Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis

F.A. Achana[1], N.J. Cooper[1], S. Dias[2], G. Lu[2], S. J.C. Rice[3], D. Kendrick[4], A.J. Sutton[1]

[1]*Department of Health Sciences, University of Leicester, Leicester LE1 7RH, U.K.*

[2]*School of Social & Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK*

[3]*Centre for Reviews and Dissemination, University of York, York, YO10 5DD, UK*

[4]*Division of Primary Care Research, Floor 13, Tower Building, University Park, Nottingham, NG7 2RD, UK*

Correspondence to: F.A. Achana, Department of Health Sciences, University of Leicester, Leicester LE1 7RH, U.K.

Email:faa7@le.ac.uk

Funding:

This study presents independent research commissioned by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10231). The views expressed are those of the author(s) and are not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health, United Kingdom

Acknowledgements:
The authors would like to thank Tony Ades for useful discussions relating to the statistical aspects of the research presented. They would also like to thank Catriona McDaid, Emma Maund and Nerys Woolacott for their contribution to the effectiveness of non-opioid analgesics dataset.

ABSTRACT

Adjusting for the baseline risk is of interest in meta-analysis as a potential source of heterogeneity because it is a proxy for unmeasured but important patient-level characteristics which may be modifiers of treatment effect. This is problematic when the observed event rate in the placebo arm is taken as the baseline risk measure because of the measurement error in the covariate. If this error is not accounted for, a biased estimate of the relationship could be obtained. Models to address this problem have been presented in the literature for pairwise meta-analysis. Our objective is to extend these methods to network meta-analysis (NMA) where it is of interest to adjust for baseline imbalances in the non intervention group event rate in order to reduce heterogeneity and possibly inconsistency. This objective is complicated in NMA by this covariate being sometimes missing, due to the fact that not all studies in a network may have a non-active intervention arm. A random effects meta-regression model allowing for inclusion of multi-arm trial and trials without a 'non intervention' arm is developed. Analyses are conducted within a Bayesian framework using the WinBUGS Software. The method is illustrated using two examples: i) Interventions to promote functional smoke alarm ownership by households with children and ii) Analgesics to reduce post-operative morphine consumption following a major surgery. The results showed no evidence of baseline effect in the smoke alarm example but the analgesics example shows the adjustment can greatly reduce heterogeneity and improve overall model fit.

Key words: network meta-analysis; mixed treatment comparison; baseline risk; underlying risk; MCMC; meta-regression

1 INTRODUCTION

In meta-analyses of clinical trials, differences in patient or trial/study level characteristics often give rise to variation in treatment effect estimates between studies also called heterogeneity [3]. Between study variance in the treatment effects is usually taken into account through including a parameter for the residual heterogeneity in a random effects meta-analysis [3, 5]. A random effects model quantifies the degree of heterogeneity but does not explain it. To explain the source of the heterogeneity, patient and study level characteristics are sometimes included in the analysis as covariates [3, 5]. A trial-level covariate of interest as a possible source of heterogeneity is the 'baseline risk' or the underlying risk of the disease. The baseline risk reflects the burden of disease in a study population and defines the average risk of a patient to experience the outcome of interest if they have not been treated [7]. It is potentially an important proxy for a number of unmeasured (and even measured) patient-level characteristics such as age, sex, medical history and disease severity that collectively influence a patient's response to treatment [6]. In addition to heterogeneity, baseline imbalances between trials may also give rise to inconsistency (i.e. variability in the treatment effect between pair-wise contrasts [9] in a network meta-analysis (NMA)). Therefore, adjusting for it may have the benefit of reducing both heterogeneity and inconsistency in NMA and improve the overall model fit.

Various measures have been used for the baseline risk in meta-analyses. Examples include the observed event rate in the placebo or *non-active intervention arm*, the observed placebo arm log odds and the average of the observed event rates in the placebo and treatment arms [1, 11, 12]. However, including observed measures of baseline risk in a meta-regression can be problematic because of the measurement error in both response (i.e. treatment effect) and explanatory variables and functional relationship between the two [6]. The problem has received considerable attention in the literature with several authors proposing alternative

model based solutions. Examples include the methods of McIntosh [2], Walter [1], Thomson et al. [6], Sharp and Thomson [8], Arends et al. [10] and van Houwelingen et al. [4]. However, these methods are directly applicable mainly in pairwise meta-analysis. Our objective is to extend these methods to network meta-analysis (NMA) where it may be of interest to adjust for baseline imbalance in the underlying event risk across studies. The main reason for doing this would be reduce between-study heterogeneity and possible inconsistency in the direct and indirect trial evidence on pairwise comparisons. This objective is complicated by missing data, due to the fact that not all studies in a network may have a placebo or *non-active treatment* control, and thus an observed covariate value.

A review of the pair-wise meta-analysis methods for investigating the relationship between treatment effect and baseline risk is presented in Section 2. In Section 3, we present an approach which primarily extends the methods of Thompson et al. [6, 8] and Arends et al. [10] from pair-wise to NMA where it is of interest to adjust for the baseline risk. The method we present compliments previous general multivariate meta-regression models suggested for NMA (see Cooper et al. [9], Salanti et al. [13, 14], Dias et al. [15]and Stijnen et al. [16]) by allowing for i) alternative distributional assumptions to be made about the nature of the true unobserved baseline risk measure, and ii) the inclusion of trials without a *non-active treatment* control and hence no baseline risk measure whilst allowing for the $treatment \times covariate$ interactions to be exchangeable or even different (i.e. as many regression coefficients as there are treatment effects). Section 4 presents application of the method to two recently published NMAs [17, 18]. The first example has a binary outcome and examines effectiveness of home safety education interventions to promote ownership of functional smoke alarm in households with children [17]. The second example has a continuous outcome measure and examines the effectiveness of analgesic treatments in reducing post-operative morphine consumption in adult

patients following major surgery [18]. In Section 5, we discuss the results and findings from the example datasets followed by the strengths and limitations of the approach outline in this paper.

2 REVIEW OF BASELINE RISK MODELS FOR PAIR-WISE META-ANALYSIS

Sharp and Thomson [8] and Arends et al. [10] both present good introductions to baseline risk adjustment and detailed review of available methods for pair-wise meta-analysis. We have summarised important features of six of the methods that we consider most relevant to our modelling approach for NMA in table 1. A common feature in these methods is to model the relationship of interest in three parts, although this was only stated explicitly in Arends et al. [10]. This involves specifying in any order i) an appropriate likelihood for the data, ii) a regression model relating the true treatment effect as explanatory variable and the true baseline risk as the covariate and iii) a model for the distribution of the baseline risk across studies [10].

Differences between approaches have mostly arisen from slightly different strategies adopted for each part of the model. For example, Thompson et al. [6] Arends et al. [10] and Sharp and Thompson [8] assumed a binomial likelihood for a binary outcome whereas McIntosh [2], Walter [1] and van Houwelingen et al. [4] used normal distribution to model a binary outcome measure (e.g. log odds or log odds ratio). Approximating a log odds ratio with a normal distribution can be mathematically and computationally convenient but the normality assumption may be inappropriate if there are trials in the meta-analysis with zero or small numbers of events [8]. Secondly, except for the method of Walter [1], all the other methods assumed a random study-specific effects. Walter's [1] model is fixed effect in that no allowance is made for any residual heterogeneity other than that explained by the baseline risk, although we believe expecting residual heterogeneity is more realistic in most applications where it is of interest to adjust for the baseline risk.

Finally, the approaches outlined in table 1 make different assumptions about the distribution of the true unobserved baseline risk across studies [2, 19]. Whilst some models assumed a vague or minimally informative normal prior distribution (e.g. Thompson et al. [6], Sharp and Thompson [8] and also in Arends et al. [10]); a common parametric formulation is to assume that the baseline risk is normally distributed across trials as in McIntosh [2], van Houwelingen et al. [4] and also Arends et al. [10]. Additionally, Arends et al. [10] also proposed a more flexible model for the distribution of the baseline risk comprising a mixture of two normal distributions with different means but common between-study variance. Whether or not to assume a model for the baseline risk is a much debated issue with as yet no clear consensus among methodologists [4, 8, 10, 20]. More recently, Ghidey et al. [19, 21] proposed semi-parametric models for the distribution of the baseline risk as well as models that do not make any distributional assumptions. In the next section, we aim to develop methods for the baseline risk adjustment in NMA that incorporate the different assumptions about the baseline risk distribution across studies in order to assess the effect of these assumptions on parameter estimates.

3 NETWORK META-ANALYSIS WITH BASELINE RISK COVARIATE

3.1 Network meta-analysis model with no covariate adjustment

Suppose in a meta-analysis of $i = 1, 2, \ldots, N$ trials, we have $k = A, B, C, \cdots, NT$ treatments being compared with one another where NT is the total number of treatments. Take treatment A as the overall baseline or reference treatment of the entire network. For a binary outcome, we assume r_{ik} events occur out of n_{ik} patients in treatment arm k of trial i according to a binomial distribution with underlying event probability p_{ik}. Standard random effects NMA for a binary outcome with no covariate can be specified using logistic regression fellows [22, 23]:

$$r_{ik} \sim Binomial(p_{ik}, n_{ik}) \text{ with } \theta_{ik} = logit(p_{ik});$$

$$\theta_{ik} = \begin{cases} \mu_{ib} & k = b; \quad b \in \{A, B, C, \cdots\} \\ \mu_{ib} + \delta_{ibk} & k > b; \quad b \in \{A, B, C, \cdots\} \end{cases} \qquad \textbf{Equation (1)}$$

$$\delta_{ibk} = d_{bk} + \varepsilon_{ibk} \text{ with } \varepsilon_{ibk} \sim N(0, \sigma_{bk}^2)$$

Note: $d_{AA} = 0$, $k > b$ implies treatment *k* comes alphabetically after *b*

where θ_{ik} is a continuous measure of the treatment effect in arm *k* of trial *i* (log odds in the case of a binary outcome), μ_{ib} is the effect of baseline treatment *b* (log odds) in trial *i* and ε_{ibk} denote a random effect indicating that the trial-specific effects (log odds ratios) of treatment *k* relative to *b*, δ_{ibk} , are normally distributed with mean d_{bk} and between-study variance σ_{bk}^2. The fundamental assumption underlying random effects network meta-analysis is that the treatments effects are exchangeable across the entire network of trials regardless of whether or not treatments *b* and *k* are included in trial *i* [24]. Validity of this assumption means that the pooled treatment effects, d_{bk}, can further be expressed as functions of basic parameters taken with reference to treatment *A*, (i.e. $d_{bk} = d_{Ak} - d_{Ab}$) [25]. Effect estimates from trials with more than 2 treatment groups will be correlated through sharing a common comparator treatment. The correlation is taken into account by assuming homogenous variances ($\sigma_{bk}^2 = \sigma^2$) so that the covariance is equal to $\frac{\sigma^2}{2}$ (Lu and Ades 2004 [23]). Alternatively, heterogeneous variance models have also been proposed (see Lu and Ades, 2007 [26]). Modelling is conducted within the framework of Bayesian analysis using Markov chain Monte Carlo simulation through the WinBUGS software [27] with minimally informative prior distributions specified for d_{Ak}, μ_{ib} and σ.

3.2 Extending the network meta-analysis to include a covariate for the baseline risk

Using the true but unobserved non-active control or placebo group log-odds, μ_{iA} (i.e. for *b* = *treatment A*) in trial *i* as a measure of the baseline risk, the trial-specific treatment effects in equation (1) can be made to depend on the baseline risk through the following regression:

$$\delta_{ibk} = d_{bk} + \beta_{bk}(\mu_{iA} - \bar{\mu}) + \varepsilon_{ibk};\ \varepsilon_{ibk} \sim N(0,\ \sigma_{bk}^2) \qquad \textbf{Equation (2)}$$

Note: d_{AA}, $\beta_{AA} = 0$

where δ_{ibk} and σ_{bk}^2 are defined as in equation (1), d_{bk} is mean effect of treatment *k* relative to baseline treatment *b* adjusted for the baseline risk and β_{bk} is the change in the log odds ratio of an event per unit change in the baseline risk for treatment *k* relative to *b* at the mean baseline risk across trials. We centred the baseline risk covariate on $\bar{\mu}$, the observed mean log odds in the non-active control group (*treatment A*) to improve convergence of the model (Draper et al, 1998, [28]; page 27). For trials with an active treatment control (baseline treatment $b \neq A$), we make use of the substitution $d_{bk} = d_{Ak} - d_{Ab}$ under consistency of evidence arising from the exchangeability assumption [25] to express equation (2) as:

$$\delta_{ibk} = (d_{Ak} - d_{Ab}) + (\beta_{Ak} - \beta_{Ak}) \times (\mu_{iA} - \bar{\mu}) + \varepsilon_{ibk}$$
$$\varepsilon_{ibk} \sim N(0,\ \sigma_{bk}^2) \qquad \textbf{Equation (3)}$$

All variables in equation (3) have the same interpretation as in the previous equations. Although treatment *A* is not actually included in trial *i* of equation (3), the fundamental assumption on exchangeability means that treatment arms can be assumed to be missing at random without loss to efficacy [25, 29]. This allows us to imagine that there would still be a baseline risk in trials without treatment *A* and thus, borrow strength from other trials. Therefore, no new parameters are needed for including for example a B *versus* C trial, and all other aspects of the model will remain the same. For multi-arm trials, the model takes the form of a multivariate regression to accommodate the within-study correlations between effect estimates arising from such trials. The multivariate form of equation (2) with bold characters denoting vectors and matrices is given by:

$$\boldsymbol{\delta_i} = \boldsymbol{x_i\beta} + \boldsymbol{\varepsilon_i};\quad \boldsymbol{\varepsilon_i} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}) \qquad \textbf{Equation (4)}$$

where $\boldsymbol{\delta_i}$ with elements $\delta_{i,1}, \delta_{i,2}, \cdots, \delta_{i,NT_i-1}$ for trial *i* is now a vector of relative effects (e.g. log odds ratios), NT_{i-1} is the total number of treatment effects in trial *i*, $\boldsymbol{\varepsilon_i}$ is a vector of random effects associated with trial *i* and $\boldsymbol{\Sigma}$ is a variance-covariance matrix (as defined for the network in equation (5) below). The design matrix $\boldsymbol{x_i}$ contain the covariate information with entries indicating the treatment effects being estimated in trial *i* and $\boldsymbol{\beta}$ is a vector of regression coefficients including the intercept and slope terms [30]. Following Salanti et al[13] as an example, consider a network of 4 trials with three treatments A, B and C in which trial 1 is AB (i.e. A versus B), trial 2 is AC, trial 3 is ABC and trial 4 is BC (i.e. no non-active control). With treatment A taken as the overall baseline treatment, assuming homogenous variances (i.e. $\sigma_{bk}^2 = \sigma^2$), equation (4) can be written in full for this network as:

$$\begin{pmatrix} \delta_{1,AB} \\ \delta_{2,AC} \\ \delta_{3,AB} \\ \delta_{3,AC} \\ \delta_{4,BC} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \mu_{1A} - \bar{\mu} & 0 \\ 0 & 1 & 0 & \mu_{2A} - \bar{\mu} \\ 1 & 0 & \mu_{3A} - \bar{\mu} & 0 \\ 0 & 1 & 0 & \mu_{3A} - \bar{\mu} \\ -1 & 1 & -(\mu_{4A} - \bar{\mu}) & \mu_{4A} - \bar{\mu} \end{pmatrix} \begin{pmatrix} d_{AB} \\ d_{AC} \\ \beta_{AB} \\ \beta_{AC} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,AB} \\ \varepsilon_{2,AC} \\ \varepsilon_{3,AB} \\ \varepsilon_{3,AC} \\ \varepsilon_{4,BC} \end{pmatrix}; \qquad \textbf{Equation (5)}$$

$$\begin{pmatrix} \varepsilon_{1,AB} \\ \varepsilon_{2,AC} \\ \varepsilon_{3,AB} \\ \varepsilon_{3,AC} \\ \varepsilon_{4,BC} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & \sigma^2/2 & 0 \\ 0 & 0 & \sigma^2/2 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 \end{pmatrix} \right)$$

where $\boldsymbol{\beta} = (d_{AB} \;\; d_{AC} \;\; \beta_{AB} \;\; \beta_{AC})^T$ is the 4×1 matrix of regression coefficients representing the pooled effects of treatments B and C relative to treatment A and the effect of baseline risk on treatment effect estimates. All that remains is to specify models for the distribution of the true baseline risk across trials and distribution of the regression coefficients. These are presented in the next section.

3.3 Models for the baseline risk and $\boldsymbol{treatment \times covariate}$ interactions

As stated in the review of previous model (section 2), there is no consensus in the literature about what form of distribution the baseline risk should take. We therefore follow the example

in Arends et al. [10] to specify models based on three different assumptions about the distribution of the baseline risk as follows:

1. Model 1 assumes that baseline risk is independent or unconstrained so that each trial has its own baseline risk measure. This is equivalent to specifying a vague normal prior distribution for the baseline risk across trials; $\mu_{i,A} \sim N(0,\ 10^3)$
2. Model 2 assumes that the baseline risk across trials is drawn from a normal distribution with common mean and between-study variance; $\mu_{i,A} \sim N(\bar{\mu},\ \sigma_{\mu}^2)$. Prior distributions are specified for $\bar{\mu}$ and σ_{μ}; $\bar{\mu} \sim N(0,\ 10^3)$ and $\sigma_{\mu} \sim Uniform(0\ ,10)$
3. Model 3 assumes the baseline risk is drawn from a mixture of two normal distributions but a common between-study variance; $\mu_{i,A} \sim p_1 N(\bar{\mu}_1,\ \sigma_{\mu}^2) + (1 - p_1) N(\bar{\mu}_2, \sigma_{\mu}^2)$ with prior distributions: $\bar{\mu}_1,\ \bar{\mu}_2 \sim N(0,\ 10^3)$; $\sigma_{\mu} \sim Uniform(0\ ,10)$; $p_1 \sim ddirch(\alpha_c = 1); c = 1,2$.

Similar to the models for the distribution of the baseline risk, we also specify models for the interaction terms based on the following three assumptions described in Cooper et al. [9]:

- **A.** Common effect $treatment \times\ covariate$ interactions; $\beta_{AK} = B;\ B \sim N(0,\ 10^3)$
- **B.** Exchangeable $treatment \times\ covariate$ interactions; $\beta_{AK} \sim N(B,\ \sigma_B^2);\ B \sim N(0,\ 10^3)$ and $\sigma_B \sim Uniform(0, 10)$ and
- **C.** Independent or unrelated $treatment \times\ covariate$ interactions; $\beta_{AK} \sim N(0,\ 10^3)$

Therefore, a total of 9 models are fitted based on the combination of assumptions about distribution of the baseline risk and the $treatment \times\ covariate$ interactions (slopes):

Model A1: Unconstrained baseline risk and common slope

Model A2: Normal distribution for baseline risk and common slope

Model A3: Mixture distribution for baseline risk and common slope

Model B1: Unconstrained baseline risk and exchangeable slopes

Model B2: Normal distribution for risk and exchangeable slopes

Model B3: Mixture distribution for baseline risk and exchangeable slopes

Model C1: Unconstrained baseline risk and independent slopes

Model C2: Normal distribution for baseline risk and independent slopes

Model C3: Mixture distribution for baseline risk and independent slopes

3.4 Goodness of fit and model selection

In the applications which follow, adequacy of model fit to the data was assessed through the residual deviance where a model is judged to adequately fit the data if the residual deviance closely matches the actual number of unconstrained data points available [9, 22]. The overall goodness of fit and model selection criteria were based on the Deviance Information Criteria (DIC), a measure of model fit that penalises model complexity, where the model with the lowest DIC is generally preferred [31] and differences of 3 or 5 are considered significant [27].

4 APPLICATION EXAMPLES

4.1 Example 1: Functional Smoke Alarm (FSA) Data

The data come from a published NMA [17] and consists of 20 randomised and non-randomised studies that evaluated the effectiveness of home safety education to increase ownership of functioning smoke alarm (FSA) systems in households with children. The outcome of interest is whether or not a household had a FSA. Thus, each study supplied arm level data on the number of households with a FSA and the total number of households surveyed. The FSA data is used here to illustrate application of the method to binary outcome

data. The full data is displayed in Appendix 1 Table B1 with Figure 1A displaying a network diagram for the 7 interventions and 40 data points from the 20 studies. The baseline or non intervention arm is the usual care intervention. Seven of the 20 studies did not have a usual care intervention and therefore no baseline risk covariate. [Note: In the data coding step, these 7 studies are included by using NA to represent missing information on the number of events in the usual care arm (see web appendix)]. Baseline functioning smoke alarm ownership in the remaining 13 studies ranged from about 3% to about 96%. Because a number of the studies are non-randomised, it should be noted that the effect of baseline risk may be a composition of many unmeasured sources of heterogeneity between the studies. Evidence of significant inconsistency was also detected [17] using the method of node-splitting [25]. Hence it is of interest to know whether baseline differences in FSA ownership across studies can explain the heterogeneity and inconsistency. For this example where the outcome is binary, a binomial likelihood was assumed for the arm-level data and NMA without covariate adjustment (model 0) fitted based on the model defined by equation (1). The relationship between intervention effect and baseline FSA ownership was then investigated using the methods described in Section 3.2. The covariate was centred on the observed mean baseline log odds of 0.81(calculated outside WinBUGS) for FSA ownership in the 13 studies with a usual care arm. In total, 10 models were fitted (the 9 models described in Section 3.3 in addition to the unadjusted model) using Markov Chain Monte Carlo (MCMC) simulation in the WinBUGS software [27]. A modified version of the NMA code from Dias et al.[15] was used and is given in the accompanying web appendix. The following prior distributions were used and intended to be minimally informative:

$$\sigma, \sigma_{\mu} \text{ and } \sigma_B \sim Uniform(0, 5)$$

$$\beta_{Ak}, B, d_{Ak}, \mu_{ib}, \bar{\mu} \sim N(0, \ 10^3)$$

Models were run for 100 000 iterations, discarding the first 30 000 iterations as burn-in samples in order to ensure convergence of the MCMC sampler. There was evidence of poor convergence for the models that assumed separate/independent $treatment \times covariate$ interactions (model C1, model C2 and model C3). This may be due to i) 7 out of 20 studies not having a usual care intervention arm and ii) hence, relatively few data points compared to the number of parameters needed to be estimated [9]. Therefore parameter estimates from models C1 to C3 are not presented in the results in section 4.3.1.

4.2 Example 2: Pain Relief Data

The second dataset consists of 56 RCTs with 116 data points from a published Health Technology Assessment (HTA) report [18]. The report examined effectiveness of 3 non-opioid analgesics (paracetamol, NSAIDs or COX-2 inhibitors) and placebo in reducing morphine consumption following major surgery in adults. The outcome of interest is the amount of morphine in milligrams (mg) consumed over a 24 hour period (continuous outcome). Each study provided arm-level information on the number of patients together with the mean 24 hour morphine consumption and its standard deviation (SD). The treatment network is given in Figure 1B. The dataset is presented in Table A2 of Appendix 1, in order of increasing mean morphine consumption in the placebo group (baseline risk). Two trials have no placebo and 4 of the trials are 3-arm studies. There is considerable variability in the 24 hour morphine consumption in the placebo arm of trials ranging from a low of 8.6 mg (SD 5.2 mg) to a high of 142 mg (SD 80mg). The average across the placebo group is 45.26 mg. Therefore, a sensitivity analysis was conducted in the original report [18] to investigate the effect of this baseline imbalance in morphine use on the treatment effects estimates. To include the two studies that did not have placebo, the original analysis in the published report was first carried out without these studies in order to derive an estimate of baseline morphine consumption for the two trials. The derived estimates were then included in the sensitivity analysis that adjusted for the baseline morphine use. In our analysis however, we make use of exchangeability assumption

mentioned earlier (see equation 5) in order to include the trials without a placebo arm and thus baseline risk measure. [Note: In the data coding step, these 2 studies are included by using NA to represent missing mean morphine consumption and 1 for its standard error (see web appendix)].

Since 24 hour morphine consumption is a continuous outcome, we replace the binomial likelihood and logistic regression model in equation (1) with a normal distribution for the observed arm-specific outcome (i.e. mean 24 hour morphine), $\hat{y}_{ik}$ in treatment arm *k* of trial *i*:

$$\hat{y}_{ik} \sim N(\theta_{ik}, S_{ik}^2), \qquad i = 1, 2, \cdots, 56 \text{ and } k = 1, 2, \cdots, 4$$

$$\theta_{ik} = \begin{cases} \mu_{ib} & k = b; \quad b \in \{1,2,3,4\} \\ \mu_{ib} + \delta_{ibk} & k > b; \quad b \in \{1,2,3,4\} \end{cases} \qquad \textbf{Equation (6)}$$

$$\delta_{ibk} = d_{bk} + \beta_{bk} \times (\mu_{i1} - \bar{\mu}) + \varepsilon_{ibk} \text{ with } \varepsilon_{ibk} \sim N(0, \sigma^2)$$

where θ_{ik} is the true unobserved mean morphine consumption in treatment arm *k* of trial *i* with variance, S_{ik}^2 assumed known but estimated from the data [4]. We centred the baseline morphine consumption, μ_{i1} on 45.26 mg, the average consumption across the placebo arms. All other aspects of the modelling assumptions and model fit remain the same as in example 1 except for the minimally informative prior distributions specified as follows:

$$\sigma, \sigma_\mu \text{ and } \sigma_B \sim Uniform(0, 30)$$

$$\beta_{Ak}, B, d_{Ak}, \mu_{ib}, \bar{\mu} \sim N(0, \ 10^6)$$

The MCMC simulations were run using WinBUGS for 50 000 iterations, discarding the first 20 000 iterations as burn-in samples in order to ensure convergence. The results are presented in section 4.3.2 below.

4.3 RESULTS

4.3.1 Example 1: FSA Model

In this example, the log odds ratio was regressed on the *true* control group log odds (usual care intervention) taken as a measure of baseline risk. Table 2 displays estimates of the residual heterogeneity σ, $treatment \times covariate$ interactions (regression slopes) and model fit statistics excluding the three models (C1, C2 and C3) which showed evidence of non convergence. Firstly, different assumptions about the distribution of the baseline risk did not seem to greatly affect estimates of the $treatment \times covariate$ interaction terms in this case. The slope of the regression lines are slightly steeper when minimally informative prior distribution were assumed for the baseline risk (models A1 and B1) than in models that assumed a normal baseline distribution (model A2) or a mixture of two normal distributions (model B2). Secondly, the posterior credible intervals for the slope terms included zero in all models indicating that none of these are statistically significant. Therefore, baseline imbalance in smoke alarm distribution across studies was not significantly related to effectiveness of home safety education to promote FSA ownership in households with children (provided this analysis is powered appropriately for effects under investigation). Consequently the heterogeneity and also the inconsistency were not significantly reduced in all models that adjusted for the baseline risk compared with the unadjusted model (table 2).

Using the posterior mean residual deviance as a measure of model fit to the data (table 2), both adjusted and unadjusted models predicted values close to the 40 unconstrained data points in the FSA data, indicating that these models fit the data equally well. Since baseline risk appears to be unrelated to intervention effect, there was very little difference to choose between these models and we report only the results from the common slope or $treatment \times covariate$ interaction models for convenience. Posterior median estimates of the slope is -0.08 (95% Credible interval (CrI); -0.41 to 0.28) from model A1, -0.03 (95% CrI; -0.41 to 0.35) from

model A2 and -0.03 (95% CrI; -0.39 to 0.34) from model A3 which all indicate non-significant decrease in intervention effectiveness with increasing baseline FSA ownership.

4.3.2 Example 2: Pain Relief Model

For pain relief data, the treatment effect, expressed as the mean difference in 24 hour morphine use was regressed on true but unobserved 24 hour mean morphine consumption in the placebo group (taken as baseline risk measure). There were no problems with convergence of the MCMC simulations and all 9 models described in Section 3.3 were fitted in addition to the unadjusted model. Parameter estimates of interest and model fit statistics are presented in Table 3. Firstly, estimates of the regression slopes from the 9 adjusted models were all negative, suggesting evidence of increasing treatment effect with increasing baseline morphine consumption. Estimate of the common regression slope is -0.34 (95% CrI; -0.41 to -0.27) for unconstrained baseline model (model A1), and -0.31 (95% CrI; -0.38 to -0.23) for models with normal (model A2) and mixture of two normal distributions (model A3) for baseline risk. Similar estimates of the relationship between treatment effect and baseline risk were also obtained from the independent and exchangeable slope models but only the effect of NSAIDS and COX-2 are statistically significant. Again, the three modelling assumptions about the distribution of the baseline risk seem to have very little impact on $treatment \times covariate$ interactions. Figure 2 plots treatment effects versus baseline 24-hour morphine use from the model with independent/separate slopes (model C1) for paracetamol, NSAIDS and COX-2. The plot shows: i) evidence of increasing effectiveness with increasing baseline morphine use for all three classes of analgesics; ii) NSAIDS and COX-2 are increasingly more effective than

paracetamol at higher baseline morphine use, and iii) little difference between NSAIDS and COX-2. The vertical distance between the line of no effect and each treatment regression line gives an estimate of the treatment effect relative to placebo at a given baseline morphine consumption. Similarly, the relative effectiveness of any two analgesics at a given baseline morphine consumption can be obtained from the plot as the vertical distance between the two regression lines. Secondly, adjusting for the baseline risk reduced the residual heterogeneity and improves the overall model fit. From Table 3, the posterior mean estimate of the residual heterogeneity σ is 5.44 mg (95% CrI; 4.5 to 5.98) in the unadjusted model and 3.48mg (95% CrI; 3.24 to 4.57) in model C1, the adjusted model with the least reduction in heterogeneity. Compared to the unadjusted models there is at least a 40% reduction in between-study heterogeneity.

5 DISCUSSION

We have shown how methods for baseline risk covariate adjustment can be extended from pair-wise to network meta-analysis (NMA) when it is of interest to account for differences in underlying risk across trial populations. This type of analysis can help identify potential treatment effect modifiers which may give rise to heterogeneity in effect estimates and/or inconsistency in the direct and indirect evidence on pair-wise contrasts in a network of trials. The pain relief example shows how adjusting for baseline risk can greatly reduce heterogeneity and improve overall model fit. Similar results and conclusions have been reported before, for example, by Lu et al. [32] in a NMA at multiple follow-up times where the baseline effects

were adjusted at different follow-up points. However, there was no evidence of baseline effect in the FSA example and the inconsistencies identified in Cooper et al. [17] were not resolved by baseline risk adjustment. In meta-analyses of studies evaluating complex and or public health interventions such as the FSA data, interventions may not always be clearly defined and studies are often of variable quality, and conducted in populations with different characteristics. These factors can introduce heterogeneity in both meta-analyses of clinical trials and studies of non-complex and or public health interventions, but the problem is more pronounced in public health. The FSA network included both RCTs and non randomised observational studies both of which are of variable quality. Although care was taken to categorise the interventions appropriately, 'lumping' of interventions within categories could not be completely ruled out[17]. Lumping of interventions creates relative contrasts that are unevenly distributed across contrast and has been cited as a possible source of heterogeneity and inconsistency in NMA [33] .

The main advantage of the approach described in this paper is that the models can be easily implemented by making simple modifications to freely available WINBUGS code for NMA [22] (see code in the accompanying web appendix file). Specifying the models in WinBUGS, and analysing them using Markov Chain Monte Carlo simulation, is beneficial as it allows the adjustment to be carried out without excluding trials with missing placebo or no treatment control group (hence no baseline risk covariate). The imputation step is implemented automatically in WinBUGS through the model jointly specified by the likelihood and prior distribution placed on the 'baseline risk' (described section 3.3). Since parameters are

considered as random variables within the Bayesian framework requiring a distribution [34], the 'missing covariate' is treated as any other unknown parameter to be estimated under exchangeability (see Mason 2009[35], page. 117). Alternatively, the analysis can also be carried outside a Bayesian framework using multivariate meta-analysis methods (for example Stijnen et al.[16]) fitted in standard statistical software or self written programs. However, validity of the results obtained from either classical or Bayesian analyses will depend on appropriateness of the assumption that the non-active intervention arm of studies without a baseline risk are missing at random. Fitting models with separate and or exchangeable regression slopes described in Cooper et al. [9], in addition to the common slope model can be useful for assessing the appropriateness of these assumptions. For example, the common slope assumption can be tested by first calculating the difference between estimates of any two slopes in the separate slope model followed by a *probability* that this difference is greater than zero using the step function in WinBUGS [27]. A two-sided P-value can then be derived using the formula $2 \times minimum(probability, 1 - probability)$ [25]. However, as shown by the FSA example, fitting models with separate/independent slopes may not always be feasible, possibly because of limited availability of data. In those circumstances, the exchangeable slope or even common slope models can be considered as a compromise [9]. Under the exchangeable regression slope assumption, power is improved by borrowing strength across regression slopes which shrinks treatment effect estimates towards each other. This can have policy implications especially in a decision making context where manufacturers of alternative interventions may see the effectiveness of their products "shrink" towards that of the competitor. Also the exchangeable slope assumption can reduce heterogeneity in the effect estimates (σ), but the

regression slopes themselves can be quite variable as illustrated by the pain relief example where the σ_B's are more variable than σ. This shows that the regression slopes are much more variable and therefore a common regression coefficient may not be the best model for this example.

Finally going back to the review of pair-wise meta-analysis models presented in section 2, a much debated issue in modelling the relationship between treatment effect and baseline risk has been whether or not to assume a parametric distribution for the baseline risk and what form if any such a distribution should take. Ghidey et al.[19] examined the issue in a recently published methods review paper using real and simulated data for pair-wise meta-analysis. The simulated results found no difference between models that assumed normality for the baseline risk and those that did not with both models producing robust/unbiased estimates of the regression slope when the baseline risk is normally distributed across studies [19]. However, the estimate of the regression slope was found to be less biased under the functional modelling approach when normality of the baselines was violated but the relative difference in bias was small. The results from the approach outlined in this paper for network meta-analysis appear consistent with the findings from Ghidey et al.[19] and also with Arends et al [10]. Estimates of regression slopes from both FSA and pain relief examples were slightly less negative, and tended to shrink towards zero in models that assumed normally distributed baselines (models A2, A3, B2, B3 in table 2 and models A2-A3, B2-B3 and C2-C3 in table 3) compared to the unconstrained or minimally informative prior distributions for the baseline risk (models A1-

C1). The effect of different distributional assumptions about the baseline risk were however, very minimal as both unconstrained and normally distributed baseline risk models produced practically identifiable estimates of regression slopes.

6 REFERENCE

1. Walter SD. Variation in baseline risk as an explanation of heterogeneity in meta-analysis. *Stat Med* 1997; **16**: 2883-2900.
2. McIntosh MW. The population risk as an explanatory variable in research synthesis of clinical trials. *Stat Med* 1996; **15**: 1713-1728.
3. Sutton A, Abrams K, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research.* Wiley: Chichester, 2000.
4. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002; **21**: 589-624.
5. Borenstein M, Hedges LV, J.P.T. H, Rothstein HR. *Introduction to Meta-Analysis.* Wiley 2009, 2009.
6. Thompson SG, Smith TC, Sharp SJ. Investigating underlying risk as a source of heterogeneity in meta-analysis. *Stat Med* 1997; **16**: 2741-2758.
7. Higgins JPT, Green SE. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* The Cochrane Collaboration 2011.
8. Sharp SJ, Thompson SG. Analysing the relationship between treatment effect and underlying risk in meta-analysis: comparison and development of approaches. *Stat Med* 2000; **19**: 3251-3274.
9. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med* 2009; **28**: 1861-1881.
10. Arends LR, Hoes AW, Lubsen J, Grobbee DE, Stijnen T. Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Stat Med* 2000; **19**: 3497-3518.
11. Brand R, Kragt H. Importance of trends in the interpretation of an overall odds ratio in the meta-analysis of clinical trials. *Stat Med* 1992; **11**: 2077-2082.
12. Sharp SJ, Thompson SG, Altman DG. The relation between treatment benefit and underlying risk in meta-analysis. *BMJ* 1996; **313**: 735-738.
13. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; **17**: 279-301.
14. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *J Clin Epidemiol* 2009; **62**: 857-864.
15. Dias S, Sutton AJ, Welton NJ, Ades AE. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. Available from http://www.nicedsu.org.uk. 2011.
16. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine* 2010; **29**: 3046-3067.
17. Cooper NJ, Kendrick D, Achana F, Dhiman P, He Z, Wynn P, Le Cozannet E, Saramago P, Sutton AJ. Network Meta-analysis to Evaluate the Effectiveness of Interventions to Increase the Uptake of Smoke Alarms. *Epidemiologic Reviews* 2011.
18. McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N. Paracetamol and selective and non-selective nonsteroidal anti-inflammatory drugs (NSAIDs) for the reduction of morphine-related side effects after major surgery:a systematic review. *Health Technology Assessment* 2010; **14**: 1-+.

19. Ghidey W, Stijnen T, van Houwelingen HC. Modelling the effect of baseline risk in meta-analysis: A review from the perspective of errors-in-variables regression. *Statistical Methods in Medical Research* 2011.
20. van Houwelingen H, Senn S. Investigating underlying risk as a source of heterogeneity in meta-analysis. *Stat Med* 1999; **18**: 110-115.
21. Ghidey W, Lesaffre E, Stijnen T. Semi-parametric modelling of the distribution of the baseline risk in meta-analysis. *Statistics in Medicine* 2007; **26**: 5434-5444.
22. Dias S, Welton NJ, Sutton AJ, Ades AE. In, Editor (ed)^(eds). City, 2011.
23. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; **23**: 3105-3124.
24. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; **331**: 897-900.
25. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **29**: 932-944.
26. Lu G, Ades AE. Assessing Evidence Inconsistency in Mixed Treatment Comparisons. *Journal of the American Statistical Association* 2006; **101**: 447-459.
27. Spiegelhalter D, Thomas AB, N. , D. L. WinBUGS User Manual Version 1.4 January 2003. Upgraded to Version 1.4.3. 2007.
28. Draper NR, Smith H. *Applied Regression Analysis.* (3rd edn): London, 1998.
29. Lu GB, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009; **10**: 792-805.
30. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Stat Med* 1998; **17**: 2537-2550.
31. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society (B)* 2002; **64**: 1-34.
32. Lu G, Ades AE, Sutton AJ, Cooper NJ, BriggS AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Statistics in Medicine* 2007; **26**: 3681-3699.
33. Caldwell DM, Welton NJ, Ades AE. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *Journal of Clinical Epidemiology* 2010; **63**: 875-882.
34. Ntzoufras I. *Bayesian Modelling Using WinBUGS.* John Wiley & Sons: New Jersey, 2009.
35. Mason AJ. Bayesian methods for modelling non-random missing data mechanisms in longitudinal studies. In Bayesian methods for modelling non-random missing data mechanisms in longitudinal studies, Editor (ed)^(eds). Imperial College: City, 2009.

Table 1: Summary of methods for modelling the relationship between treatment effect and baseline risk in pair-wise meta-analysis with a binary outcome

Method	Outcome data	Likelihood model	Distribution of baseline risk	Method of estimation	Further notes
Method 1 (RE): Walter (1997) [1]	Arm level	**Two normal distributions** (Observed treatment and control group log-odds with normal errors	None	ML or WSL with bias correction	Gives only fixed effects results as no allowance for excess heterogeneity. Narrow standard errors for regression slope β
Method 2 (RE): McIntosh (1996) [2]	Trial level	**Bivariate normal (BVN) approximation: -** Log-OR & control group log-odds assumed bivariate normal with known variance and covariance-matrix estimated from data)	Normal (RE on baseline risk)	ML & Bayesian	BVN assumption may be inappropriate if there are trials with small number of events. May result in more extreme estimates of slope β and lower estimates of between-study heterogeneity, σ^2 Normality of baseline risk across trials may be hard to justify
Method 3 (RE): van Houwelingen et al. (2002) [4]	Arm level	**Bivariate normal (BVN) approximation** for binary outcome data (treatment & Control group log-odds *or* observed treatment effect & control group log-odds assumed BVN with known covariance-matrix estimated from data)	Normal (RE on baseline risk)	EM algorithm & SAS Proc Mix	
Method 4a (RE): Thompson et al. (1997 [6], 2000[8]) ; Arends et al. (2000) [10]	Arm level	**Exact binomial model** (Observed number of events in each treatment-arm assumed binomial)	Fixed, flat prior	Bayesian	Eliminates need for zero-cell corrections Useful situations where trials with small sampled sizes are included in the meta-analysis
Method 4b (RE): Arends et al. (2000) [10]	Arm level	**Exact binomial model** (Observed number of events in each treatment-arm assumed binomial)	Normal (RE)	Bayesian	
Method 4c (RE): Arends et al. (2000)[10]	Arm level	**Exact binomial model** (Observed number of events in each treatment-arm assumed binomial)	Mixture of two normal distributions	Bayesian	Eliminates need for zero-cell corrections Useful situations where trials with small sampled sizes are included in the meta-analysis Flexible distributional assumptions for the baseline risk measure

ML = Maximum Likelihood, WLS = Weighted least square, Log-OR = Log-odds ratio; MA=Meta-analysis, FE=Fixed effects, RE=Random effects

Table 2: NMA with baseline risk adjustment applied to functional smoke alarm data: - Estimates of residual heterogeneity σ and regression coefficients; β_k of intervention k relative to usual care (control intervention) measuring the relationship between intervention effect (log-odds ratio) and baseline risk (control group log-odds)

	Model 0: Unadjusted model	Model A1: Unconstrained baseline; common slope	Model A2: Baseline normally distributed; common slope	Model A3: Baseline mixture of two normal; common slope	Model B1: Unconstrained baseline; Exchangeable slopes	Model B2: Baseline normally distributed; exchangeable slopes	Model B3:Mixture model; Exchangeable slopes
Regression slopes	Median (95% CrI)	Median (95% CrI)	Median (95% CrI)	Median (95% CrI)	Median (95% CrI)	Median (95% CrI)	Median (95% CrI)
Common β	-	-0.08 (-0.41, 0.28)	-0.03 (-0.41, 0.35)	-0.03 (-0.39, 0.34)	-	-	-
Education (β_2)					-0.13 (-0.42, 0.22)	-0.09 (-0.39, 0.25)	-0.09 (-0.40, 0.27)
Education + Equipment (β_3)					0.19 (-0.59, 1.38)	0.26 (-0.57, 1.54)	0.25 (-0.57, 1.50)
Education + Equipment + HIS (β_4)					-1.08 (-2.75, 0.201)	-1.20 (-2.81, 0.163)	-1.16 (-2.63, 0.28)
Education + Equipment + Fitting (β_5)					0.26 (-0.32, 1.05)	0.35 (-0.25, 1.20)	0.35 (-0.28, 1.22)
Education + HIS (β_6)					-0.07 (-3.34, 3.02)	-0.07 (-3.45, 3.08)	-0.08 (-3.55, 2.54)
Education + Equipment + Fitting + HIS (β_7)					0.09 (-1.69, 2.48)	0.19 (-2.01, 2.66)	0.165 (-1.98, 2.38)
Mean random effects β	-	-	-	-	-0.09 (-1.55, 1.28)	-0.07 (-1.58, 1.42)	-0.07 (-1.63, 1.30)
Residual heterogeneity, σ	0.77 (0.34,1.47)	0.83 (0.39, 1.56)	0.84 (0.40, 1.59)	0.84 (0.40, 1.59)	0.59 (0.16, 1.35)	0.57 (0.167, 1.30)	0.59 (0.14, 1.38)
SD for random effects β	-	-	-	-	0.88 (0.07, 3.31)	1.02 (0.12, 3.45)	0.97 (0.09, 3.41)
Model fit statistics							
Residual deviance ($\bar{D}$)	41.72	41.49	40.86	40.99	40.33	39.98	40.04
Effective number of parameters (pD)	35.28	35.85	35.70	36.65	35.53	34.97	35.091
Deviance information criteria (DIC)	77.00	77.34	76.56		75.86	74.95	75.131

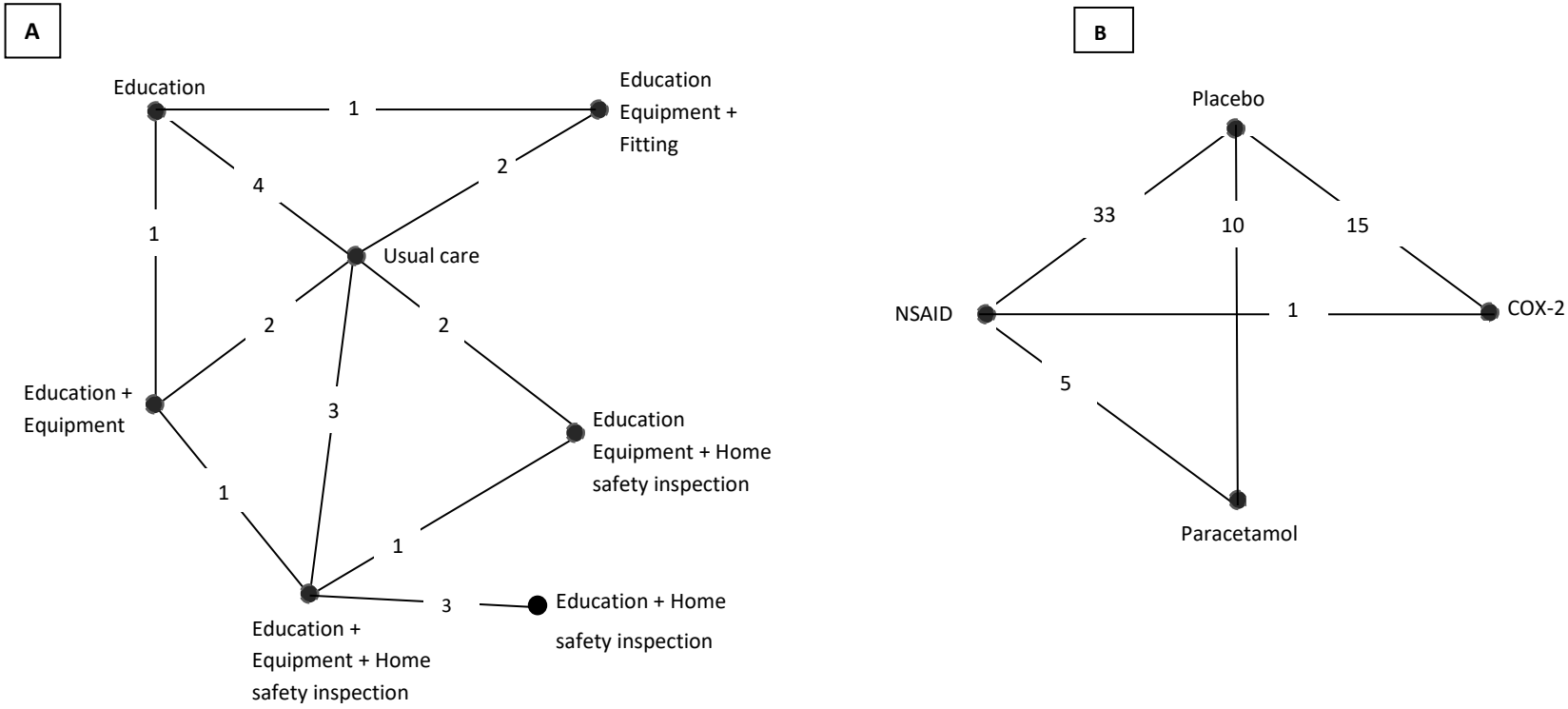
β_2= interaction term for Education relative to Usual care; SD = standard deviation in treatment effect estimate; CrI = Credible interval

Table 3: NMA with baseline risk adjustment applied to pain relief data: - Estimates of the residual heterogeneity, σ , and the regression coefficients; β_k of intervention *k* relative to usual care (control intervention) measuring the relationship between treatment effect (mean difference in 24 hour morphine use in mg) and placebo group morphine use in mg

	Model 0: Unadjusted model	Model A1: Unconstrained baseline; common slope	Model A2: Baseline normally distributed; common slope	Model A3: Baseline mixture of two normal; common slope	Model B1: Unconstrained baseline; exchangeable slopes	Model B2: Baseline normally distributed; exchangeable slopes	Model B3: Mixture model; exchangeable slopes	Model C1: Unconstrained baseline; separate slopes	Model C2: Baseline normally distributed; separate slopes	Model C3: Mixture model; separate slopes
Regression slopes	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)	Mean (95% CrI)
Common β		-0.34 (-0.41, -0.27)	-0.31 (-0.38, -0.23)	-0.31 (-0.38, -0.23)	-	-	-	-	-	-
Paracetamol (β_2)					-0.22 (-0.39, 0.003)	-0.18 (-0.35, 0.04)	-0.19 (-0.36, 0.01)	-0.16 (-0.36, 0.04)	-0.15 (-0.34, 0.06)	-0.13 (-0.34, 0.10)
NSAIDS (β_3)					-0.36 (-0.44, -0.28)	-0.34 (-0.42, -0.25)	-0.33 (-0.42, -0.25)	-0.36 (-0.45, -0.28)	-0.34 (-0.43, -0.26)	-0.34 (-0.43, -0.26)
COX-2 (β_4)					-0.34 (-0.450, -0.18)	-0.27 (-0.43, -0.09)	-0.27 (-0.42, -0.09)	-0.35 (-0.53, -0.19)	-0.25 (-0.44, -0.06)	-0.26 (-0.44, -0.05)
Random effects mean β					-0.30 (-0.93, 0.34)	-0.26 (-1.00, 0.47)	-0.27 (-0.98, 0.48)			
Residual heterogeneity, σ	5.44 (4.50, 5.98)	3.19 (2.15, 4.47)	3.19 (2.14, 4.51)	3.22 (2.15, 4.56)	3.20 (2.15, 4.49)	3.13 (2.06, 4.50)	3.13 (2.04, 4.50)	3.28 (2.20, 4.57)	3.16 (2.06, 4.51)	3.20 (2.06, 4.53)
SD for random effects β	-	-	-	-	0.35 (0.01, 2.18)	0.39 (0.01, 2.37)	0.36 (0.01, 2.32)	-	-	-
Model fit statistics										
Residual deviance ($\bar{D}$)	124	119.5	121.90	121.1	117.60	121.10	120.30	116.40	120.60	119.70
Effective number of parameters (*pD*)	90.63	84.11	81.97	82.31	85.27	82.58	82.57	85.56	82.58	82.96
Deviance information criteria (*DIC*)	214.63	202.61	203.87	203.41	202.87	2	202.87	201.964	203.18	202.66

Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis" by F.A Achana et al.

Figure 1: Intervention network for A) Possession of Functional Smoke Alarm and B) Post-operative pain relief



Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis" by F.A Achana et al.

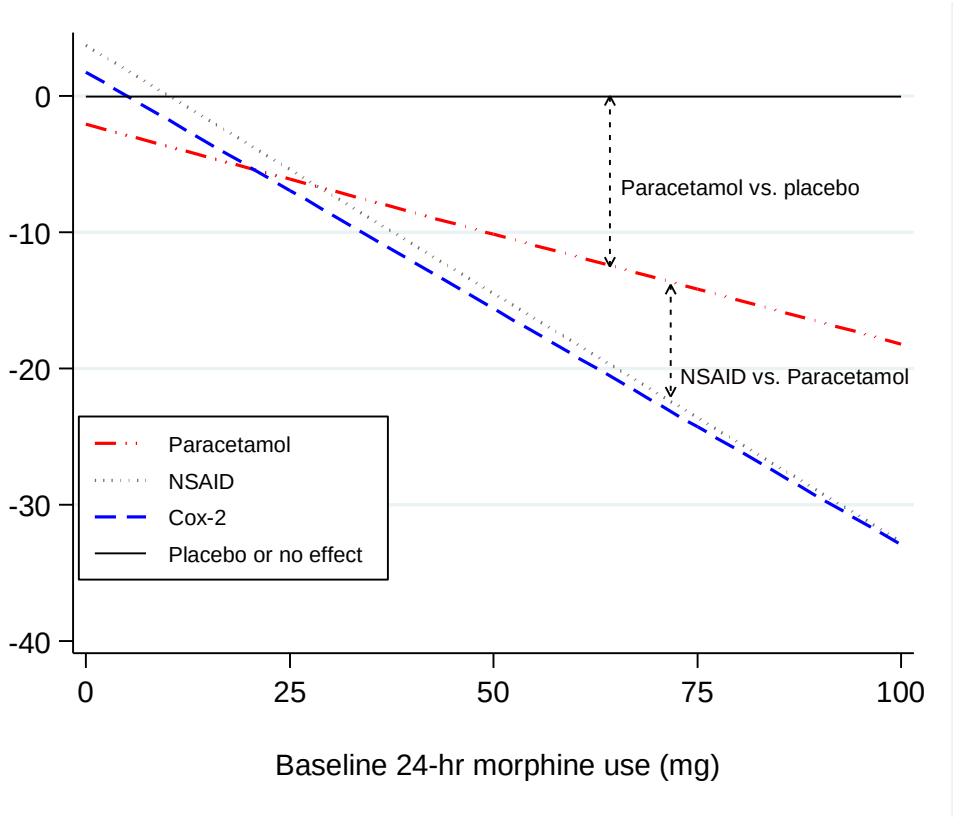


Figure 2: NMA adjusting for baseline morphine use assuming independent slopes for different treatment effects and unconstrained baseline risk