

Tackling poorly selected, collected, and reported outcomes in obstetrics and gynecology research.

James M. N. Duffy MBChB MRes ¹

Sue Ziebland DPhil MSc ¹

Peter von Dadelszen FRCOG DPhil ²

Richard J. McManus FRCGP FRCP ¹

¹ Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, United Kingdom.

² Department of Women and Children's Health, School of Life Course Sciences, King's College London, London SE1 7EU, United Kingdom.

Conflicts of interest:

Richard McManus has received blood pressure monitors for research from Omron. The remaining authors no conflict of interests.

Sources of funding:

This report is arising from a doctoral fellowship (DRF-2014-07-051) supported by the National Institute for Health Research. Richard McManus was supported by a National Institute for Health Research Professorship (NIHR-RP-R2-12-015) and the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care Oxford. The funders had no role in the commentary.

Correspondence to:

Dr James M. N. Duffy MBChB MRes BSc (Hons) PG HCL

Nuffield Department of Primary Care Health Sciences, University of Oxford

Oxford OX2 6GG

United Kingdom

james.duffy@balliol.ox.ac.uk

@jamesmnduffy

Word count: 1,370 words

Abstract

Clinical research should ultimately improve patient care. To enable this, randomized controlled trials must select, collect, and report outcomes which are both relevant to clinical practice and genuinely reflect the perspectives of key stakeholders including healthcare professionals, researchers, and patients. Unfortunately, many randomized controlled trials fall short of this requirement. Complex issues, including a failure to take into account the perspectives of key stakeholders when selecting outcomes, variations in outcome definitions and measurement instruments, and outcome reporting bias make research evidence difficult to interpret, undermining the translation of research into clinical practice. Problems with poor outcome selection, measurement, and reporting can be addressed by developing, disseminating, and implementing core outcome sets. A core outcome set represents a minimum data set of outcomes developed using robust consensus science methods engaging diverse stakeholders including healthcare professionals, researchers, and patients. Core outcomes should be routinely utilized by researchers, collected in a standardized manner, and reported consistently in the final publication. They are currently being developed across our specialty including infertility, endometriosis, and preeclampsia. Recognizing poorly-selected, -collected, and -reported outcomes as serious hindrances to progress in our specialty, over 80 journals including the *American Journal of Obstetrics and Gynecology*, have come together to support the Core Outcomes in Women's and Newborn Health (CROWN) initiative. The consortium supports researchers to develop, disseminate, and implement core outcome sets. Implementing core outcome sets could make a profound contribution to addressing poorly-selected, -collected, and -reported outcomes. Implementation should ensure future randomized controlled trials hold the necessary reach and relevance to inform clinical practice, enhance patient care, and improve patient outcomes.

Keywords: [1] Core outcome sets; [2] Gynecology; [3] Obstetrics; [4] Outcomes; [5] Outcome measures; and [6] Research waste.

The problem:

Research should ultimately improve patient care. For this to be possible, randomized controlled trials must select, collect, and report outcomes that are relevant to clinical practice and genuinely reflect the perspectives of diverse stakeholders including professionals, researchers and patients.

The solution:

Developing, disseminating, and implementing core outcome sets should standardize outcome selection, collection, and reporting to ensure future randomized controlled trials inform clinical practice, enhances patient care, and improves patient outcomes.

Poorly-Selected: Outcomes Lacking Relevance to Clinical Practice and Key Stakeholders.

Selecting appropriate outcomes is a critical step in designing randomized trials. However, the relative importance of individual outcomes may become a secondary consideration for researchers, as they accommodate sample size requirements, costs, and time constraints.

The primary outcome should be the outcome of greatest therapeutic importance to the study's prospective hypothesis. When considering randomized trials evaluating potential treatments for infertility, there is a consensus regarding the selection of live birth as the primary outcome. Unfortunately, a recent systematic review of 142 infertility trials published in 2013 and 2014 found that only 52 included trials (37%) reported live birth.¹

Poorly-selected outcomes extend beyond the primary outcome. Researchers should select secondary outcomes which reflect the complex pathophysiology of the healthcare condition of interest. For example, when considering preeclampsia, secondary outcomes should include blood pressure thresholds, clinical symptoms, maternal morbidity, and neonatal morbidity. Many preeclampsia trials fall short of this requirement. A systematic review of 79 preeclampsia trials demonstrated widespread variation in the reporting of outcomes. A minority of randomised trials reported outcomes concerning maternal morbidity including eclampsia (39 trials, 49%), pulmonary edema (15 trials, 19%), and renal failure (10 trials, 13%).² Over the past three decades, the scope of obstetric research has widened, with long-term childhood follow-up becoming increasingly prioritized by patients, healthcare professionals, and researchers. As the importance of assessing long-term outcomes gains momentum, challenging decisions regarding the selection of long-term outcomes, follow-up duration, and methods will need to be made.

Poorly-Collected: Different Definitions and Measurement Instruments.

Evidence synthesis can be limited by variation in methods of measurement or definition, even when outcomes have been consistently collected and reported across trials. In the absence of a standardized approach, researchers can choose from a variety of different definitions and measurement instruments. For example, researchers have defined stillbirth using different combinations of gestational ages, birth weights, and crown-heel lengths. When considering endometriosis trials, dysmenorrhea has been measured by ten different measurement instruments including visual analogue scales, Likert scales, and questionnaires.³ Such variation can make it difficult to synthesize the results of individual randomized trials using secondary research, including individual patient data meta-analysis, pair-wise meta-analysis, and network meta-analysis.

Poorly-Reported: Outcome Reporting Bias.

Outcome reporting bias is defined as the selection for publication of a subset of the original recorded outcome variables on the basis of the results. Several systematic reviews have confirmed outcome reporting bias and quantified its impact when pooling data from individual trials in a meta-analysis. A systematic review of 519 randomized controlled trials identified substantial deficiencies in outcome reporting: when compared with the trial registry record, 360 trials (75%) did not fully report efficacy outcomes and 196 (64%) did not fully report harm outcomes in the final trial report.⁴ Statistically significant outcomes had higher odds of being fully reported when compared with non-significant outcomes (odds ratio (OR) = 2.0, 95% confidence interval (CI) 1.6 to 2.7). Further problems arise with outcome reporting when the results of individual trials are combined within a meta-analysis. A systematic review of 157 Cochrane systematic reviews published in 2007, identified more than a third of systematic reviews contained at least one randomized trial at high risk of outcome reporting bias.⁴ A sensitivity analysis, taking into account outcome reporting bias, demonstrated a reduction of over 20% in the treatment effect of the primary outcome.

Problems with poor outcome selection, measurement, and reporting can be addressed by developing, disseminating, and implementing a core outcome set to standardize outcome selection, collection, and reporting across future randomized controlled trials within individual healthcare conditions.

Call to Action: Developing, Disseminating, and Implementing Core Outcome Sets.

A core outcome set represents a minimum data set of outcomes developed using robust consensus science methods. Core outcomes should be consistently selected, collected in a standardised manner, and reported consistently in the final publication. The existence of a core outcome set does not imply that outcomes in a particular trial should be restricted. Rather, there is an expectation that the core outcomes will be collected and reported as a minimum with other outcomes included as appropriate for a given study. This would make it easier for the core outcomes from individual trials to be compared, contrasted and combined as appropriate. Over 50 core outcome sets are currently being developed across our specialty including infertility, endometriosis, and preeclampsia.⁵

When developing a core outcome set, a diverse range of stakeholders, including healthcare professionals, researchers, and patients, representing various disciplines, geographical regions, and methodological expertise should be recruited. A high number of varied participants is desirable to secure the generalizability of the final core outcome set and increase its credibility with other researchers.

The first step in core outcome set development is to develop a long list of potential core outcomes by undertaking a systematic review of published randomized controlled trials (Figure 1). Outcomes identified through systematic reviews of published trials largely reflect

outcomes healthcare professionals and researchers have considered important to collect and measure, particularly where research pre-dates the recent emphasis on patient and public involvement. Outcomes reported in historic trial reports may not hold the same relevance for other stakeholders including patients, and, therefore, the evidence base to support shared decision making is likely to be suboptimal to inform clinical practice. Core outcome set developers should use qualitative methods to provide reassurance that the long list of potential core outcomes includes those with relevance to patients. Unfortunately, only a few core outcome set developers have used in-depth patient interviews to identify relevant outcomes.⁵ Future core outcome set developers should engage with methodological research to evaluate different qualitative methods, including secondary analysis of archive interviews, focus groups, and free text questionnaires, to identify treatment outcomes that hold relevance to patients.

The next step is to reduce the long list of potential core outcomes to a core outcome set using consensus science methods. Core outcome set developers have used the modified Delphi method enabling individuals to participate in an iterative process which assesses the extent of agreement (consensus measurement) and then resolves disagreement (consensus development).⁵ Potential core outcomes have been entered into sequential online surveys through which participants score the importance of individual outcomes. Repeated reflection and rescoring encourages convergence towards consensus 'core' outcomes.

The final step is to determine how the core outcomes should be defined and measured. Potential definitions and measurement instruments have been inventoried across formal definition development initiatives, international and national guidelines, Cochrane systematic reviews, and randomized controlled trials.⁵ Such definitions can be entered into a consensus development workshop to prioritize standardised definitions for individual core outcomes. Potential measurement instruments have been quality assessed using the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative

quality assessment framework in order to associate high-quality measurement instruments with individual core outcomes.⁴

When a core outcome set has been developed it should be regularly reviewed to confirm its ongoing validity. Future core outcome set developers should carefully design a strategy to ensure a core outcome set remains fit for purpose. There is uncertainty regarding the frequency of reviews and whether specific triggers should prompt a review. Methods including assessments of the uptake of an individual core outcome set, interviews with the intended users of the core outcome set, and monitoring outcome reporting in future trials could provide a framework when considering review triggers.

Core Outcome Sets Supported by Women's Health Journals.

Embedding core outcome sets within randomized controlled trials could make a profound contribution to advancing the usefulness of research to inform clinical practice, enhance patient care, and improve patient outcomes. Recognizing that poorly-selected, -collected, and -reported outcomes are a serious hindrance to progress in women's health, over 80 journals, including the *American Journal of Obstetrics and Gynecology*, have come together to support the Core Outcomes in Women's and Newborn Health (CROWN) initiative. The consortium supports researchers to develop, disseminate, and implement core outcome sets (Figure 2). Implementing the aims of the CROWN initiative across participating journals should encourage researchers committed to creating harmony in outcome selection, collection, and reporting. Over time, research consumers including healthcare professionals, researchers, and patients, should become accustomed to reviewing complete core outcome sets in research reports. As a result, the temptation for selective reporting based on statistical significance should be substantially reduced. Implementation of core outcome sets is supported by other key stakeholders including opinion leaders, health research funders, and health research regulators. For example, the Standard Protocol Items:

Recommendations for Interventional Trials (SPIRIT) statement, implemented by funders of health research including the National Institutes of Health, European Commission, and the National Institute of Health Research, recommend the use of core outcome sets wherever possible.⁴ Developing collaborative relationships with other stakeholders would be useful, for example, clinical trial registries, such as clinicaltrials.gov, could mandate the prespecified collection of individual core outcome sets.

Conclusions

Developing, disseminating, and implementing core outcome sets could make a profound contribution to addressing poorly-selected, -collected, and -reported outcomes. Implementation should ensure that future randomised controlled trials hold the reach and relevance to inform clinical practice, enhance patient care, and improve patient outcomes.

References

1. Wilkinson J, Roberts SA, Showell M, Brison DR, Vail A. No common denominator: a review of outcome measures in IVF RCTs. *Human Reproduction*. 2016;31(12):2714-2722.
2. Duffy JMN, Hirsch M, Kawsar A, et al. Outcome reporting across randomised controlled trials evaluating therapeutic interventions for pre-eclampsia. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2017;124(12):1829-1839.
3. Hirsch M, Duffy JMN, Kuszniir JO, et al. Variation in outcome reporting in endometriosis trials: a systematic review. *American Journal of Obstetrics and Gynecology*. 2016;214(4):452-464.
4. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: Version 1.0. *Trials*. 2017;18(S3):280.
5. Duffy JMN, Rolph R, Gale C, et al. Core outcome sets in women's and newborn health: A systematic review. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2017;124(10):1481-1489.