 04/2010 END DATE: 06/2010 CONTROL DESCRIPTION: CHWs received nothing.
Outcomes	PATIENTS: Change in percentage of closed referrals CHWs: None TIMEPOINTS COLLECTED/REPORTED: Referral and follow-up information was collected automatically by CHWs using CommCare, and sent immediately to the server by the CHW. 24 days
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 39 and 24 days long, respectively. Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: " participants were randomly assigned" (p. 27) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS notifications)
Blinding of outcome	Unclear risk	Comment: insufficient information to permit judgement.

assessment (detection bias)		
Incomplete outcome data (attrition bias)	Low risk	Comment: “Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period.”
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012B

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW (stratified by baseline performance)
Participants	LOCAL TERM FOR CHW: community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified, after attrition: I =34, C = 40), number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: Escalating reminder system that sent SMS reminders directly to the CHW before notifying the CHW's supervisor after several overdue days (augments existing supervision structures). INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: CHWs received nothing after baseline input.
Outcomes	PATIENTS: Mean time chronic care patients overdue. CHWs: None. TIMEPOINTS COLLECTED/REPORTED: Data collected on a rolling basis, sent from CHWs post closed referral (40 days)
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 40 days each. BASELINE INPUT: Pathfinder CHWs are all volunteers; they are provided with a “generous transportation stipend” during their monthly meetings and incentivized with personal use of the mobile phone they are given for CommCare. CHANGES FROM PILOT: In the pilot study (DeRenzi 2012A), the earliest reminder was sent on the evening that the CHW's visit was due, too late to make an on-time visit. For the Dar es Salaam studies (DeRenzi 2012B, DeRenzi 2012C), <i>proactive</i> reminders were incorporated, with the first one being sent the day before a visit is due. Second, if a referral was not closed during the pilot, the system continued to send SMS messages but had no means of determining <i>why</i> the CHW was not reporting a follow-up. This issue directly inspired the escalation to supervisor evaluated in DeRenzi 2012B.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: " CHWs in each bin were then randomly assigned to the intervention or control group." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS messages and supervisor intervention).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.
Incomplete outcome data (attrition bias)	High risk	Quote: "Although 87 CHWs participated in this study, 13 participants were removed before analysis for one of the following reasons: no data submitted during the intervention period (11 CHWs), no active cases (1 CHW), and technical issues with CommCare or the phone (1 CHW). This removal left 34 CHWs in the SMS+Supervisor group and 40 CHWs in the control group." "During the course of the intervention, there were occasional phone problems (e.g. accidentally deleting the CommCare application) and discrepancies (e.g. a CHW reporting s/he had already sent data) reported from CHWs in the field. If there was a discrepancy, we removed the client from our analysis. This happened for 27 clients during the intervention period, which represented 3.7% of the total number of clients visited during that period."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012C

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF ALLOCATION: CHW
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified), number of patients unspecified

Interventions	OBJECTIVE: Although only 6.5% of cases in DeRenzi 2012B required the involvement of the supervisor, this escalation step is still more costly and burdensome than sending SMS messages alone. The objective of this study was to quantify the effect of escalating to the supervisor versus SMS messages alone in the context of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION: Intervention group (SMS-Only) was provided only with proactive and overdue reminders, but escalation to the supervisor never occurred. MODE OF DELIVERY: The system, implemented on top of CommCare, first sends proactive reminders to intervention CHWs (the day before (x1) and the day of (x2) a scheduled routine visit. Daily reminders are sent while the visit remains overdue, informing the CHW of the # of days the visit is overdue. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL: CHWs (SMS+ Supervisor) received exactly the same input as the intervention group in DeRenzi 2012B.
Outcomes	PATIENTS: None. CHWs: Mean time overdue. TIMEPOINTS COLLECTED/REPORTED: 90 days
Funding source	Not specified.
Notes	90-day baseline period and a 90-day intervention period. BASELINE INPUT: Since the baseline period followed directly after DeRenzi 2012B and with the same CHWs, some CHWs continued to receive escalating SMS reminders during the baseline, while others did not. H1: For CHWs who received escalating SMS reminders during the baseline period, the SMS-Only intervention (removing the supervisor) would decrease performance compared to continuing in the SMS+Supervisor condition. H2: For CHWs who did <i>not</i> any receive SMS reminders during the baseline period; the SMS+Supervisor intervention would result in a larger increase in performance than the SMS-Only intervention. The studies presented here measure reported follow-ups and are not correlated with ground truth of actual follow-ups.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: " CHWs were then randomly assigned to one of two conditions" Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received SMS and/or additional supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.
Incomplete outcome	High risk	Quote: "Of the original 87 CHWs, 26 were excluded

data (attrition bias)		from analysis for the following reasons: incorrect phone number (2 CHWs), having no data during baseline, intervention, or both periods (22 CHWs), and technical problems with the phone (2 CHWs). In the end, there were 32 CHWs originally from the intervention group in Study 1 (16 SMS+Supervisor, 16 SMS-Only) and 29 from the control group in Study 1 (14 SMS+Supervisor, 15 SMS-Only).”
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

Gautham 2015

Trial reg.	N.S.
Country	India
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Rural health provider CHWS: 8 male, 8 female; existing rural health providers; Males: mean age of 48.75 years, 10 years of education or more. Females: mean age 24.75 years, 10 years of education or more PATIENTS: Paediatric and Adult patients HEALTHCARE SETTING: Community (personal communication with S. Iyengar) INTERVENTION SAMPLE SIZE: 8 CONTROL SAMPLE SIZE: 8
Intervention	OBJECTIVE: (a) Use a randomized control design to measure changes in protocol compliance (PC) by health workers in their everyday work settings, (b) assess the usability and acceptability of the mobile application with health workers in the experimental group and (c) obtain patient feedback on health workers' use of the mobile system during treatment. INTERVENTION DESCRIPTION: Mobile phone-based, media-rich procedural guidance applications TRAINING DURATION (DAYS): 2 days TRAINING DESCRIPTION: Training programme in guideline-based care and in use of the mHealth system by physicians who spoke the local language. The study team assessed all RHPs on their understanding of guideline-based care using a simple assessment instrument developed by the research team, on which every RHP received a pre-and post-training test score (PTTS). REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: A set of paper guidelines to use in the field
Outcomes	PATIENTS: None CHWs: Protocol compliance DURATION OF FOLLOW UP: 2 months
Funding	Information Society Innovation Fund (ISIF)

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Blinding of participants and personnel (performance bias)	High risk	Quote: "after the training, the RHPs in the experimental group went back to their field sites with the applications on their mobile phones, while for the ones in the control group, the data card with the application was taken out" Comment: Unclear if patients were blinded. CHWs were not blinded.
Blinding of outcome assessment (detection bias)	High risk	Quote: "PC was evaluated by direct observations of the RHPs' performance as they were treating the patients."
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Not described
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Ginsburg 2014

Trial reg.	N.S.
Country	Bangladesh
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Community health workers CHWS: Female; age N.S.; high school education; recruited from rural villages by study team PATIENTS: Adult women aged 25 and older and living in the catchment area of Khulna Division, Bangladesh. Exclusion criteria were women younger than age 25 and women unable to give verbal consent. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: A: 246 B: 240 CONTROL SAMPLE SIZE: C: 70
Intervention	OBJECTIVE: Test whether CHWs guided by smart phone applications would be more efficient and effective than CHWs without phones at encouraging women with an abnormal breast exam to adhere to advice regarding a clinic appointment INTERVENTION DESCRIPTION: Arm A: smart phone with applications to guide interview, report data, show motivational video, and offer an appointment for women with an abnormal CBE. Arm B: smart phone platform as in arm A plus additional "patient navigation training." TRAINING DURATION (DAYS): 3 TRAINING DESCRIPTION: CHWs and supervisors attended an intensive training programme on breast anatomy, physiology, and pathology, as well as the 2-hour training session on CBEs. CHWs were taught in a gentle, culturally sensitive manner how to obtain informed verbal consent, to ask the structured interview questions, and to perform a CBE. Only the 10 CHWs randomized to arm B were given an extra day of training in patient navigation. REFRESHER OR ONGOING TRAINING: N.S.

	SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Three female CHW supervisors, also local women but with postsecondary education and prior experience in the health or social services sector, were hired to oversee daily workflow and to liaise with our field operations team. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:19 START DATE: July, 2012 END DATE: October, 2012 CONTROL DESCRIPTION: No smart phone; same interview recorded on paper
Outcomes	PATIENTS: The "adherence" (to advice regarding a clinic appointment) for women with an abnormal CBE CHWs: N.S. DURATION OF FOLLOW UP: 1 month
Funding	Rising Stars in Global Health award from Grand Challenges Canada.

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Low Risk	Quote: "The 30 CHWs were randomized by manual lottery (by O.M.G., eyes closed, hand in jar with identical folded papers with CHW names) to one of the three study arms."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding of participants and personnel (performance bias)	Unclear	Comment: Unclear if patients were blinded. CHWs could not be blinded.
Blinding of outcome assessment (detection bias)	Low risk	Comment: clinic records used, so low risk of bias
Incomplete outcome data (attrition bias)	Unclear risk	Comment: No attrition reported
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Omer 1998

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Enumeration area CLUSTERS: Intervention: 5, Control: 5
Population	LOCAL TERM FOR CHW: Lady health workers (LHWs) CHWs: 100% female. Mean age and education level not reported. PATIENTS: 100% women, of whom 45% report receiving "some formal education." Mean age not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 52 CHWs, 529 patients CONTROL SAMPLE SIZE: Unspecified number of CHWs, 541 patients
Interventions	OBJECTIVE: Demonstrate the effective use of community-based evidence for health promotion by CHWs in Sindh, Pakistan. INTERVENTION DESCRIPTION: In a consultation forum convened for the purpose, CHWs and government health officials developed a targeted tool to guide household discussions. It was designed to be "user-friendly, indigenous, and attractive" to both CHWs and their clients. The tool was based on ajrak, a traditional Sindh fabric, with the message in embroidery. The design team also

	consulted with the provincial team of the National Programme, CHWs, women in different urban and rural communities, local designers and handicraft experts. The embroidery depicted maternal practices like attending and not attending antenatal check-ups, giving colostrum after birth and not doing heavy work. The team pre-tested the embroidery in five communities not included in this study, ensuring input from CHWs and women in the communities to adjust the tools. In intervention communities, CHWs showed these panels to pregnant women during their household visits and invited them to describe what was going on in the pictures. They encouraged their clients to express their views of the situations depicted on the cloths and discussed any difficulties they might face. To go for prenatal check-ups, for example, the discussion might include when and how the person would go. Training in the tool included an explanation of the tools, role modelling and practice in the community. INTERVENTION TYPE: Training/integration TRAINING DURATION (DAYS): Part of intervention TRAINING DESCRIPTION: Part of the intervention. REFRESHER OR ONGOING TRAINING: Part of the intervention. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Monitoring relied on the existing LHW supervisory structure in both intervention and control communities. As part of routine control visits, supervisors did random checks to see if LHWs used the tools in intervention communities. LHSs held monthly meetings at district level, where they would get feedback on progress and problems they were not able to rectify on their own. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 06/2000 END DATE: 04/2001 CONTROL DESCRIPTION: CHWs (N = unspecified) in control communities did not use the embroidered cloths and received no training in how to use them, but continued their usual work visiting women before and after childbirth to discuss safe practices in pregnancy based on the standard training of the CHW programme. (N = 541 patients)
Outcomes	PATIENTS: (i) attendance at prenatal checkups; (ii) stopping of routine heavy work during pregnancy (iii) giving colostrum to newborn babies (iv) maintaining exclusive breastfeeding for four months. CHWs: Performance (# of visits to pregnant women) DURATION OF FOLLOW UP: All outcomes were collected and reported at baseline and in a household survey after CHWs in intervention communities had used the tools for 10 months.
Funding source	Canadian International Development Agency's (CIDA) Canada Fund for Local Initiatives in Pakistan.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We allocated these randomly using a computerised random number generator" (p. 180)
Allocation concealment (selection bias)	Unclear risk	Quote: "...using a computerised random number generator" (p. 180) Comment: Unclear if enumeration areas were assessed before number appeared.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training on a novel communication tool).
Blinding of outcome	Low risk	"Fieldworkers, who did not know which communities were

assessment (detection bias)		intervention and which were control, collected data" (p. 180)
Incomplete outcome data (attrition bias)	Low risk	Quote: "Analysis followed the intention to treat principle, with all women in the intervention communities analysed as though exposed, whether or not they had seen the intervention materials." (p. 180)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Other bias	Low risk	In cluster-randomized trials, particular biases to consider include: (i) Recruitment bias Comment: Low risk; no individuals recruited to enumeration area post randomisation. (ii) Baseline imbalance Comment: Unclear risk; baseline comparability of clusters not reported, not adjusted for. (iii) Loss of clusters Comment: Low risk; no clusters were lost (iv) Unit of analysis issues Comment: Unclear risk; not clear if clustering taken into account in the analysis. (v) Comparability with individually randomized trials Comment: Irrelevant, all studies in this review are C-RCTs.
Recruitment bias	Low risk	Comment: No individuals recruited to enumeration area post randomisation.
Baseline imbalance	Unclear risk	Comment: Baseline comparability of clusters not reported, not adjusted for.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Unclear risk	Comment: Not clear if clustering taken into account in the analysis.

Winch 2003

Trial reg.	None
Country	Mali
Methods	STUDY DESIGN: Pair matched C-RCT UNIT OF RANDOMISATION: Health zones CLUSTERS: Intervention: 5; Control: 5 DURATION: 15 months (Visual aids: 11/2001-02/2002; Referral mechanism: 03/2001-06/2002)
Population	LOCAL TERM FOR CHW: Gérant (village drug kit managers) CHWs: Sex, age and education level not reported. PATIENTS: The intervention was directed at children <14 years. Mean schooling of carers was 0.85 years. Sex of all patients treated was not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: ~102 CHWs in both arms, 152 patients in the intervention arm, 134 in the control arm
Interventions	OBJECTIVE: Evaluate an intervention to (i) improve the skills of CHWs to counsel parents on correct home administration of chloroquine (CQ), and (ii) increase the referral of sick children to community health centres (CHC). INTERVENTION DESCRIPTION: (1) Use of visual aids for drug administration and danger sign recognition (2) Introduction of a referral/counter referral recordkeeping mechanism (3) Refresher training (4) Community engagement INTERVENTION TYPE: Training-Reporting- Integration TRAINING DURATION (DAYS): Part of the intervention TRAINING DESCRIPTION: Part of the intervention REFRESHER OR ONGOING TRAINING: a 5 day additional training course was conducted in a local elementary school that covered: (i) doses of CQ for different age groups (also covered in the standard course), (ii) how to counsel carers of young children regarding the correct administration of CQ in the home and the importance

	of going to the CHC when referred, (iii) indications for referral including danger signs and failure of the child's symptoms to respond to treatment, (iv) basic evaluation of a sick child for danger signs such as dehydration requiring referral, and (v) how to use the referral notebook and the visual aids SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: ~1:3 START DATE: 03/2001 END DATE: 06/2002 CONTROL DESCRIPTION: When a new village drug kit is established, all managers go through standard training which consists of a 35-d literacy course in which they are taught to read and write the national language, Bambara, using a phonetic alphabet, followed by a 1-week course on how to manage the drug kit including the correct doses to recommend and sell for each product in the kit. All of the drug kit managers in both the intervention and comparison groups had had this standard training prior to the study.
Outcomes	PATIENTS: The proportion of children with indications for referral to a health facility who were referred to and subsequently reached the facility. CHWs: (i) The proportion of children who were treated with a full course of CQ in the home DURATION OF FOLLOW UP: At least twice a week an interviewer visited the CHW, talked to him/her and examined the register of sales maintained, noting down the cases where CQ had been sold to treat a child aged <6 years. The interviewer assessed the effect of the intervention on the quality and content of the counselling received from the CHW through direct questions during follow-up in the home 5 days after the sale. Data from health centre registers and CHW referral notebooks was collected at the end of the trial. All data was reported at last observation.
Funding source	United States Agency for International Development through the Family Health and Child Survival Cooperative Agreement with Johns Hopkins Bloomberg School of Public Health

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "1 health zone from each pair was randomly assigned to receive the intervention." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra materials, refresher training and referral recording materials).
Blinding of outcome assessment (detection bias)	Unclear risk	Quote: "Fifteen field interviewers were recruited. In 6 of the 10 health zones there was 1 interviewer only, while the other 4 health zones had 2 or 3 interviewers due to their large size. Each interviewer was assigned a set of village drug kits and their clients to study. At least twice a week the interviewer visited the manager of the village drug kit, talked to him/her and examined the register of sales maintained by him/her, noting down the cases where CQ had been sold to treat a child aged <6 years. On the fifth day after the sale, the field interviewer went to the child's house and interviewed the person who had administered the treatment to the child." (p. 483) Comment: Insufficient information to permit

		judgement of 'Low risk' or 'High risk'
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Insufficient reporting of attrition/exclusions to permit judgement.
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to health zones post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Comment: Intraclass correlation coefficient taken into account in the analysis.

Characteristics of excluded studies

Record name	Year	Reason for exclusion
Afenyadu	2005	Did not meet the definition for CHW (CHWs were teachers)
Ahmed	1993	Not a randomised controlled trial
Akogun	2001	Not a randomised controlled trial
Alam	2012	Not a randomised controlled trial
Amouzou	2014	Non-CHW comparator
Ashraf	2014	Did not meet the definition for CHW (CHWs were hair stylists)
Avula	2011	Not a randomised controlled trial
Baqui	2008	Non CHW comparator
Begum	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Bertrand	1986	Non CHW comparator
Bhatt	1981	Not a randomised controlled trial
Bhutta	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Björkman	2009	Health workers were primarily facility-based
Bojang	2011	Non CHW comparator
Boone	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Brieger	2000	Health workers were primarily facility-based
Chin-Quee	2016	Not a randomised controlled trial
Curtale	1995	Not a randomised controlled trial
Das	2014	Effects of the CHW performance intervention cannot be disaggregated
Davies-Adetugbo	1997	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Elder	1992	Not a randomised controlled trial
Findley	2013	Not a randomised controlled trial
Florez-Arango	2011	Did not meet the definition for CHW (one quarter of assessed health workers were primarily facility-based)

Foreit	1984	Health workers were primarily facility-based
George	2012	Did not meet the definition for CHW (Not part of health system)
Harvey	2008	Not a randomised controlled trial
Iyengar	2013	One quarter of assessed health workers were primarily facility-based
Jintaganont	1988	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Joos	2016	Performance interventions in both arms
Jung	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Kalyango	2012	Intervention not designed to improve performance
Morris	2015	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Owais	2011	Tested the efficacy of an intervention for mothers rather than efficacy of an intervention to improve CHW performance
Pineda	1983	Not a randomised controlled trial
Pradeep	2014	Non-CHW comparator
Puett	2012	Not a randomised controlled trial
Sranacharoen-pong	2012	Intervention not designed to improve performance
Turab	2014	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Zafar	2016	Intervention not designed to improve performance
Zurovac	2011	Health workers were primarily facility based

Characteristics of studies awaiting classification

Author	Year	Title
Carrón	1993	Desarrollando instrumentos de bajo costo para mejorar la calidad de los servicios rurales de planificación familiar: informe final (Enero 1992 - Diciembre 1993).
Uwima	2013	Community-based intervention to enhance provision of integrated TB-HIV and PMTCT services in South Africa.

Characteristics of ongoing studies

Hackett 2015

Trial Registration	N.S.
Trial name	Job satisfaction, self-efficacy, and performance of community health workers participating in a mobile health (mHealth) programme to improve maternal, newborn and child health (MNCH) in Rural Tanzania

Objective	Evaluate the process and impact of an mHealth intervention
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: CHW pairs
Country	Tanzania
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: CHWs were trained on a smartphone application designed to improve data management, patient tracking, and delivery of key health/nutrition messages to pregnant women and mothers.
Control	Ministry of Health's national community- MNCH program.
Outcomes	PATIENTS: Satisfaction CHWs: Performance
Starting date (dd/mm/yyyy)	N.S.
Contact information	Kristy Hackett PhD Candidate. University of Toronto Toronto, ON. M5T 3M7. Canada. +1-416-978-2058 kristy.hackett@mail.utoronto.ca
Funding Source	Foreign Affairs, Trade and Development Canada (DFATD) W. Garfield Weston Doctoral Fellowship Program

Källander 2012

Trial Registration	NCT01972321
Trial name	Evaluating the effect of innovative motivation and supervision approaches on community health worker performance and retention in Uganda and Mozambique: study protocol for a randomised controlled trial
Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Sub-counties
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: The mHealth interventions in Uganda and Mozambique encompass three main activities: 1) closed user groups to enable free two-way communication between CHWs and their supervisors; 2) weekly ICCM data submission using phones with automated motivational feedback, SMS to supervisors flagging problems for target supervision, and summary ICCM statistics made accessible online to district statisticians; and 3) monthly motivational

	and constructive SMS to CHWs.
Control	Children living in the control areas continued to benefit from the routine Ministry of Health iCCM package provided by the CHWs who were supported by the national and sub-national health services with funding from Malaria Consortium's CIDA-iCCM project.
Outcomes	PATIENTS: Proportion of sick children appropriately treated. CHWs: (1) Performance, (2) Motivation
Starting date (dd/mm/yyyy)	11/2009
Contact information	Karin Källander, PhD +256772744126 k.kallander@malariaconsortium.org
Funding source	Bill & Melinda Gates Foundation and DFID Programme Partnership Agreement (DFID-PPA)

Rabbani 2015

Trial registration	ACTRN12613001261707
Trial name	Improving community case management of diarrhoea and pneumonia in district Badin, Pakistan through a cluster randomised study
Objective	Evaluate the affect of new supervisory strategies employed by Lady Health Supervisors
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Lady health supervisors
Country	Pakistan
Population	LOCAL TERM FOR CHW: Lady health workers CHWs: Women. Age and education level not reported. PATIENTS: Children. Sex and education level not reported.
Interventions	INTERVENTION DESCRIPTION: The intervention consists of training to build LHS knowledge and skills, clinical mentorship and written feedback to LHWs KEY CHW COMPETENCY: Community case management
Control	Standard care
Outcomes	PATIENTS: None CHWs: Improvement in CCM practices of childhood diarrhoea and pneumonia
Starting date (dd/mm/yyyy)	11/2014
Contact information	Fauziah Rabbani fauziah.rabbani@aku.edu Department of Community Health Sciences, The Aga Khan University, Stadium Road, 3500, Karachi 74800, Pakistan
Funding source	WHO Geneva, Department of Maternal, Newborn, Child and Adolescent Health.

Saha 2013

Trial registration	N.S.
---------------------------	------

Trial name (registration)	A Comprehensive Training And Supervision Program Improved Frontline Health Workers' Knowledge Of Infant And Young Child Feeding In Rural Bangladesh
Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Not reported
Country	Bangladesh
Population	LOCAL TERM FOR CHW: Shasthya Sebika & Pushti Kormi CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: Education, Incentives, Supervision KEY CHW COMPETENCY: A comprehensive infant and young child feeding practices training and performance improvement package (3-day classroom discussion, 3-day field practice, monthly refreshers, routine supervision and performance-based incentives) plus a mass media campaign.
Control	Mass media campaign
Outcomes	PATIENTS: None CHWs: CHW knowledge of infant and young child feeding practices
Starting date (dd/mm/yyyy)	2010
Contact information	Kuntal Kumar Saha, International Food Policy Research Institute (IFPRI) kuntalkumar38@gmail.com
Funding source	Bill and Melinda Gates Foundation

Vedanthan 2014

Trial registration	NCT01844596
Trial name	Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study): study protocol for a randomized controlled trial
Objective	Investigate whether community health workers, equipped with a tailored behavioural communication strategy and smartphone technology, can increase linkage and retention of hypertensive individuals to a hypertension care programme and significantly reduce blood pressure among them.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Community
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: 1) community health

	workers with an additional tailored behavioural communication strategy; and 2) community health workers with a tailored behavioural communication strategy who are also equipped with smartphone technology. KEY CHW COMPETENCY: Motivational interviewing and tailored communication- strategies
Control	Usual care (community health workers with the standard level of hypertension care training)
Outcomes	PATIENTS: The co-primary outcome measures are: 1) linkage to hypertension care, and 2) one-year change in systolic blood pressure among hypertensive individuals. Cost-effectiveness analysis will be conducted in terms of costs per unit decrease in blood pressure and costs per disability-adjusted life year gained. CHWs: None.
Starting date (dd/mm/yyyy)	04/2014
Contact information	Rajesh Vedanthan rajesh.vedanthan@mssm.edu Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1030, 10029 New York, USA
Funding source	National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 1U01HL114200, United States, under The Global Alliance for Chronic Diseases programme

Quality of included economic evaluations

Item		Carrón 1994	Chang 2011	DeRenzi 2012/ABC
1	Is the study population clearly described?	+	+	+
2	Are competing alternatives clearly described?	+	+	+
3	Is a well-defined research question posed in answerable form?	-	-	-
4	Is the economic study design appropriate to the stated objective?			
5	Is the chosen time horizon appropriate to include relevant costs and consequences?			
6	Is the actual perspective chosen appropriate?	+	+	+
7	Are all important and relevant costs for each alternative identified?	+	+	-

8	Are all costs measured appropriately in physical units?	+	+	+
9	Are costs valued appropriately?	+	+	+
10	Are all important and relevant outcomes for each alternative identified?			
11	Are all outcomes measured appropriately?			
12	Are outcomes valued appropriately?			
13	Is an incremental analysis of costs and outcomes of alternatives performed?			
14	Are all future costs and outcomes discounted appropriately?	-	-	-
15	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	-	-	-
16	Do the conclusions follow from the data reported?	+	+	+
17	Does the study discuss the generalizability of the results to other settings and patient/ client groups?			
18	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	+	+	-
19	Are ethical and distributional issues discussed appropriately?	-	-	-

Appendix 6e: Differences Between Protocol and Review

Cost effectiveness was added as an outcome for reasons explained in-text.