

Original citation:

Achana, Felix A., Cooper, Nicola S., Bujkiewicz, Sylwia, Hubbard, Stephanie J., Kendrick, Denise, Jones, David R and Sutton, Alex J. (2014) Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. BMC Medical Research Methodology, 14 (1). p. 92.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/78528>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 2.0 Generic (CC BY 2.0) license and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/2.0/>

A note on versions:

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

RESEARCH ARTICLE

Open Access

Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes

Felix A Achana^{1*}, Nicola J Cooper¹, Sylwia Bujkiewicz¹, Stephanie J Hubbard¹, Denise Kendrick², David R Jones¹ and Alex J Sutton¹

Abstract

Background: Network meta-analysis (NMA) enables simultaneous comparison of multiple treatments while preserving randomisation. When summarising evidence to inform an economic evaluation, it is important that the analysis accurately reflects the dependency structure within the data, as correlations between outcomes may have implication for estimating the net benefit associated with treatment. A multivariate NMA offers a framework for evaluating multiple treatments across multiple outcome measures while accounting for the correlation structure between outcomes.

Methods: The standard NMA model is extended to multiple outcome settings in two stages. In the first stage, information is borrowed across outcomes as well across studies through modelling the within-study and between-study correlation structure. In the second stage, we make use of the additional assumption that intervention effects are exchangeable between outcomes to predict effect estimates for all outcomes, including effect estimates on outcomes where evidence is either sparse or the treatment had not been considered by any one of the studies included in the analysis. We apply the methods to binary outcome data from a systematic review evaluating the effectiveness of nine home safety interventions on uptake of three poisoning prevention practices (safe storage of medicines, safe storage of other household products, and possession of poison centre control telephone number) in households with children. Analyses are conducted in WinBUGS using Markov Chain Monte Carlo (MCMC) simulations.

Results: Univariate and the first stage multivariate models produced broadly similar point estimates of intervention effects but the uncertainty around the multivariate estimates varied depending on the prior distribution specified for the between-study covariance structure. The second stage multivariate analyses produced more precise effect estimates while enabling intervention effects to be predicted for all outcomes, including intervention effects on outcomes not directly considered by the studies included in the analysis.

Conclusions: Accounting for the dependency between outcomes in a multivariate meta-analysis may or may not improve the precision of effect estimates from a network meta-analysis compared to analysing each outcome separately.

Keywords: Network meta-analysis, Mixed treatment comparisons, Multiple outcomes, Multivariate, WinBUGS

Background

Meta-analysis or the quantitative synthesis of evidence, usually from systematic reviews, has become a popular tool in healthcare evaluations [1,2]. Largely driven by a desire for more realistic synthesis of complex healthcare evidence, increasingly sophisticated methodology has been developed.

One area of meta-analysis that has seen significant methodological development is the application of multivariate statistical methods for the comparison of treatments on two or more endpoints (usually known as multivariate meta-analysis) [3-8]. These methods are appealing because many studies and systematic reviews focus on broad health effects and therefore typically report several outcome measures [4,6,9]. In such instances, the multivariate approach offers some advantages over separate univariate analyses including the ability to account for the inter-relationship between

* Correspondence: faa7@le.ac.uk

¹Biostatistics Group, Department of Health Sciences, University of Leicester, University Road, Leicester LE1 7RH, UK

Full list of author information is available at the end of the article

outcomes and borrow strength across studies as well as across outcomes [10] through modelling the correlation structure [7,11]. This can potentially reduce outcome reporting bias [12] and the uncertainty with which intervention effects are estimated. Additionally, in a decision making context where the synthesis is meant to inform a health economic evaluation, accounting for the correlations between effect estimates on different outcomes is important as the dependence between outcomes may have implication for estimating quality of life or economic consequences associated with treatment [13]. An example is the situation where a particularly effective treatment for a disease condition is associated with a large side effect profile. Ignoring information about the inter-relationships between beneficial and 'side effect' endpoints in such instances may have implications for quantifying the benefits associated with treatment.

When summarising effectiveness evidence, correlations between the effectiveness estimates typically arise at either within-study and or between-study levels. At the within-study level, correlations arise mainly due to differences in patient-level characteristics. They are rarely reported in the published literature and usually have to be estimated from external sources such as individual patient level data if available or elicited from expert opinion [8,11,14,15]. At the between-study level, correlations arise from i) differences in the distribution of patient-level characteristics across studies, in which case they will be related to the within-study correlations and/or ii) differences in the distribution of other study-level characteristics such as study design, population and baseline disease severity [16]. The within-study correlations thus give an indication of the association between multiple endpoints within a study while the between-study correlations indicate how the underlying true study-specific effects on different outcomes vary jointly across studies.

A second area of rapid methodological development is network meta-analysis (NMA) [17], also known as mixed treatment comparison meta-analysis [18-20] or multiple treatment meta-analysis [21-23]. NMA methods extend standard pairwise meta-analysis to enable simultaneous comparison of multiple treatments while maintaining randomisation of individual studies [18]. The method enables 'direct' evidence (i.e. evidence from studies directly comparing two interventions of interest) and 'indirect' evidence (i.e. evidence from studies that do not compare the two interventions directly) to be pooled under the assumption of evidence consistency [24]. Estimates of intervention effects can then be obtained, including effects between treatments not directly compared within any one individual study [19]. NMA methods thus provide a coherent framework for appraising all available evidence relevant to a specific decision problem. The results from such analyses are increasingly being used to inform economic evaluations in healthcare decision making where coherent decisions (about judicious use of scarce resource)

need to be made based on sound appraisal of all available evidence.

Approaches to extend NMA methodology to multiple outcome settings have been proposed in the literature [13,25-27], initially focusing on mutually exclusive competing risk outcomes [13] or a single outcome measured at multiple time points [26,28]. More recently, Efthimiou *et al.* [14] proposed a method for modelling multiple correlated outcomes in networks of evidence with binary outcome measures. The proposed method accounts for both the within-study and between-study correlation structure and includes a strategy for eliciting expert opinion to inform the within-study correlations. This paper contributes to the growing literature on the simultaneous evaluation of correlated outcomes. We do this in two stages. In the first stage (labelled as model 2 in the remainder of the paper), information is borrowed across studies as well as across outcomes through modelling the correlations between effectiveness estimates on different outcomes. In the second stage (labelled as model 3 in the remainder of the paper), additional information is borrowed across outcomes based on ideas for combining evidence across human and animal studies originally proposed by DuMouchel and Harris [29] and also revisited by Jones *et al.* [30]. The proposed second stage analysis methods allows: i) disconnected treatments to be incorporated as nodes in a network of evidence and ii) prediction of intervention effects for outcomes where evidence from primary studies is either sparse or not directly available from any one study included in the analysis. The motivating application area is injury prevention in children where a broad array of outcomes and intervention packages have been evaluated with the aim of increasing safety practices around the home (to ultimately reduce household injuries).

The remainder of this paper is structured as follows: the example dataset is first described followed by a Methods section describing the statistical models developed and implementation of the models. These are followed by sections presenting the results of applying the methods to the motivating dataset and a discussion.

Dataset

The example data comes from a recently updated Cochrane systematic review of home safety education and provision of safety equipment for injury prevention in children [31]. The models developed in this paper are applied to a subset of the review evidence relating to the prevention of poisoning injuries. Table 1 presents the data from 22 studies for the following outcomes:

- a) Safe storage of medicines
- b) Safe storage of other household products (e.g. cleaning products) and
- c) Possession of a poison control centre (PCC) telephone number.

Table 1 Summary of the available evidence

Comparison	First author and year of publication	IPD	Outcome information (no. of events/no. of households in control versus (vs.) treatment arm)		
			Safe storage of medicines	Safe storage of other household products	Possession of a PCC number
Usual care (1) vs. Education (2)	Gielen 2007	Yes	178/271 vs. 188/249 [†]	44/62 vs. 57/73 [†]	
	Nansel 2002	Yes	83/89 vs. 79/85	65/89 vs. 66/85	59/89 vs. 63/85
	Nansel 2008	Yes	72/74 vs. 140/144 [†]	59/73 vs. 117/144 [†]	50/59 vs. 90/119 [†]
	Kelly B 1987	No	54/54 vs. 55/55	43/54 vs. 49/55	
	McDonald 2005	No	4/57 vs. 6/60	3/57 vs. 6/61	
	Kelly N 2003	No			45.56/136.68 vs. 112.95/137.63 [*]
Usual care (1) vs. Education + free/low cost safety equipment (3)	Clamp 1998	Yes	68/82 vs. 79/83	49/82 vs. 59/83	
	Woolf 1987	No			29/143 vs. 47/119
	Woolf 1992	No		60/151 vs. 89/150	59/151 vs. 117/150
Usual care (1) vs. Education + equipment (3) vs. Education + equipment + home safety inspection (4)	Babul 2007	Yes	147/149 vs. 171/173 vs. 160/163		
Usual care (1) vs. Education + equipment + home safety inspection (4)	Hendrickson 2002	Yes		14/40 vs. 34/38	8/40 vs. 34/38
	Swart 2008	No	70.26/79.58 vs. 74.07/80 [*]	46.86/57.96 vs. 50.87/58.27 [*]	
	Kendrick 1999	Yes		317/367 vs. 322/363	
Usual care (1) vs. Education + equipment + fitting (5)	Watson 2005	Yes	683/738 vs. 712/762	327/669 vs. 368/693	
Usual care (1) vs. Education + home safety inspection (6)	Petridou 1997	No			67.26/100.12 vs. 71.08/97.83 [†]
Usual care (1) vs. Education + equipment + home safety inspection + fitting (7)	Schwarz D 1993	No	88.42/248.37 vs. 128.16/248.37 [*]		
	Phelan 2011	No			16/138 vs. 71/139
Usual care (1) vs. Home visit (8)	Johnson 2006	No			82/91 vs. 222/232 [†]
Education (2) vs. Education + equipment (3)	Posner 2004	Yes	14/47 vs. 19/49	22/47 vs. 34/49	27/47 vs. 35/49
Education (2) vs. Education + equipment + fitting (5)	Sznajder 2003	Yes	44/49 vs. 43/45	32/41 vs. 40/48	
Education + equipment + home safety inspection (4) vs. Education + equipment + home safety inspection + fitting (7)	King J 2001	No		261/469 vs. 273/482	
Education + equipment (3) vs. Equipment (9)	Dershewitz 1979	No	22/101 vs. 20/104	1/101 vs. 0/104	

Treatment abbreviation and codes:

Usual care = UC (1).

Education = E (2).

Education + free/low cost equipment = E + FE (3).

Education + equipment + home safety inspection = E + FE + HSI (4).

Education + equipment + fitting = E + FE + F (5).

Education + home safety inspection = E + HSI (6).

Education + equipment + home safety inspection + fitting = E + FE + HSI + F (7).

Education + home visit = E + HV (8).

Free/low cost equipment = FE (9).

^{*}Effective sample size reported for cluster randomised studies after adjusting clustering, hence not whole numbers (details given in Kendrick *et al.* 2012 [31]).

[†]The IPD for Gielen 2007 shows information on safe storage of medicines and safe storage of other household products was collected from different sets of households in this study (i.e. all the households that provided information for storage of medicines had missing data for safe storage of other household products and vice versa). Hence the Gielen 2007 IPD was not used to estimate the within-study correlations. [†]The intervention arms of Nansel 2008 and Johnson 2006 [32] comprises two groups that received different versions of a home safety intervention. The two versions were considered to be similar, hence combined into one intervention group for the analysis reported here.

Thirteen of the 22 studies considered at least two of the three outcomes. Of these, 8 considered storage of medicines and storage of other household products, 2 considered storage of other household products and possession of a PCC telephone number, and 3 considered all three outcome measures. Individual patient data (IPD) were available for 8 of the 13 studies, of which 7 were in a format suitable for the analysis reported here as explained by the footnotes in Table 1.

We classified the interventions trialled in the 22 studies into 9 relatively homogenous treatment packages:

- 1) Usual care (UC)
- 2) Education (E)
- 3) Education + provision of free/low cost equipment (E + FE)
- 4) Education + provision of free/low cost equipment + home safety inspection (E + FE + HSI)
- 5) Education + provision of free/low cost equipment + fitting of equipment (E + FE + F)
- 6) Education + home safety inspection (E + HSI)
- 7) Education + provision of free/low cost equipment + home safety inspection + fitting of equipment (E + FE + HSI + F)
- 8) Education + home visit (E + HV)
- 9) Provision of free/low cost equipment (FE).

- 7) Education + provision of free/low cost equipment + home safety inspection + fitting of equipment (E + FE + HSI + F)
- 8) Education + home visit (E + HV)
- 9) Provision of free/low cost equipment (FE).

Figure 1 shows the comparisons between the interventions that were made by individual studies and the number of comparisons in each network. All studies compared 2 intervention strategies, except Babul *et al.* (2007) [33] which compared 3 strategies. Data on each outcome was not available for all interventions; i.e. for the storage of medicines and other household products outcomes, interventions E + HSI and E + HV were not investigated in any of the included studies, and for possession of a PCC number interventions, E + FE + F and FE were not available.

Methods

In this section, we first present the NMA statistical model for one binary outcome measure and then extend it to

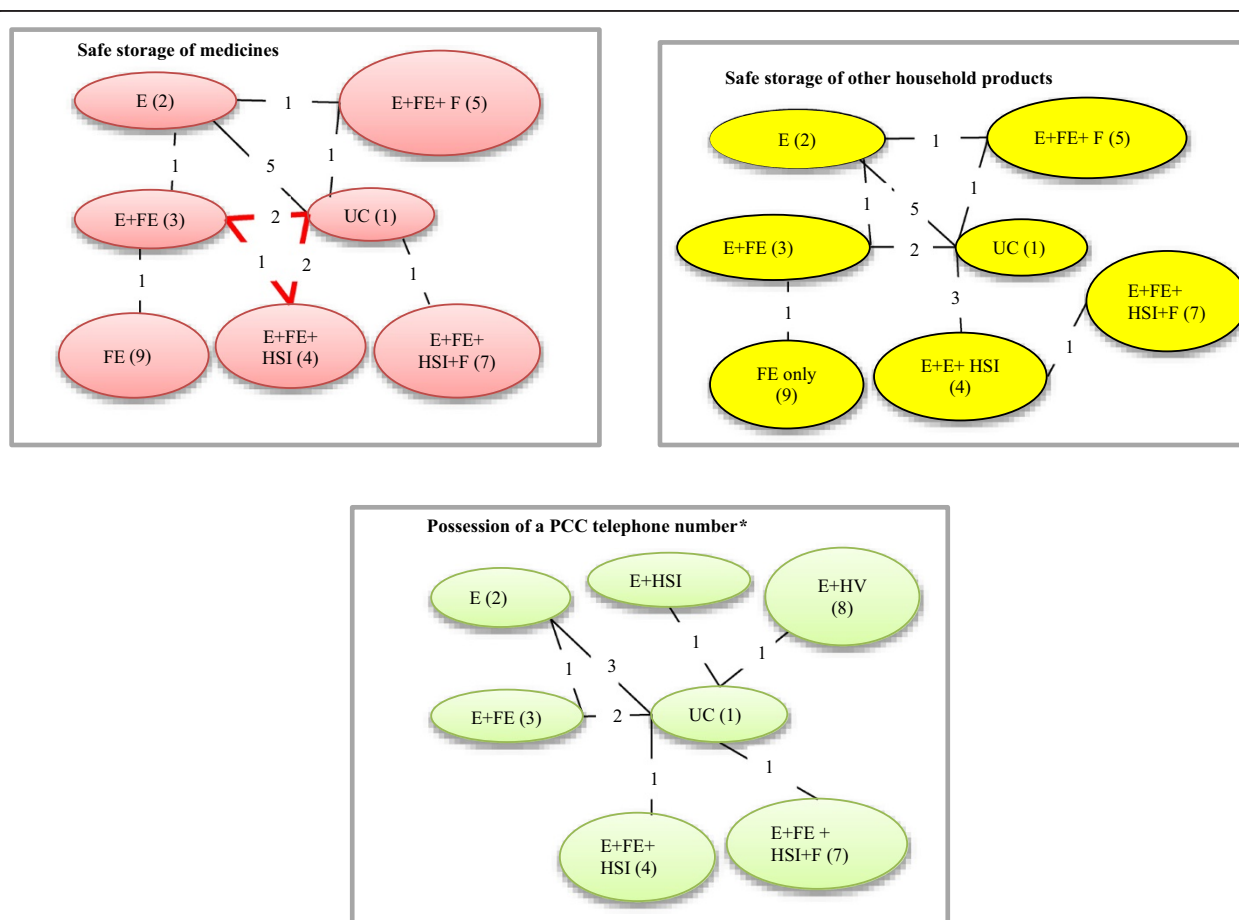


Figure 1 Intervention networks for the poisoning prevention outcomes (thick red lines indicate multi-arm comparisons). *PCC= poison control centre telephone number.

compare multiple interventions across multiple outcomes. Throughout the paper, we refer to these single and multiple outcome models as univariate and multivariate NMAs respectively. Where studies report multiple outcomes, these will not be independent as each household provides information on the different outcome measures within intervention arms. The multivariate model takes this correlation structure into account by allowing the intervention effects measured by one outcome to be correlated with the intervention effects measured by other outcomes.

Model 1: Univariate NMA

Given arm-level binary data of the form presented in Table 1, a random effects NMA may be specified using the method of Lu and Ades [20]. It is assumed that the occurrence of r_{ik} events out of a total of n_{ik} households in the k th-arm ($k = A, B, C, \dots$) of the i th-study follow a binomial distribution with underlying event probability p_{ik} :

$$\begin{aligned} r_{ik} &\sim \text{Binomial}(p_{ik}, n_{ik}); \text{logit}(p_{ik}) = \theta_{ik} \\ \theta_{ik} &= \begin{cases} \mu_{ib} & \text{if } k = b \\ \mu_{ib} + \delta_{i(bk)} & \text{if } k > b \end{cases} \quad (1) \\ b &= A, B, C, \end{aligned}$$

$$\delta_{i(bk)} \sim \text{Normal}\left(d_{(bk)} = d_{(Ak)} - d_{(Ab)}, \sigma^2_{(bk)}\right) \quad (2)$$

Note : $d_{AA} = 0$

where μ_{ib} is a study-specific baseline effect (i.e. the log-odds for the control group in study i with baseline treatment b), $\delta_{i(bk)}$ is a study-specific log-odds ratio, $d_{(bk)}$ is the pooled effect of treatment k relative to treatment b (a quantity usually of interest in a meta-analysis) and $\sigma^2_{(bk)}$ is the between-study variance or heterogeneity parameter. Random effects NMA assumes that intervention effects are exchangeable across the network regardless of whether or not treatments b and k are included in study i [18]. This assumption implies that the pooled effects $d_{(bk)}$ can be expressed as functions of basic parameters with reference to a common comparator or baseline treatment (i.e. $d_{(bk)} = d_{(Ak)} - d_{(Ab)}$) [24]. Throughout this paper, we take usual care (UC) intervention to be the reference or 'baseline' treatment (i.e. UC is taken as treatment A of equation (2) above). Multi-arm studies (i.e. studies with more than 2 treatment groups) present a special problem in network meta-analysis because they produce evidence on multiple treatment effects that are correlated through sharing a common reference or 'baseline' treatment. Under a homogeneous variance assumption ($\sigma^2_{(bk)} = \sigma^2$), the covariance between any two effects that share a common reference treatment is $\frac{\sigma^2}{2}$ [20]. The homogeneous variance assumption allows for the distribution of effects (in a study with an arbitrary number of arms) to be expressed as a univariate marginal distribution and a series of univariate conditional distributions. Specifically, for the i th-study

with $p + 1$ arms and p treatment effect estimates relative to the reference treatment, if

$$\begin{pmatrix} \delta_{i(bk_1)} \\ \delta_{i(bk_2)} \\ \vdots \\ \delta_{i(bk_p)} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} d_{(bk_1)} \\ d_{(bk_2)} \\ \vdots \\ d_{(bk_p)} \end{pmatrix}, \begin{pmatrix} \sigma^2 & \frac{\sigma^2}{2} & \dots & \frac{\sigma^2}{2} \\ \frac{\sigma^2}{2} & \sigma^2 & \dots & \frac{\sigma^2}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma^2}{2} & \frac{\sigma^2}{2} & \dots & \sigma^2 \end{pmatrix} \right) \quad (3)$$

then the marginal and conditional univariate distributions for arm j , given the previous $1, \dots, (j-1)$ arms are:

$$\begin{aligned} \delta_{i(bk_1)} &\sim \text{Normal}(d_{(bk_1)}, \sigma^2) \text{ for } j = 1. \\ \delta_{i(bk_j)} | \begin{pmatrix} \delta_{i(bk_1)} \\ \vdots \\ \delta_{i(bk_{j-1})} \end{pmatrix} &\sim \text{Normal} \left(d_{(bk_j)} + \frac{1}{j} \sum_{i=1}^{j-1} (\delta_{i(bk_i)} - d_{(bk_i)}), \frac{(j+1)}{2j} \sigma^2 \right) \\ &\text{for } j = 2, \dots, p \end{aligned} \quad (4)$$

The analysis is conducted within a Bayesian framework requiring prior distributions to be specified for all model parameters. Accordingly, we specified minimally informative prior distributions corresponding to a Normal $(0, 10^3)$ prior distribution for the pooled mean effects relative to usual care, $d_{(Ak)}$ and the study-specific baseline effects, μ_{ib} and a Uniform $(0, 2)$ prior distribution for the between-study standard deviation on the log odds ratio scale σ [34].

Model 2: Multivariate NMA

We extend the univariate NMA model defined above to the multiple outcomes settings in order to account for correlations between intervention effects on different outcomes. The method presented here follows from Ades *et al.* (2010) NMA with competing risks model [13] where only the within-study correlations are taken into account. We extend their method to account for the between-study correlation as well.

Note that in Ades *et al.* (2010), a multinomial likelihood was appropriate as the three binary outcomes (relapse during treatment for Schizophrenia, discontinuation because of intolerable side effects, and discontinuation for other reasons) are mutually exclusive and event probabilities sum to 1 across outcomes. A multinomial likelihood will not be appropriate for our example dataset because each household can have one, two or all three outcomes simultaneously so that the event probabilities do not sum to 1

across outcomes. Instead, we use a Normal distribution on the log-odds scale to take account of the within-study correlations between outcomes. We assume that in each study i and for each k -th arm, the estimates $\hat{\theta}_{ikm}$ of the observed log-odds of an event on the m th outcome ($m = 1, 2, \dots, M$) jointly follow a multivariate normal distribution:

$$\begin{pmatrix} \hat{\theta}_{ik1} \\ \vdots \\ \hat{\theta}_{ikM} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} \theta_{ik1} \\ \vdots \\ \theta_{ikM} \end{pmatrix}, \mathbf{S}_{ik} \right), \quad (5)$$

$$\mathbf{S}_{ik} = \begin{pmatrix} s_{ik1}^2 & \cdots & r_{ik}^{1M} s_{ik1} s_{ikM} \\ \vdots & \ddots & \vdots \\ s_{ikM}^2 & \cdots & s_{ikM}^2 \end{pmatrix}$$

$$\begin{pmatrix} \theta_{ik1} \\ \vdots \\ \theta_{ikM} \end{pmatrix} = \begin{cases} \begin{pmatrix} \mu_{ib1} \\ \vdots \\ \mu_{ibM} \end{pmatrix} & \text{if } k = b \\ \begin{pmatrix} \mu_{ib1} + \delta_{i(bk)1} \\ \vdots \\ \mu_{ibM} + \delta_{i(bk)M} \end{pmatrix} & \text{if } k > b \end{cases}$$

for $b = A, B, C, \dots$

where $(\mu_{ib1}, \mu_{ib2}, \dots, \mu_{ibM})$ and $(\delta_{i(bk)1}, \delta_{i(bk)2}, \dots, \delta_{i(bk)M})$ represent vectors of 'true' baseline and study-specific effects in study i with baseline treatment b respectively. The quantities $(\hat{\theta}_{ik1}, \hat{\theta}_{ik2}, \dots, \hat{\theta}_{ikM})$ and $(\theta_{ik1}, \theta_{ik2}, \dots, \theta_{ikM})$ represent vectors of observed and 'true' log-odds of response in arm k of study i and \mathbf{S}_{ik} is the associated within-study covariance matrix usually assumed known but estimated in practice from the data [35]. We calculated $(\hat{\theta}_{ik1}, \hat{\theta}_{ik2}, \dots, \hat{\theta}_{ikM})$ and the diagonal elements of \mathbf{S}_{ik} using standard formulae for log-odds and variance of the log-odds [2]. We applied a continuity correction by adding 0.5 to the numerators and 1 to the denominators of studies with 0% or 100% event rate in one of the treatment arms [36,37]. The off-diagonals of \mathbf{S}_{ik} were calculated from estimates of within-study correlations r_{ik}^{mn} between outcomes m and n ($m \neq n$) in arm k of study i obtained from studies with IPD (see Additional file 1: Table S1). The method used to estimate the correlations from the IPD is described in the implementation section below.

When summarising evidence across multiple end-points, it is common to encounter instances where some studies do not report information for all outcomes of interest leading to incomplete vectors with missing study-specific effects for the outcomes not reported [5,10]. Such studies can be included in our model under the assumption that the effects for outcomes not reported are missing at random. When implemented using the WinBUGS software, the missing study effects and standard errors are coded as NA in the data, a strategy previously outlined in Bujkiewicz *et al.* [10] and Dakin *et al.* [28]. This enables WinBUGS to automatically

'impute' values for the missing information under missing at random assumption with predicted distributions.

We refer to equation (5) as the within-study model and the model describing the distribution of the 'true' effects across studies (presented below) as the between-study model following standard terminology in multivariate meta-analysis [5,6,10,11,38,39]. For the network of two-arm trials, the between-study model for the i th study is thus given by:

$$\begin{pmatrix} \delta_{i(bk)1} \\ \vdots \\ \delta_{i(bk)M} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} d_{(bk)1} = d_{(Ak)1} - d_{(Ab)1} \\ \vdots \\ d_{(bk)M} = d_{(Ak)M} - d_{(Ab)M} \end{pmatrix}, \Sigma_{(bk)} \right) \quad (6)$$

$$\Sigma_{(bk)} = \begin{pmatrix} \sigma_{(bk)1}^2 & \cdots & \rho_{bk}^{1M} \sigma_{(bk)1} \sigma_{(bk)M} \\ \vdots & \ddots & \vdots \\ \sigma_{(bk)M}^2 & \cdots & \sigma_{(bk)M}^2 \end{pmatrix}$$

where the 'true' effects $\delta_{i(bk)m}$ ($m = 1, 2, \dots, M$) jointly follow a Normal distribution with mean effects $d_{(bk)m}$. The parameters in equation (6) have the same interpretation as in equation (2) except that they are now specific to each outcome. The covariance matrix $\Sigma_{(bk)}$ contains terms for the between-study variances, $\sigma_{(bk)m}^2$ for each outcome m and the between-study correlations ρ_{bk}^{mn} between effects measured by outcome m and n ($m \neq n$) specific to each k versus b comparison. Fitting the full model would thus require a large number of possibly multi-arm studies in order to make $\Sigma_{(bk)}$ identifiable [5,13]. The number of parameters in $\Sigma_{(bk)}$, can however be reduced if reasonable assumptions can be made about the covariance structure. In particular, most practical applications of NMA methods involve the assumption of a common between-study variance across treatment arms, often referred to as a homogenous variance assumption [18,40,41]. Therefore, to simplify $\Sigma_{(bk)}$ we make the additional assumption in this context of a common between-study correlation ($\rho_{bk}^{mn} = \rho^{mn}$) leading to the following simplified between-study covariance structure for two-arm studies:

$$\begin{pmatrix} \delta_{i(bk)1} \\ \vdots \\ \delta_{i(bk)M} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} d_{(bk)1} = d_{(Ak)1} - d_{(Ab)1} \\ \vdots \\ d_{(bk)M} = d_{(Ak)M} - d_{(Ab)M} \end{pmatrix}, \Sigma_{(M \times M)} \right)$$

$$\Sigma_{(M \times M)} = \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \sigma_M^2 & \cdots & \sigma_M^2 \end{pmatrix} \quad (7)$$

where, as in the univariate case, σ_m represent the common between-study standard deviation or heterogeneity parameter specific to outcome m .

To include multi-arm studies in our model, we extend equations (3) and (4) to the multiple outcome setting. We show in Appendix A, that under evidence consistency and the homogenous between-study covariance structure ($\sigma_{(bk)m}^2 = \sigma_m^2$ and $\rho_{bk}^{mn} = \rho^{mn}$), equation (3) can be extended

to the multiple outcome settings by formulating the distribution of effects in a multi-arm study i with $p + 1$ arms reporting on $m = 1, 2, \dots, M$ outcomes as follows:

$$\begin{pmatrix} \begin{pmatrix} \delta_{i(bk_1)1} \\ \vdots \\ \delta_{i(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} \delta_{i(bk_2)1} \\ \vdots \\ \delta_{i(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \delta_{i(bk_p)1} \\ \vdots \\ \delta_{i(bk_p)M} \end{pmatrix} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} \begin{pmatrix} d_{(bk_1)1} \\ \vdots \\ d_{(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} d_{(bk_2)1} \\ \vdots \\ d_{(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} d_{(bk_p)1} \\ \vdots \\ d_{(bk_p)M} \end{pmatrix} \end{pmatrix}, \Sigma_{(Mp \times Mp)} \right) \quad (8)$$

$$\Sigma_{(Mp \times Mp)} =$$

$$\begin{pmatrix} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \cdots & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \cdots & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} \end{pmatrix}$$

where p is the number of treatment effect estimates. The corresponding marginal and conditional distributions for arm j , given the previous $1, \dots, (j - 1)$ arms are:

$$\begin{pmatrix} \delta_{i(bk_1)1} \\ \vdots \\ \delta_{i(bk_1)M} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} d_{(bk_1)1} \\ \vdots \\ d_{(bk_1)M} \end{pmatrix}, \Sigma_{(M \times M)} = \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M} \sigma_1 \sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M} \sigma_1 \sigma_M & \cdots & \sigma_M^2 \end{pmatrix} \right) \text{ for } j = 1$$

$$\begin{pmatrix} \begin{pmatrix} \delta_{i(bk_1)1} \\ \vdots \\ \delta_{i(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} \delta_{i(bk_2)1} \\ \vdots \\ \delta_{i(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \delta_{i(bk_j)1} \\ \vdots \\ \delta_{i(bk_j)M} \end{pmatrix} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} \begin{pmatrix} d_{(bk_1)1} \\ \vdots \\ d_{(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} d_{(bk_2)1} \\ \vdots \\ d_{(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} d_{(bk_j)1} \\ \vdots \\ d_{(bk_j)M} \end{pmatrix} \end{pmatrix}, \Sigma' = \frac{(j+1)}{2j} \Sigma_{(M \times M)} \right) \text{ for } j = 2, \dots, p$$

(9)

To complete model 2, μ_{ibm} and $d_{(1k)m}$ are given minimally informative prior distributions:

$$\mu_{ibm}, d_{(1k)m} \sim \text{Normal}(0, 10^3)$$

Prior distributions also need to be specified for $\Sigma_{(M \times M)}$ which, in general, is non-trivial because of the positive definite constraint. Initially we specified an Inverse-Wishart distribution [42]:

$$\Sigma_{(M \times M)} \sim \text{Inverse - Wishart}(\mathbf{K}, M)$$

where \mathbf{K} is $M \times M$ scale matrix and M is the total number of outcomes. Specifying minimally informative Inverse-Wishart prior distribution is, however, problematic, especially when the amount of data is small relative to the dimensions of $\Sigma_{(M \times M)}$ as is the case for our example data. Therefore, to allow for flexibility in formulating a prior distribution for $\Sigma_{(M \times M)}$, we also followed a strategy outlined by Lu and Ades (2009) [43] and more recently by Wei and Higgins (2013) [39] to express $\Sigma_{(M \times M)}$ in terms of a diagonal matrix of standard deviations $\mathbf{V}^{1/2}$ and squared positive semi-definite matrix of correlations \mathbf{R} based on a separation strategy (Barnard et al. [44]):

$$\Sigma_{(M \times M)} = \mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2}$$

where the off-diagonal elements of \mathbf{R} contain correlation terms and diagonal elements equal 1. Lu and Ades [43] and also Wei and Higgins [39] showed that \mathbf{R} can be written as $\mathbf{R} = \mathbf{L}^T \mathbf{L}$ using Cholesky decomposition where \mathbf{L} is an upper triangular matrix. The spherical parameterization technique [39,43] can be used to express \mathbf{R} in terms of sine and cosine functions of the elements in \mathbf{L} . Using this later technique, we specified Uniform $(0, \pi)$ prior distributions for the spherical coordinate φ_{mm} in our model to ensure that elements of the correlation matrix \mathbf{R} lie in the interval $(-1, 1)$. Finally, the elements of $\mathbf{V}^{1/2}$ correspond to the between-study standard deviation terms in $\Sigma_{(M \times M)}$ and are given independent Uniform $(0, 2)$ prior distributions as in the univariate case (model 1).

Model 3: Borrowing strength across interventions and outcomes

From Table 1, it can be seen that none of the studies had considered the interventions E + HSI and E + HV for storage of medicines and storage of other household products. Similarly, interventions E + FE + F and FE were not trialled by any of the included studies on possession of a PCC number. To estimate the full set of 24 basic intervention effects relative to usual care from 9 interventions on 3 outcomes, we applied methods originally proposed by DuMouchel and Harris [29] and revisited by DuMouchel and Groer [45] and Jones *et al.* [30]. We assume that the pooled effects of treatment k relative to usual care intervention $d_{(Ak)m}$ can be expressed as a sum of treatment-specific effect α_k and outcome-specific effect γ_m . This assumption replaces the minimally informative prior distribution $N(0, 10^3)$ specified for $d_{(Ak)m}$ in model 2 with:

$$d_{(Ak)m} \sim \text{Normal}(\alpha_k + \gamma_m, \tau^2) \quad (10)$$

$$k = 2, 3, \dots, K; m = 1, 2, M$$

where K is the total number of treatments being evaluated across M outcomes, and for $k = 1$ (i.e. reference treatment A), $d_{(Ak)m}$ equal to zero. Note that on the logarithmic scale, this would imply that the ratio of any intervention effects is constant across outcomes as the γ_m cancel, i.e.

$$d_{(bk)m} = (d_{(Ak)m} - d_{(Ab)m}) \sim \text{Normal}(\alpha_k - \alpha_b, 2\tau^2) \quad (11)$$

Equation (10) thus embodies an assumption of equal or constant relative potency of treatments across outcomes which imply exchangeability of the relative effects between the non-reference/baseline treatments indicated by equation (11). For our example dataset, this implies that missing intervention effects for comparisons with the usual care intervention can be predicted directly

from equation (10) as a linear combination of γ_m and α_k assuming that each treatment effect relative to usual care is reported on at least one outcome. The missing intervention effects between non-reference/baseline treatments if required can similarly be predicted directly from the model as linear combinations of the intervention effects relative to usual care.

The parameter τ controls the accuracy of the constant relative potency assumption. Values of τ close to zero would thus indicate a high degree of confidence (and support from the data) in the parallelism of effect profiles across outcomes and the constant relative potency assumption. Conversely, larger values of τ would indicate otherwise.

Multi-arm studies are included in model 3 based on equations (8) and (9) in the same way as in model 2. To complete model 3, the parameters α_k , γ_m and τ are given minimally informative prior distributions. For the mean effects, this is a normal distribution with zero mean and large variance:

$$\alpha_k, \gamma_m \sim \text{Normal}(0, 10^3)$$

We give τ a Uniform $(0, 2)$ prior distribution, considered to be minimally informative on the log-odds ratio scale. Sensitivity analyses were conducted to assess the impact of specifying alternative prior distributions for τ that are also considered minimally informative [46]:

- Normal prior distribution centred on 0 with large variance and constrained to be positive, $\tau \sim N(0, 10^2)$, $\tau \geq 0$
- Gamma prior distribution placed on the precision: $\tau^{-2} \sim \text{Gamma}(0.001, 0.001)$.

The results of the sensitivity analyses are presented in Additional file 1: Figure S1.

There is a limitation to the number of data (i.e. intervention effects relative to the usual care) on outcomes allowed to be missing for the model hyper-parameters to be identifiable. For K interventions and M outcomes, there will be $(K - 1) \times M$ equation (10) that are used to estimate a total of $(K - 1) + M$ hyper-parameters (i.e. $(K - 1)$ of α_k and M of γ_m hyper-parameters). Therefore no more than $((K - 1) \times M) - ((K - 1) + M)$ missing values in total are allowed. For example, for $K = 3$ treatments and $M = 2$ outcomes, data has to be available on both outcomes for both treatment comparisons with the baseline when the prior distributions are non-informative. When large number of data on outcomes is missing, placing informative prior distributions on the hyper-parameters can improve convergence.

Implementation of the models

We fitted a total of four models, models 1 and 3 as described above and two versions of model 2. In model 2a,

we specified an inverse-Wishart prior distribution for the between-study covariance matrix $\Sigma_{(M \times M)}$ whilst in model 2b, we specified a prior distribution for $\Sigma_{(M \times M)}$ based on the separation strategy. All four models allowed for multi-arm trials to be included in the analysis. To fit the multivariate NMA models, the quantities $(\theta_{ik1}, \theta_{ik2}, \theta_{ik3})$ and the diagonals of S_{ik} were estimated using standard 2×2 -table formulae [2]. Next, we obtained estimates of the within-study correlations from the IPD studies using the following three methods: i) Pearson correlation coefficient between the observed outcome events ii) Bootstrapping as described in Daniel and Hughes (1998), and iii) Generalised Estimating Equations (See details in Additional file 1). All three methods produced identical estimates of the correlations between pairs of outcome specific log-odds of event from each IPD study (Additional file 1: Table S1). Therefore, we formulated informative prior distributions for the correlation terms in S_{ik} of equation (5) using the estimates of the correlations between the observed outcome events (Pearson correlation) as follows:

$$r_{ik}^{mn} \sim \text{Uniform}(a^{mn}, b^{mn}) \text{ with}$$

$$a^{mn} = \bar{r}^{mn} - \left(\frac{\sqrt{12 \times \text{var}(r^{mn})}}{2} \right) \text{ and}$$

$$b^{mn} = \bar{r}^{mn} + \left(\frac{\sqrt{12 \times \text{var}(r^{mn})}}{2} \right)$$

where r_{ik}^{mn} is the within-study correlation between the outcomes m and n effects measured on the log-odds scale in arm k of study i , and \bar{r}^{mn} and $\text{var}(r^{mn})$ are the mean and variance of the within-study correlation between outcomes m and n effects measured by IPD respectively.

We also assessed consistency of the evidence within each network using a method based on node splitting [24]. We found no evidence of conflict between the direct and indirect sources (Additional file 1: Table S2) in all three outcome networks.

We fitted all models described above in WinBUGS [47] using Markov Chain Monte Carlo (MCMC) simulations. The univariate models were fitted separately for each outcome using WinBUGS code available from Dias *et al.* [48]. The WinBUGS code for the multivariate models is provided in the appendices 1 and 2 of the Additional file 1. Convergence was assessed by examination of the trace and autocorrelation plots and the Rubin-Gelman statistic after running 400 000 simulations and discarding the first 200 000 samples as 'burn in samples'.

Results

Univariate and multivariate analyses

Parameters of interest were the posterior median estimate (and 95% credible intervals) of the pooled intervention effects relative to the usual care intervention and the posterior median estimate (and 95% credible intervals) of the between-study standard deviation and correlation terms. Summary forest plots displaying effectiveness estimates relative to usual care on the odds ratio (OR) scale are presented in Figure 2. It can be seen that, all 4 models produced broadly similar estimates when the treatment effect is not extreme compared to the other effect estimates for the same outcome. Compared to the univariate analysis, the multivariate models produced noticeably less extreme estimates of intervention effects. This can be seen in the effect of FE + HSI (3) on possession of PCC number being shifted towards the line of no effect from an OR of 39.35 (95% CrI 2.37 to 732.30) in model 1 to 23.55 (95% CrI 1.39, to 456.80) in model 2a, 20.37 (95% CrI 0.72, to 706.00) in model 2b and 4.20 (95% CrI 1.59 to 13.16) in model 3. Similarly, the OR for FE (9) on safe storage of other household products shifted from 0.37 (95% CrI 0.00 to 15.10) in model 1 to 1.81 (95% CrI 0.63, to 5.37) in model 3.

Posterior median and 95% credible intervals of the between-study standard deviations and correlations are presented in Table 2. The posterior medians of the between-study correlations from the multivariate models were small and estimated with considerable uncertainty (i.e. all had large variances). Estimates of the between-study standard deviations were broadly similar for the univariate NMA (model 1) and the multivariate NMA using the separation strategy (model 2b), and relatively high for multivariate NMA using the inverse-Wishart prior distribution (model 2a).

Borrowing strength across outcomes

It can be seen from Figure 2 that the effect of E + HSI and E + HV relative to usual care intervention on safe storage of medicines and safe storage of other household products, and E + FE + F and FE on possession of a PCC telephone number were only estimated in model 3 as none of the studies had trialled these interventions on the respective outcomes. In this model, estimates of relative effects between non-reference/baseline treatments were assumed to be exchangeable across outcomes, which enabled estimates to be obtained for all outcomes by predicting effects where the interventions have not been considered for the particular outcome of interest. For the intervention/outcome pair where data from trials were available, this extrapolation step had the additional effect of producing more precise estimates of the treatment effect in comparison to the models that do not assume exchangeability effects across outcomes.

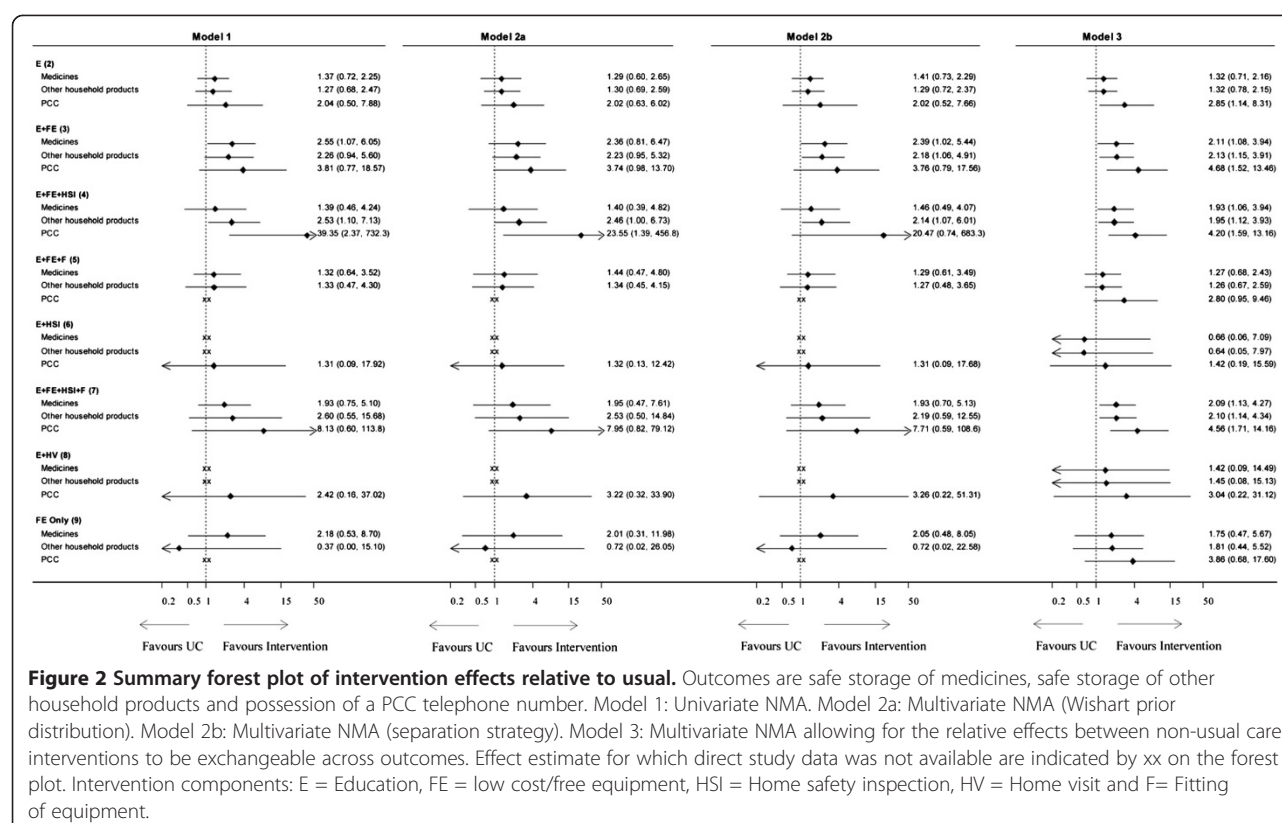


Table 2 Posterior median and 95% credible intervals of the between-study standard deviation and correlation parameters

Parameter	Description/prior distribution	Model 1: univariate	Model 2a: Multivariate using inverse-Wishart prior distribution for $\Sigma_{(M \times M)}$	Model 2b: Multivariate using a separation strategy to specify priors for elements of $\Sigma_{(M \times M)}$	Model 3: Multivariate with extrapolation of effects across outcomes
σ_1	Between-study standard deviation: safe storage of medicines	0.26 (0.03, 1.02)	0.58 (0.33, 1.18)	0.27 (0.01, 1.08)	0.23 (0.01, 0.80)
σ_2	Between-study standard deviation: safe storage of other household products	0.56 (0.13, 1.27)	0.62 (0.35, 1.15)	0.47 (0.04, 1.18)	0.31 (0.01, 0.81)
σ_3	Between-study standard deviation: PCC	1.16 (0.57, 1.93)	0.94 (0.53, 1.99)	1.18 (0.57, 1.93)	1.08 (0.58, 1.85)
τ	Primary analysis: $\tau \sim \text{Uniform}(0, 2)$	-	-	-	0.10 (0.01, 0.53)
τ	Sensitivity analysis: $\tau \sim \text{Normal}(0, 10^2)$, $\tau \geq 0$	-	-	-	0.11 (0.00, 0.56)
τ	Sensitivity analysis: $\tau^2 \sim \text{Inverse-Gamma}(0.001, 0.001)$	-	-	-	0.08 (0.02, 0.36)
ρ^{12}	Between-study correlation [medicines, other household products]	-	0.03 (-0.73, 0.76)	0.05 (-1.00, 1.00)	0.45 (-0.99, 1.00)
ρ^{13}	Between-study correlation [medicines, PCC]	-	0.06 (-0.80, 0.81)	0.20 (-1.00, 1.00)	0.50 (-0.98, 1.00)
ρ^{23}	Between-study correlation [Other household products, PCC]	-	0.08 (-0.81, 0.83)	0.13 (-0.97, 0.98)	0.60 (-0.87, 0.99)

The posterior median and 95% credible intervals of intervention effects relative to usual care were unaffected by placing alternative minimally informative prior distributions on τ (Additional file 1: Appendix 2). The posterior median and credible intervals for τ (Table 2) were similarly not sensitive to the choice of prior distribution placed on τ in the primary and sensitivity analyses. The posterior median estimates of τ were all close to zero, which suggest that assumptions about the parallelism of effect profiles across outcomes is supported by the data. The estimates of τ would thus suggest with 95% probability, that based on the information in our example dataset, the extrapolation model could be accurate to within a factor of about $e^{(2 \times 0.10)} = 1.24$ (95% CrI 1.01 to 2.87).

Discussion

We have presented methods for simultaneous comparison of multiple treatments across multiple outcome measures while preserving the internal randomisation of individual studies. Our method may be viewed as an extension of Ades *et al.*'s (2010) NMA with competing risks paper [13] wherein only the within-study correlation is taken into account. We have extended their method to account for the dependency between outcome effects across studies as well as within-studies.

In this particular application of the multivariate approach, accounting for the correlation between outcomes alone (models 2a and 2b) did not reduce the uncertainty around estimates of intervention effects compared to analysing each outcome separately (model 1). Assuming that intervention effects are exchangeable across outcome did however lead to a modest reduction in uncertainty around effectiveness estimates (model 3).

The between-study correlations were estimated with considerable uncertainty (Table 2) and appear to have little impact on overall effect estimates. This may be because the between-study correlation arises due to, among other things, differences in study-level characteristics that also give rise to between-study heterogeneity in a meta-analysis. Based on a criterion outlined in Spiegelhalter *et al.* [49] the posterior median estimates of the between-study standard deviations, σ_1 and σ_2 on the log odds ratio scale (Table 2) could be interpreted as indicating evidence of low to moderate heterogeneity for storage of medicines and storage of other household products outcomes. Only the estimates for possession of poison control centre number exhibited a considerable degree of heterogeneity. Consequently, the posterior medians of the between-study correlations were small. There was therefore very little gain (in terms of increasing the precision of estimates) from formulating the between-study covariance structure described for the analysis

presented here. Accounting for the between-study correlation is likely to be beneficial in situations where the between-study variance (heterogeneity) is large relative to within-study variances.

We opted to incorporate the within-study correlation through the arm-specific effects (log-odds) rather than the study-specific treatment difference (log-odds ratio) as is often done in multivariate meta-analysis [3,4,15,38]. This approach greatly simplifies the likelihood for multi-arm studies because treatment arms can be considered independent as a consequence of randomisation. Hence, there is no requirement to account for the additional correlations between effect estimates which share a common comparator treatment in the model likelihood [50]. The arm-based approach is also likely to be useful when (as is typical with many practical application of multivariate meta-analysis) the within-study correlations are not available [10,12,15,51] and have to be obtained from an external source such as expert opinion [14]. In such situations, formulating questions about correlations between outcome-specific event probabilities (which can be used directly in an arm-based approach) is more likely to be intuitive and easily understood by non-statistician healthcare experts than questions about correlations between intervention effects. It is acknowledged however, that the correlations between the intervention effects if required can easily be obtained from the correlations between the outcomes [14,51].

At the between-study level, we assumed a common correlation structure across treatments in addition to the common variance assumption underlying most practical application of NMA methods. The common correlation assumption implies that if several separate multivariate meta-analyses were conducted with the same outcomes, each with a different set of k versus b comparison, the assumption is that the between-study correlations would be the same across the different sets of bk comparisons. We suggested this structure to simplify the covariance structure and reduce the number of parameters in the model. Appropriateness of such modelling assumptions would need to be considered carefully and assessed when it is feasible to do so.

Initially we specified an inverse-Wishart distribution for the between-study covariance matrix $\Sigma_{(M \times M)}$. However, we believe this prior distribution to be influential due to the small number of studies in our example dataset relative to the number of outcomes. Under these conditions, the inverse-Wishart prior distribution produced upwardly-biased estimates of σ_1 and σ_2 and downward bias in the estimate for σ_3 when compared to the corresponding estimates obtained from the univariate model (Table 2). These findings are consistent with observations in the univariate case where the use of a

Gamma prior distribution (which is the univariate analogue of the Inverse-Wishart prior distribution) can lead to an overestimation of the heterogeneity parameter when the true value is close to 0 [46,52]. As an alternative to an inverse-Wishart prior distribution therefore, we followed the spherical decomposition technique suggested by Lu and Ades [43]. This parameterization offered greater flexibility in formulating independent prior distributions for the standard deviation and correlation terms in $\Sigma_{(M \times M)}$.

An obvious limitation to implementation of the multivariate models presented in this paper is the limited availability of data including i) the problem of missing within-study correlations and ii) the requirement for a relatively large number of studies to estimate all model parameters. The problem of missing within-study correlations has traditionally hampered the widespread application of multivariate meta-analysis [7,10,15]. In our example, IPD was available from a proportion of the included studies and we have used correlations estimated from the IPD to formulate informative prior distributions for the within-study model. Alternative approaches to dealing with missing within-study correlations when IPD is not available include: i) using the observed correlation from the summary study-specific effects [12], ii) eliciting information about the correlations from external sources such as clinical experts [14] and iii) specifying 'vague' prior distributions for analysis conducted within a Bayesian framework [6].

The second data issue concerns the number of studies needed to estimate the full unstructured between-study covariance matrix presented in equation (6). We anticipate a large number of multi-arm studies reporting across the three outcomes will be needed to identify $\Sigma_{(bk)}$ and estimate all model parameters. This can be problematic considering the fact that most applications of network meta-analysis typically include mostly two-arm studies with very small numbers of multi-arm studies. Even with the simplification of the between-study covariance matrix given in equation (5), a relatively large number of studies in comparison to the total number of outcomes being considered may still be needed. We are unable to answer the question of how many studies should be considered large enough for a NMA with multiple outcomes. As a guide, Wei and Higgins [39] recently estimated 15, 27 and 42 studies as a minimum for multivariate pairwise meta-analysis with two, three and four-outcomes respectively. Hence, we believe an even larger number of studies will be required for the NMA with multiple outcomes.

Another limitation of the multivariate models presented here is that they rely on the normal approximation to binomial distribution to incorporate the within-study

correlations in the model. The normal approximation frequently fails and may not provide adequate fit to the data in the presence of studies with zero or a small number of events, necessitating use of continuity corrections. We were unable to use the exact binomial distribution as our primary interest was to develop models for summary binary data where outcomes are not mutually exclusive, and where it is not reasonable to assume that within-study correlations are zero so that the likelihood factorises easily as in Arends *et al.* [3]. Further methodological investigations into modelling multivariate summary data that is not normally distributed will therefore be useful. An example is provided in Chu *et al.* [53] where parameterization of the within-study model enabled the special case of diagnostic sensitivity and specificity to be jointly modelled with disease prevalence using a trivariate binomial likelihood. In the interim, an alternative formulation which bypasses the need for approximating normal distributions is to directly model the IPD where this is available. This will require extending Saramago *et al.*'s [54] NMA model with aggregate and individual participant level data from single outcome to multiple outcome settings.

We assessed the consistency of each outcome network separately using the method of node splitting [24]. We found no evidence of conflict between the direct and indirect sources on pairwise contrasts that have both sources of evidence in model 1. We did not assess the consistency of the multivariate estimates partly because we are unaware of current methods for carrying out this type of assessment. We are investigating extensions of the node-split method to multiple outcome networks and investigate the effect of jointly synthesising evidence across multiple endpoints on evidence consistency in a simulation study.

Our initial motivation for a multiple outcome NMA was to estimate intervention effects for all the outcomes, including effects of interventions on outcomes not considered by any of the studies included in the analysis. This requires the correlation structure between effects on multiple outcomes to be appropriately modelled and also ensuring the mechanism of "borrow strength" across outcomes through the assumption of exchangeability of the random effect across outcomes. This implies a priori assumption that outcomes are related but different and that there is no way of knowing the order of magnitude of effects on outcomes. If this assumption does not hold, it may potentially lead to worse or more biased effectiveness estimates. In our example, the outcomes are similar and measured on the same scale. It would be clearly inappropriate to assume that intervention effects are exchangeable across outcomes that are different in some important respects such as

$$\Sigma_{\text{FULL}} = \begin{pmatrix} \sigma_{(AB)1}^2 & \cdots & \rho_{(AB,AB)}^{1M} \sigma_{(AB)1} \sigma_{(AB)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AB,AB)}^{1M} \sigma_{(AB)1} \sigma_{(AB)M} & \cdots & \sigma_{(AB)M}^2 \end{pmatrix} \begin{pmatrix} \rho_{(AB,AC)}^{11} \sigma_{(AB)1} \sigma_{(AC)1} & \cdots & \rho_{(AB,AC)}^{1M} \sigma_{(AB)1} \sigma_{(AC)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AB,AC)}^{1M} \sigma_{(AB)1} \sigma_{(AC)M} & \cdots & \rho_{(AB,AC)}^{MM} \sigma_{(AB)M} \sigma_{(AC)M} \end{pmatrix} \cdots \begin{pmatrix} \rho_{(AB,AK)}^{11} \sigma_{(AB)1} \sigma_{(AK)1} & \cdots & \rho_{(AB,AK)}^{1M} \sigma_{(AB)1} \sigma_{(AK)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AB,AK)}^{1M} \sigma_{(AB)1} \sigma_{(AK)M} & \cdots & \rho_{(AB,AK)}^{MM} \sigma_{(AB)M} \sigma_{(AK)M} \end{pmatrix} \\ \begin{pmatrix} \sigma_{(AC)1}^2 & \cdots & \rho_{(AC,AC)}^{1M} \sigma_{(AC)1} \sigma_{(AC)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AC,AC)}^{1M} \sigma_{(AC)1} \sigma_{(AC)M} & \cdots & \sigma_{(AC)M}^2 \end{pmatrix} \vdots \begin{pmatrix} \rho_{(AC,AK)}^{11} \sigma_{(AC)1} \sigma_{(AK)1} & \cdots & \rho_{(AC,AK)}^{1M} \sigma_{(AC)1} \sigma_{(AK)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AC,AK)}^{1M} \sigma_{(AC)1} \sigma_{(AK)M} & \cdots & \rho_{(AC,AK)}^{MM} \sigma_{(AC)M} \sigma_{(AK)M} \end{pmatrix} \\ \vdots \begin{pmatrix} \sigma_{(AK)1}^2 & \cdots & \rho_{(AK,AK)}^{1M} \sigma_{(AK)1} \sigma_{(AK)M} \\ \vdots & \ddots & \vdots \\ \rho_{(AK,AK)}^{1M} \sigma_{(AK)1} \sigma_{(AK)M} & \cdots & \sigma_{(AK)M}^2 \end{pmatrix}$$

The homogenous between-study correlation assumption implies $\rho_{(Ah,Ah)}^{mn} = \rho_{(Ak,Ak)}^{mn} = \rho^{mn}$ and $\rho_{(Ak,Ah)}^{mn} = \frac{1}{2}\rho^{mn}$ for the inequality $-1 \leq (\rho_{(Ak,Ak)}^{mn} + \rho_{(Ah,Ah)}^{mn} - 2\rho_{(Ak,Ah)}^{mn}) \leq 1$ to hold. Substituting these expressions into equation (A1), we see that the between-study correlation terms equal ρ^{mn} in the diagonal block of matrices and $\frac{1}{2}\rho^{mn}$ in the off-diagonal block of matrices of in Σ_{FULL} leading to the following simplification following simplification of the between-study covariance matrix:

$$\begin{pmatrix} \begin{pmatrix} \delta_{i(AB)1} \\ \vdots \\ \delta_{i(AB)M} \end{pmatrix} \\ \begin{pmatrix} \delta_{i(AC)1} \\ \vdots \\ \delta_{i(AC)M} \end{pmatrix} \\ \begin{pmatrix} \delta_{i(AK)1} \\ \vdots \\ \delta_{i(AK)M} \end{pmatrix} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} \begin{pmatrix} d_{(AB)1} \\ \vdots \\ d_{(AB)M} \end{pmatrix} \\ \begin{pmatrix} d_{(AC)1} \\ \vdots \\ d_{(AC)M} \end{pmatrix} \\ \begin{pmatrix} d_{(AK)1} \\ \vdots \\ d_{(AK)M} \end{pmatrix} \end{pmatrix}, \Sigma_{(Mp \times Mp)} \right)$$

with $\Sigma_{(Mp \times Mp)} = \begin{pmatrix} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \cdots & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} \\ \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \vdots & \frac{1}{2} \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} & \\ & & \begin{pmatrix} \sigma_1^2 & \cdots & \rho^{1M}\sigma_1\sigma_M \\ \vdots & \ddots & \vdots \\ \rho^{1M}\sigma_1\sigma_M & \cdots & \sigma_M^2 \end{pmatrix} \end{pmatrix}$

(A3)

Finally by relabeling the reference treatment A as b , $(\delta_{i(AB)1}, \dots, \delta_{i(AK)M})$ as $(\delta_{i(bk_1)1}, \dots, \delta_{i(bk_j)M})$ and $(d_{(AB)1}, \dots, d_{(AK)M})$ as $(d_{(bk_1)1}, \dots, d_{(bk_j)M})$, equation (A3) can be rewritten as equation (A4).

$$\begin{pmatrix} \begin{pmatrix} \delta_{i(bk_1)1} \\ \vdots \\ \delta_{i(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} \delta_{i(bk_2)1} \\ \vdots \\ \delta_{i(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \delta_{i(bk_p)1} \\ \vdots \\ \delta_{i(bk_p)M} \end{pmatrix} \end{pmatrix} \sim \text{Normal} \left(\begin{pmatrix} \begin{pmatrix} d_{(bk_1)1} \\ \vdots \\ d_{(bk_1)M} \end{pmatrix} \\ \begin{pmatrix} d_{(bk_2)1} \\ \vdots \\ d_{(bk_2)M} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} d_{(bk_p)1} \\ \vdots \\ d_{(bk_p)M} \end{pmatrix} \end{pmatrix}, \Sigma_{(Mp \times Mp)} \right)$$

(A4)

Additional file

Additional file 1: (WinBUGS Code for model 2): Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes.

Abbreviations

CrI: Credible interval; IPD: Individual patient data; MCMC: Markov Chain Monte Carlo; NMA: Network meta-analysis; PCC: Poisson centre control number.

Competing interests

The authors declare that they have no competing interests.

Author's contribution

FAA drafted the initial manuscript, conducted the analysis and coordinated contributions from all authors in drafting the final manuscript. NJC designed the study and revised the manuscript to improve general readability. SB contributed to developing the statistical analysis and interpretation of the results. SJH contributed to preparing the data for analysis, drafting the network diagrams and revision of the initial and final drafts. DK provided the data for analysis, defined the outcomes and interventions and revision of the final drafts. DRJ participated in designing the study and revision of the final manuscript. AJS conceived the original idea and participated developing the statistical model. All authors read and approved the final manuscript.

Acknowledgements

FAA is funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research programme (Reference Number RP-PG-0407-10231) to pursue a PhD in Public Health Evaluation of Interventions.

SB is supported by the MRC Methodology Research Programme (grant number MR/L009854/1).

Author details

¹Biostatistics Group, Department of Health Sciences, University of Leicester, University Road, Leicester LE1 7RH, UK. ²Division of Primary Care Community Health Sciences, Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham NG7 2RD, UK.

Received: 8 January 2014 Accepted: 2 July 2014

Published: 21 July 2014

References

- Cooper NJ, Sutton AJ, Ades AE, Paisley S, Jones DR, W. G. O. U. o. E. i. E. D: **Models, Use of evidence in economic decision models: Practical issues and methodological challenges.** *Health Economics* 2007, **16**(12):1277-1286.
- Sutton A, Abrams K, Jones DR, Sheldon TA, Song F: *Methods for Meta-analysis in Medical Research*. Chichester: Wiley; 2000.
- Arends LR, Voko Z, Stijnen T: **Combining multiple outcome measures in a meta-analysis: an application.** *Statistics in Medicine* 2003, **22**(8):1335-1353.
- Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA: **Meta-analysis of multiple outcomes by regression with random effects.** *Statistics in Medicine* 1998, **17**(22):2537-2550.
- Jackson D, Riley R, White IR: **Multivariate meta-analysis: Potential and promise.** *Statistics in Medicine* 2011, **30**(20):2481-2498.
- Nam IS, Mengersen K, Garthwaite P: **Multivariate meta-analysis.** *Statistics in Medicine* 2003, **22**(14):2309-2333.
- Riley RD, Abrams KR, Sutton AJ, Lambert PC, Thompson JR: **Bivariate random-effects meta-analysis and the estimation of between-study correlation.** *BMC Medical Research Methodology* 2007, **7**.
- Riley RD, Thompson JR, Abrams KR: **An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown.** *Biostatistics* 2008, **9**(1):172-186.
- Kendrick D, Coupland C, Mulvaney C, Simpson J, Smith SJ, Sutton A, Watson M, Woods A: **Home safety education and provision of safety equipment for injury prevention.** *Cochrane Database Systematic reviews* 2007, **1**:CD005014.
- Bujkiewicz S, Thompson JR, Sutton AJ, Cooper NJ, Harrison MJ, Symmons DPM, Abrams KR: **Multivariate meta-analysis of mixed outcomes: a Bayesian approach.** *Statistics in Medicine* 2013, **32**(22):3926-3943.
- Riley RD, Abrams KR, Lambert PC, Sutton AJ, Thompson JR: **An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes.** *Statistics in Medicine* 2007, **26**(1):78-97.
- Kirkham JJ, Riley RD, Williamson PR: **A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews.** *Statistics in Medicine* 2012, **31**(20):2179-2195.
- Ades AE, Mavranzeouli I, Dias S, Welton NJ, Whittington C, Kendall T: **Network meta-analysis with competing risk outcomes.** *Value Health* 2010, **13**(8):976-983.
- Efthimiou O, Mavridis D, Cipriani A, Leucht S, Bago P, Salanti G: **An approach for modelling multiple correlated outcomes in a network of interventions using odds ratios.** *Statistics in Medicine* 2014, **33**(13):2275-2287.
- Riley RD: **Multivariate meta-analysis: the effect of ignoring within-study correlation.** *Journal of the Royal Statistical Society Stat Soc* 2009, **172**:789-811.
- Rodgers M, Sowden A, Petticrew M, Arai L, Roberts H, Britten N, Popay J: **Testing Methodological Guidance on the Conduct of Narrative Synthesis in Systematic Reviews.** *Evaluation* 2009, **15**(1):49-73.
- Lumley T: **Network meta-analysis for indirect treatment comparisons.** *Statistics in Medicine* 2002, **21**(16):2313-2324.
- Caldwell DM, Ades AE, Higgins JP: **Simultaneous comparison of multiple treatments: combining direct and indirect evidence.** *British Medical Journal* 2005, **331**(7521):897-900.
- 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**(8):875-882.
- Lu G, Ades AE: **Combination of direct and indirect evidence in mixed treatment comparisons.** *Statistics in Medicine* 2004, **23**(20):3105-3124.
- Salanti G, Ades AE, Ioannidis JP: **Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.** *Journal of Clinical Epidemiology* 2010, **64**(2):163-171.
- Salanti G, Higgins JP, Ades AE, Ioannidis JP: **Evaluation of networks of randomized trials.** *Statistical Methods in Medical Research* 2008, **17**(3):279-301.
- Salanti G, Marinho V, Higgins JP: **A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered.** *Journal of Clinical Epidemiology* 2009, **62**(8):857-864.
- Dias S, Welton NJ, Caldwell DM, Ades AE: **Checking consistency in mixed treatment comparison meta-analysis.** *Statistics in Medicine* 2010, **29**(7-8):932-944.
- Hong H, Carlin BP, Chu H, Shamliyan TA, Wang S, Kane RL: *Methods Research Report: A Bayesian Missing Data Framework for Multiple Continuous Outcome Mixed Treatment Comparisons*; 2013. Available online from <http://www.ncbi.nlm.nih.gov/books/NBK116689/pdf/TOC.pdf>.
- 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**(20):3681-3699.
- Welton NJ, Cooper NJ, Ades AE, Lu G, Sutton AJ: **Mixed treatment comparison with multiple outcomes reported inconsistently across trials: evaluation of antivirals for treatment of influenza A and B.** *Statistics in Medicine* 2008, **27**(27):5620-5639.
- Dakin HA, Welton NJ, Ades AE, Collins S, Orme M, Kelly S: **Mixed treatment comparison of repeated measurements of a continuous endpoint: an example using topical treatments for primary open-angle glaucoma and ocular hypertension.** *Statistics in Medicine* 2011, **30**(20):2511-2535.
- DuMouchel WH, Harris JE: **Bayes Methods for Combining the Results of Cancer Studies in Humans and Other Species.** *Journal of the American Statistical Association* 1983, **78**(382):293-308.
- Jones DR, Peters JL, Rushton L, Sutton AJ, Abrams KR: **Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach.** *Regulatory Toxicology Pharmacology* 2009, **53**(3):217-225.
- Kendrick D, Young B, Mason-Jones AJ, Ilyas N, Achana FA, Cooper NJ, Hubbard SJ, Sutton AJ, Smith S, Wynn P, Mulvaney C, Watson MC, Coupland C: **Home safety education and provision of safety equipment for injury prevention.** *Cochrane Database Systematic reviews* 2012, **9**.
- Johnston BD, Huebner CE, Anderson ML, Tyll LT, Thompson RS: **Healthy steps in an integrated delivery system - Child and parent outcomes at 30 months.** *Archives of Pediatrics and Adolescent Medicine* 2006, **160**(8):793-800.

33. Babul S, Olsen L, Janssen P, McIntee P, Raina P: **A randomized trial to assess the effectiveness of an infant home safety programme.** *International Journal of Control and Safety Promotion* 2007, **14**(2):109–117.
34. Dias S, Sutton AJ, Ades AE, Welton NJ: **A Generalized Linear Modeling Framework for Pairwise and Network Meta-analysis of Randomized Controlled Trials.** *Medical Decision Making* 2012, **33**(5):607–17.
35. van Houwelingen HC, Arends LR, Stijnen T: **Advanced methods in meta-analysis: multivariate approach and meta-regression.** *Statistics in Medicine* 2002, **21**(4):589–624.
36. Dershowitz RA, Williamson JW: **Prevention of Childhood Household Injuries - Controlled Clinical-Trial.** *American Journal of Public Health* 1977, **67**(12):1148–1153.
37. Kelly B, Sein C, McCarthy PL: **Safety education in a pediatric primary care setting.** *Pediatrics* 1987, **79**(5):818–824.
38. Mavridis D, Salanti G: **A practical introduction to multivariate meta-analysis.** *Statistical Methods in Medical Research* 2013, **22**(2):133–158.
39. Wei Y, Higgins JPT: **Bayesian multivariate meta-analysis with multiple outcomes.** *Statistics in Medicine* 2013, **32**(17):2911–2934.
40. Carter P, Achana F, Troughton J, Gray LJ, Khunti K, Davies MJ: **A Mediterranean diet improves HbA1c but not fasting blood glucose compared to alternative dietary strategies: a network meta-analysis.** *Journal of Human Nutrition and Dietetics* 2014, **27**(3):280–297.
41. Cooper NJ, Sutton AJ, Lu G, Khunti K: **Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation.** *Archives Internal Medicine* 2006, **166**(12):1269–1275.
42. Spiegelhalter D, Thomas A, Best N, Lunn D: *WinBUGS User Manual*. Available from http://www.politicalbubbles.org/bayes_beach/manual14.pdf, 2003. Accessed 10th December 2013.
43. Lu GB, Ades A: **Modeling between-trial variance structure in mixed treatment comparisons.** *Biostatistics* 2009, **10**(4):792–805.
44. Barnard J, McCulloch R, Meng XL: **Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage.** *Statistica Sinica* 2000, **10**(4):1281–1311.
45. DuMouchel W, Groer PG: **A Bayesian Methodology for Scaling Radiation Studies from Animals to Man.** *Health Physics* 1989, **57**:411–418.
46. Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR: **How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS.** *Statistics in Medicine* 2005, **24**(15):2401–2428.
47. Lunn DJ, Thomas A, Best N, Spiegelhalter D: **WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility.** *Statistical Computing* 2000, **10**(4):325–337.
48. Dias S, Welton NJ, Sutton AJ, Ades AE: *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-analysis of Randomised Controlled Trials*; 2011. Available from <http://www.nicesdu.org.uk>.
49. Spiegelhalter DJ, Abrams KR, Myles JP: *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* *Statistics in Practice*. Chichester: John Wiley & Sons; 2004.
50. Franchini AJ, Dias S, Ades AE, Jansen JP, Welton NJ: **Accounting for correlation in network meta-analysis with multi-arm trials.** *Research Synthesis Methodology* 2012, **3**(2):142–160.
51. Wei Y, Higgins JPT: **Estimating within-study covariances in multivariate meta-analysis with multiple outcomes.** *Statistics in Medicine* 2013, **32**(7):1191–1205.
52. Gelman A: **Prior distributions for variance parameters in hierarchical models(Comment on an Article by Browne and Draper).** *Bayesian Analysis* 2006, **1**(3):515–533.
53. Chu HT, Nie L, Cole SR, Poole C: **Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: Alternative parameterizations and model selection.** *Statistics in Medicine* 2009, **28**(18):2384–2399.
54. Saramago P, Sutton AJ, Cooper NJ, Manca A: **Mixed treatment comparisons using aggregate and individual participant level data.** *Statistics in Medicine* 2012, **31**(28):3516–3536.

doi:10.1186/1471-2288-14-92

Cite this article as: Achana et al.: Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. *BMC Medical Research Methodology* 2014 **14**:92.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

