Statistics
in Medicine

Research Article

Received 22 November 2013, Accepted 24 February 2015 Published online 24 March 2015 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6475

A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis

N. L. Turner, S. Dias, A. E. Ades and N. J. Welton*†

Missing outcome data are a common threat to the validity of the results from randomised controlled trials (RCTs), which, if not analysed appropriately, can lead to misleading treatment effect estimates. Studies with missing outcome data also threaten the validity of any meta-analysis that includes them. A conceptually simple Bayesian framework is proposed, to account for uncertainty due to missing binary outcome data in meta-analysis. A pattern-mixture model is fitted, which allows the incorporation of prior information on a parameter describing the missingness mechanism. We describe several alternative parameterisations, with the simplest being a prior on the probability of an event in the missing individuals. We describe a series of structural assumptions that can be made concerning the missingness parameters. We use some artificial data scenarios to demonstrate the ability of the model to produce a bias-adjusted estimate of treatment effect that accounts for uncertainty. A meta-analysis of haloperidol versus placebo for schizophrenia is used to illustrate the model. We end with a discussion of elicitation of priors, issues with poor reporting and potential extensions of the framework. Our framework allows one to make the best use of evidence produced from RCTs with missing outcome data in a meta-analysis, accounts for any uncertainty induced by missing data and fits easily into a wider evidence synthesis framework for medical decision making. © 2015 The Authors. *Statistics in Medicine* Published by John Wiley & Sons Ltd.

Keywords: missing data; meta-analysis; Bayesian; bias; pattern-mixture model; decision making

1. Introduction

Systematic reviews identify, appraise and synthesise all information relevant to a specific research question [1] and are increasingly used to guide healthcare policy and clinical decisions [2]. The results from randomised controlled trials (RCTs) identified in a systematic review may be summarised by pooling them in a meta-analysis to obtain a single estimate of treatment effect that reflects uncertainty based on the existing RCT evidence [3]. Whilst RCTs are considered the highest quality evidence to inform relative treatment effect estimates, they may still be subject to bias [4], and any bias in an individual RCT will also be present in any meta-analysis that includes it [5]. One common threat to the validity of a trial is missing outcome data, which can lead to biased relative effect estimates [6]. In some cases, it may be appropriate to assume the outcomes are missing at random (MAR), meaning that missingness is dependent on observed data (e.g. covariates), but not dependent on the unobserved data. Under MAR, an analysis that restricts to the observed data only (complete case analysis) provides an unbiased central estimate of the treatment effect, and there is increased uncertainty due to the smaller numbers included. However, the existence of incomplete outcome data brings with it additional uncertainty concerning treatment effectiveness, which is not reflected in a complete case analysis [6] and which can lead to misleadingly precise estimates. If the MAR assumption does not hold, then a complete case analysis will produce biased, as well as overly precise estimates.

School of Social and Community Medicine, University of Bristol, Bristol, U.K.
**Correspondence to: Nicky J. Welton, School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Rd, Bristol, BS8 2PS, U.K.*
†E-mail: Nicky.Welton@bristol.ac.uk

The Cochrane Collaboration [1] and National Institute for Health and Care Excellence [7, 8] provide recommendations on how missing data should be handled within a meta-analysis in an attempt to account for this potential bias and uncertainty around treatment estimates induced by missing data. The first recommendation is that a judgement on the level of risk of bias due to missing data in each individual trial be made. The next recommendation is to explore the impact of this potential bias. Finally, it is suggested that a number of sensitivity analyses are carried out making different assumptions about the missing outcomes. For example, for binary outcomes, where we are measuring whether treatment is a success or a failure, first carry out the meta-analysis assuming that all of the missing outcomes in the data are treatment failures, that is, non-events (an 'all failures' analysis) and then repeat the analysis assuming that all missing outcomes are treatment successes, that is, events. Another option is to carry out treatment arm-specific sensitivity analyses; a 'worst-case' meta-analysis assumes that all missing data in the intervention arms of trials are treatment failures, but the missing outcomes in the control arms are successes. Conversely, a 'best-case' meta-analysis makes the opposite assumption; that is, missing data in the intervention arms are treated as successes, whilst the missing data in the control arms are assumed to be failures [1]. These approaches are useful to put limits on how extreme the effects of missing outcome data may be; however, they focus mainly around data manipulation to provide (deterministic) sensitivity analyses to assumptions about the distribution of events and non-events in missing participants, which are unlikely to be realistic in practise. For decision making, what is required is a single credible estimate of treatment effect, together with an estimate of its uncertainty accounting for the strength of evidence and missing data, which can be used as an input to a decision model.

In this paper, we present a statistical model to account for the bias and quantify uncertainty in treatment effect estimates induced by missing data [1]. We take a decision-making perspective, so that the quantity (estimand) we are interested in estimating is the effect that we would expect to see in the population of all patients that were randomised, which is attributable to the treatment they were initially randomised to. This focus on effectiveness, rather than efficacy, is termed a *de facto* hypothesis in the literature on longitudinal RCTs [9–11], and our estimand of interest corresponds to estimand 6 described by Mallinckrodt *et al.* [10].

A Bayesian statistical approach naturally allows for uncertainty due to missing data in the meta-analysis through the use of suitable prior distributions [12, 13]. A two-stage Bayesian approach to account for uncertainty due to missing binary outcomes has previously been proposed [5] that firstly adjusts study-specific estimates of treatment effect and its associated variance using prior beliefs on the missingness mechanism (using the informative missingness odds ratio (IMOR); see Section 3.3.2), and then pools these adjusted estimates in a meta-analysis [5]. A key disadvantage of taking such a two-stage approach is that the adjustment is 'fixed', and there is no potential for the observed data to inform the missingness parameter estimates nor to borrow strength across studies. Other limitations of this approach include that it makes an a priori assumption that the IMOR is independent of the amount of missingness and of other parameters, and it requires arbitrary 'fixes' when cell counts are zero. Furthermore, a one-stage approach that simultaneously incorporates prior beliefs and pools evidence across trials is conceptually simpler and allows the methods to be extended to more complex situations, for example, network meta-analysis [14]. A one-stage Bayesian approach has previously been proposed that puts a hierarchical model on the IMORs and attempts to estimate them from the observed data for both pairwise [15] and network meta-analyses [16]; however, this approach is complex to apply, is very computationally intensive and is found to have limited ability to identify the IMORs from the observed data [15, 16].

To date, the work on incorporating uncertainty due to missingness in binary outcomes in meta-analysis has focussed on the IMOR [5, 15]. However, there may be other missingness parameters on which we have prior information. In particular, if we have no prior information at all on the missingness mechanism, the most natural way to reflect this is through a flat prior on the probability of an event conditional on being missing. This will directly propagate the prior uncertainty in the true value of the missing outcomes into the analysis.

The aim of this paper is to present a general framework to reflect the uncertainty arising from missing binary outcome data in RCTs included in the meta-analysis. We present a conceptually simple one-stage Bayesian approach that allows priors (including minimally informative priors) to be given for parameters describing the missingness mechanism in a variety of ways, including the probability of success given a subject was missing, the IMOR [17], probability success ratio and the response probability ratio [18].

The paper is structured as follows. In Section 2, we introduce a motivating example meta-analysis of haloperidol versus placebo for schizophrenia with missing binary outcome data [19]. In Section 3, we begin by presenting the standard pairwise meta-analysis model for binary data and common deterministic

sensitivity analyses that are conducted. We then present our general framework to account for uncertainty due to missing binary outcome data in meta-analysis when no prior information on the missingness mechanism is available and outline how to apply the framework for some common missingness parameters for which prior information may be available or elicited. In Section 4, we explore the ability of the model to learn about the missingness parameters using artificial data scenarios. We then present the results from the motivating example in Section 5 and end with a discussion.

2. Motivating example: haloperidol versus placebo for treatment of schizophrenia

We illustrate our framework using a meta-analysis of 17 RCTs comparing haloperidol with placebo in the treatment of schizophrenia [17]. Trials involving schizophrenic patients often report high levels of missing data because of treatment side effects, poor treatment compliance and strict execution of the RCT protocols. For each study, i, we have for each arm, k, the number of observed events (defined as the number of patients 'improved'), $r_{i,k}$, the total number of individuals randomised in each arm, $n_{i,k}$, and the number of missing individuals, $m_{i,k}$, where $k = 1$ indicates the placebo arm and $k = 2$ the haloperidol arm. The number of individuals in each arm for whom an outcome was actually observed (the complete cases) is therefore $c_{i,k} = n_{i,k} - m_{i,k}$ (Table I). In this dataset, six of the 17 RCTs contained no missing data, and overall, there was proportionately more missing data in the control group, possibly because of lack of efficacy [16].

3. Methods

3.1. Pairwise meta-analysis in the absence of missing data

In the absence of missing outcome data, the pairwise meta-analysis model for binary outcomes [14,20,21] is as follows. The number of events in study i arm k, $r_{i,k}$, is assumed to have a binomial likelihood determined by the total number of individuals randomised to that arm, $n_{i,k}$, and the probability of an event occurring in all randomised individuals, $\pi^{\text{all}}_{i,k}$:

$$r_{i,k} \sim \text{Binomial}\left(\pi^{\text{all}}_{i,k}, n_{i,k}\right) \tag{1}$$

Table I. Data from a meta-analysis of 17 trials comparing haloperidol with placebo for the treatment of schizophrenia [19].

	Placebo				Haloperidol			
Trial	No. of events	No. missing	No. of complete cases	Total no. randomised	No. of events	No. missing	No. of complete cases	Total no. randomised
i	$r_{i,1}$	$m_{i,1}$	$c_{i,1}$	$n_{i,1}$	$r_{i,2}$	$m_{i,2}$	$c_{i,2}$	$n_{i,2}$
1	18	0 (0%)	51	51	25	2 (4%)	50	52
2	20	34 (50%)	34	68	29	22 (32%)	47	69
3	2	1 (3%)	30	31	12	1 (3%)	29	30
4	0	0 (0%)	12	12	3	0 (0%)	12	12
5	3	0 (0%)	22	22	10	0 (0%)	21	21
6	1	0 (0%)	15	15	11	0 (0%)	19	19
7	4	1 (4%)	25	26	7	1 (4%)	25	26
8	3	0 (0%)	13	13	8	0 (0%)	17	17
9	14	2 (3%)	64	66	19	2 (3%)	64	66
10	0	0 (0%)	10	10	1	1 (10%)	9	10
11	0	0 (0%)	13	13	11	3 (8%)	34	37
12	2	0 (0%)	11	11	20	0 (0%)	29	29
13	7	18 (62%)	11	29	17	11 (38%)	18	29
14	0	1 (7%)	13	14	4	0 (0%)	14	14
15	0	1 (13%)	7	8	2	0 (0%)	16	16
16	1	0 (0%)	12	12	11	0 (0%)	12	12
17	0	1 (3%)	29	30	9	1 (3%)	29	30

The logit link function is used to model $\pi_{i,k}^{\text{all}}$ to ensure that the probabilities lie on (0,1) as follows:

$$\text{logit}\left(\pi_{i,k}^{\text{all}}\right) = \begin{cases} \mu_i & k = 1 \\ \mu_i + \delta_i & k = 2 \end{cases} \tag{2}$$

where μ_i are the study-specific log-odds of the outcome on treatment 1 (the control) and are treated as unrelated nuisance parameters.

For a fixed effects model, we assume each study i to be estimating the same underlying treatment effect. For a random effects model, we instead assume that the study-specific treatment effects, δ_i, are exchangeable and come from a common normal population of treatment effects. Thus,

$$\begin{array}{ll} \delta_i \sim N(d, \sigma^2) & \text{Random effects model} \\ \delta_i = d & \text{Fixed effects model} \end{array} \tag{3}$$

where d is the pooled log-odds ratio and σ is the between studies standard deviation.

3.2. *Sensitivity analyses in the presence of missing data*

When data are missing, different analyses can be carried out to assess the sensitivity of results to varying assumptions on the missing outcomes. A complete case analysis can be carried out using the standard pairwise meta-analysis model (Equations 1–3) where the denominator $n_{i,k}$ in Equation 1 is replaced by the number of complete cases $c_{i,k}$. Similarly, best-case and worst-case scenarios can be obtained by replacing $r_{i,k}$ in Equation 1 by

$$r_{i,k}^{\text{best}} = \begin{cases} r_{i,k} & \text{if } k = 1 \\ r_{i,k} + m_{i,k} & \text{if } k = 2 \end{cases} \text{ and } r_{i,k}^{\text{worst}} = \begin{cases} r_{i,k} + m_{i,k} & \text{if } k = 1 \\ r_{i,k} & \text{if } k = 2 \end{cases}$$

Finally, an analysis where all missing data are assumed to be failures (non-events) can be obtained using Equations 1–3 directly.

3.3. *General framework to account for missing data*

We extend the standard meta-analysis model described earlier to account for the uncertainty introduced by missing outcome data. We propose a pattern-mixture model where the outcome is modelled conditional on whether or not it is missing [22]. The approach assumes that the data are MAR with non-ignorable missingness [23]. Our general model is illustrated in a directed acyclic graph in Figure 1. We assume that $m_{i,k}$, the number of missing observations in arm k of study i, has a binomial likelihood:

$$m_{i,k} \sim \text{Binomial}\left(q_{i,k}, n_{i,k}\right) \tag{4}$$

where $q_{i,k}$ is the probability of being missing. Then, conditional on being observed, the number of events, $r_{i,k}$, are assumed to have a binomial likelihood with the total number of patients observed, $c_{i,k}$, as the denominator:

$$r_{i,k} \sim \text{Binomial}\left(\pi_{i,k}^{\text{obs}}, c_{i,k}\right) \tag{5}$$

where $\pi_{i,k}^{\text{obs}}$ is the probability of an event conditional on an individual being observed.

The model for the probability of an event for the population of all individuals (whether observed or missing) in arm k of trial i, $\pi_{i,k}^{\text{all}}$, is as set out in Equation 1. Our challenge is therefore to link the parameters we can estimate from the observed data, $\left\{\pi_{i,k}^{\text{obs}}, q_{i,k}\right\}$, to the key parameter of interest, $\pi_{i,k}^{\text{all}}$. This is achieved by defining a 'missingness' parameter $\theta_{i,k}$, which describes our belief about the missingness process and allows us to write $\pi_{i,k}^{\text{obs}}$ as a function of $\left\{\pi_{i,k}^{\text{all}}, q_{i,k} \text{ and } \theta_{i,k}\right\}$ (Figure 1), so that the observed data can be used to estimate $\pi_{i,k}^{\text{all}}$, the key parameter of interest.

Various options for the missingness parameter $\theta_{i,k}$ have been proposed [15, 17]. If we have no prior information on the missingness process, then the most straightforward option is to set $\theta_{i,k} = \pi_{i,k}^{\text{miss}}$ the probability of an event in the individuals with missing outcomes, which can be given a flat prior to reflect the uncertainty in this probability. If we do have prior information on the missingness process (e.g. from

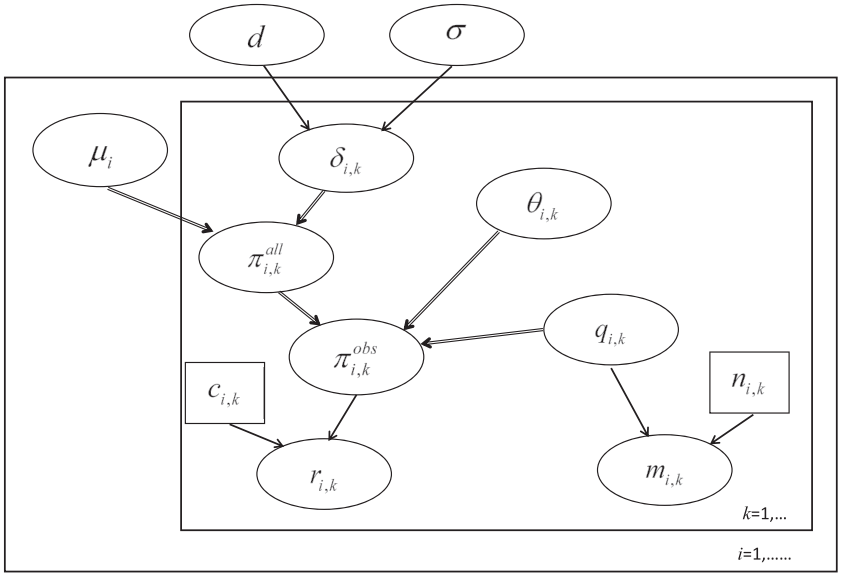

Figure 1. Directed acyclic graph (DAG) of our missing data framework: $\mu_{i,k}$ is the study-specific log-odds of the outcome in treatment 1 (the control), $\delta_{i,k}$ is the log-odds ratio, $\pi^{\text{all}}_{i,k}$ = p(success), $\pi^{\text{obs}}_{i,k}$ = p(success | observed), $r_{i,k}$ is the number of observed events, $c_{i,k}$ is the number of observed outcomes, $\theta_{i,k}$ is the missing data parameter, $q_{i,k}$ = p(missing), $m_{i,k}$ is the number of missing outcomes and $n_{i,k}$ is total number randomised. Ellipses denote stochastic parameters or observed data. Small boxes denote constants. The two large boxed ('plates') represent indexing of studies i and treatments k. Single line arrows denote stochastic relationships, and double-line arrows denote logical relationships. Note that $c_{i,k} = n_{i,k} - m_{i,k}$; however, we omit the logical relationship between these parameters in the DAG, as $c_{i,k}$ is constant conditional on the observed number of missing outcomes.

experts), then it may be easier to use other missingness parameters that allow beliefs to be expressed on the likelihood of an event in the missing individuals *relative* to the observed individuals. We first describe our approach for the case where $\theta_{i,k} = \pi^{\text{miss}}_{i,k}$. We then describe how the framework can also incorporate alternative definitions for $\theta_{i,k}$ in which we may have, or be able to elicit, prior information and show how to link $\pi^{\text{obs}}_{i,k}$ to $\left\{\pi^{\text{all}}_{i,k}, q_{i,k} \text{ and } \theta_{i,k}\right\}$ for some specific definitions of $\theta_{i,k}$ that have been used in the literature.

3.3.1. Probability of success given missing, $\pi^{\text{miss}}_{i,k}$. When $\theta_{i,k} = \pi^{\text{miss}}_{i,k}$, we can write the probability of an event in all individuals as a weighted average of an event in the missing and observed individuals:

$$\pi^{\text{all}}_{i,k} = q_{ik}\pi^{\text{miss}}_{i,k} + (1 - q_{i,k})\pi^{\text{obs}}_{i,k}$$

Rearranging this formula gives us the following linking equation:

$$\pi^{\text{obs}}_{i,k} = \frac{\pi^{\text{all}}_{i,k} - \pi^{\text{miss}}_{i,k} q_{i,k}}{(1 - q_{i,k})} \tag{6}$$

Equations 2–6, with either a fixed or random effects model for the δ_i, determine the model.

Beta priors are a natural choice for probability parameters, such as $\pi^{\text{miss}}_{i,k}$, because they are constrained to (0,1). When no prior information is available, (unrelated) Beta(1,1) priors for each $\pi^{\text{miss}}_{i,k}$ will appropriately reflect the uncertainty in the probability of an event in the missing individuals.

If the observed data do not provide any information on the missingness mechanism, then the uncertainty in the prior will be propagated to the treatment effect estimate. In this case, values for $\theta_{i,k}$ are simulated 'forward' from their prior (as in probabilistic sensitivity analysis commonly used in economic evaluation), as opposed to the deterministic sensitivity analyses where values for $\theta_{i,k}$ are assumed fixed at a particular value. However, in some cases, the information available on the relative treatment effects from studies where few or no data are missing, will allow the prior distribution for $\pi^{\text{miss}}_{i,k}$ to be updated, so that we 'learn' about the missingness mechanism from the data. This approach means that trials with

no missing values will have relatively more influence on the estimated treatment effect than those with large amounts of missing data.

3.3.2. Alternative definitions of the missingness parameter, $\theta_{i,k}$. Although $\theta_{i,k} = \pi_{i,k}^{\text{miss}}$ is a natural parameter on which to place an uninformative prior, it is not a natural parameter on which to elicit informative priors, because it is an absolute, rather than relative measure. If prior information is available on the missingness process, then it may be easier to elicit that information using an alternative definition for the missingness parameter. In this section, we describe some alternative missingness parameters and give the linking equation to replace Equation 6, in our general framework.

3.3.3. $\theta_{i,k}$ = Informative missingness odds ratio. The IMOR [5, 15] is defined as the ratio of the odds of the outcome among participants for whom the outcome is missing to the odds of the outcome among observed participants:

$$IMOR_{i,k} = \varphi_{i,k} = \frac{\left(\pi_{i,k}^{\text{miss}} \Big/ \left(1 - \pi_{i,k}^{\text{miss}}\right)\right)}{\left(\pi_{i,k}^{\text{obs}} \Big/ \left(1 - \pi_{i,k}^{\text{obs}}\right)\right)}$$

Appendix A shows that the linking equation can be written as

$$\pi_{i,k}^{\text{obs}} = \frac{-\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right) - \sqrt{\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right)^2 - 4\pi_{i,k}^{\text{all}}(1-q_{i,k})(1-\varphi_{i,k})}}{2(1-q_{i,k})(1-\varphi_{i,k})} \quad (7)$$

Equations 2–5 and 7, with either a fixed or random effect model for δ_i, determine the model. Normal priors on $\log(IMOR_{i,k})$ are a natural choice; however, sensible and informative priors are required for robust results to be obtained.

3.3.4. $\theta_{i,k}$ = Response probability ratio. Defining $\omega_{i,k}$ as the response probability ratio (RPR) [18], we obtain

$$RPR_{i,k} = \omega_{i,k} = \frac{p(\text{observed}|\text{ success})}{p(\text{observed}|\text{ failure})}$$

Applying Bayes' rule and rearranging, we obtain the linkage function:

$$\pi_{i,k}^{\text{obs}} = \frac{\omega_{i,k}\left(\pi_{i,k}^{\text{all}} \Big/ 1 - \pi_{i,k}^{\text{all}}\right)}{1 + \omega_{i,k}\left(\pi_{i,k}^{\text{all}} \Big/ 1 - \pi_{i,k}^{\text{all}}\right)} \quad (8)$$

Equations 2–6 and 8, with either a fixed or random effect model for δ_i, determine the model. Normal priors on $\log(\omega_{i,k})$ are a natural choice.

3.3.5. θ_{ik} = Success probability ratio. Another option could be to instead define $\theta_{i,k} = \rho_{i,k}$ as the ratio of the probability of success given a subject was missing to the probability of success given a subject was observed

$$\rho_{i,k} = \frac{\pi_{i,k}^{\text{miss}}}{\pi_{i,k}^{\text{obs}}}$$

Substituting into Equation 6 and rearranging, we obtain the linkage function:

$$\pi_{i,k}^{\text{obs}} = \frac{\pi_{i,k}^{\text{all}}}{1 - q_{i,k}(1 - \rho_{i,k})} \quad (9)$$

Equations 2–6 and 9, with either a fixed or random effect model for δ_i, determine the likelihood. Normal priors on $\log(\rho_{i,k})$ are a natural choice.

3.3.6. More structured models for the missingness parameter. Earlier, we have assumed independent but identical priors for $\theta_{i,k}$ across studies and arms. These can be modelled in a variety of more structured ways. Here, we give some suggested structural assumptions that could be made, but note that these should

be informed by expert opinion, and different sets of assumptions will be appropriate in different applications. Prior distributions could be (i) the same across trial arms but different between trials, (ii) different between trial arms but the same between studies or (iii) different between arms and studies. Furthermore, these different prior distributions across arm and study may be independent, hierarchical or common across trials/arms.

Estimates of treatment effect will be affected by the structure of the missingness model, including prior distributions, as this will determine how much can be 'learnt' about the missingness parameters. The stronger the assumptions made in the priors, the more potential for learning, but the validity of the results relies on the validity of the assumptions. Therefore, the choice of model structure should be an informed one and be a part of any elicitation exercise to obtain priors for the missingness parameters [24]. For example, if from clinical experience, those who drop out of an active treatment arm are expected to have stopped taking the medication (non-compliers); it may be considered reasonable to use a model structure where there is an arm-specific missingness parameter, but that the missingness parameter on the active arm is set equal (or similar) to that for *observed* outcomes in a placebo arm [25].

3.4. Priors and implementation

For all models, we assign a Uniform(0,1) prior distribution to the probability of being missing $q_{i,k}$, and a Normal(0,100^2) prior distribution for the trial-specific baselines μ_i. In random effects models, we give a Uniform(0,5) prior for the between studies standard deviation, σ. However, it is well documented that results can be sensitive to the form of prior assumed for random effects variances [26]. We therefore explored the sensitivity of the results to the prior for σ using the alternative priors: Uniform(0,2) distribution for σ and also an informative log Normal(-2.34, 1.62^2) prior distribution for σ^2, as suggested for a meta-analysis of a pharmacological treatment versus placebo/control where the outcome is subjective (as is the case here) by Turner *et al.* [27]. Possible priors for the missingness parameters, $\theta_{i,k}$, have already been indicated in Section 3.3; the specific priors that are used in the applications are described in Sections 4 and 5, and we return to consider prior specification further in the discussion.

We estimated the model using Markov Chain Monte Carlo simulation implemented in WinBUGS 1.4.3 [28]. The WinBUGS code to fit our missing data models for each of the proposed definitions of the missingness parameter can be found in Appendix B. The stability of the parameter estimates, Brooks–Gelman–Rubin statistics, auto-correlation and level of Monte Carlo (MC) error was assessed for each model to determine convergence and the simulation sample size for inference [21].

3.5. Model fit and comparison

To assess the fit of an individual model, we calculated the posterior mean of the residual deviance, $\bar{D}_{\text{res}}$. We sum over data points (conditional on being observed) in the calculation of residual deviance, so that models with $\bar{D}_{\text{res}}$ close to the number of observed independent data points (conditional on being observed) can be considered an adequate fit to the observed data, whereas models with $\bar{D}_{\text{res}}$ much bigger than this indicate a lack of fit [3, 14, 21, 29].

4. Exploring uncertainty and learning

We explored the effect of different proportions of missing data and distribution of missing data across trial arms on the resulting estimates of $\pi_{i,k}^{\text{miss}}$ using two fictitious data scenarios. We chose an illustrative meta-analysis of 11 trials with a binary outcome (mortality) [30] for which a fixed effect model is known to be appropriate [30] and manipulated the data in two ways. In the first scenario, we investigate the effect of fitting our model to a dataset where missing outcomes were not associated with the outcome or treatment arm. We therefore removed 20% of outcomes evenly across arms for eight of the 11 trials.

In the second scenario, we investigate the effect of fitting our model to a dataset where missing outcomes were associated with outcome and treatment arm, but only some trials had missing data. We expect that there is more scope for learning about the treatment effect in this scenario because (i) the missing data depends on outcome and arm, and (ii) there are more trials with complete data to help identify the missingness parameters. We removed 20% of outcomes from each arm of some of the trials (trials 2, 5, 6, 8, 9 and 11), but the proportion of events missing differed between arms for some studies. See Appendix C for the data used in these scenarios.

4.1. Models

We carried out a fixed effects meta-analysis on each of our two scenarios listed in Section 4. For each scenario, a complete case model was run and the results compared with those obtained for our miss-

Table II. Results from artificial data scenarios.

Scenario	Model	Posterior mean [a] residual deviance	Posterior median odds ratio (95% CI)
Scenario 1: Missingness not associated with outcome or treatment arm	Complete case analysis	19.7	0.80 (0.62, 1.04)
	Missing data framework, prior: $\pi_{i,k}^{\text{miss}} \sim \text{Beta}(1,1)$	19.4	0.80 (0.50, 1.23)
Scenario 2: Missingness associated with arm and outcome	Complete case analysis	26.0	0.74 (0.58, 0.95)
	Missing data framework, prior: $\pi_{i,k}^{\text{miss}} \sim \text{Beta}(1,1)$	18.5	0.80 (0.57, 1.11)

[a]Compare with 22 data points. Values larger than this are indicative of lack of fit.

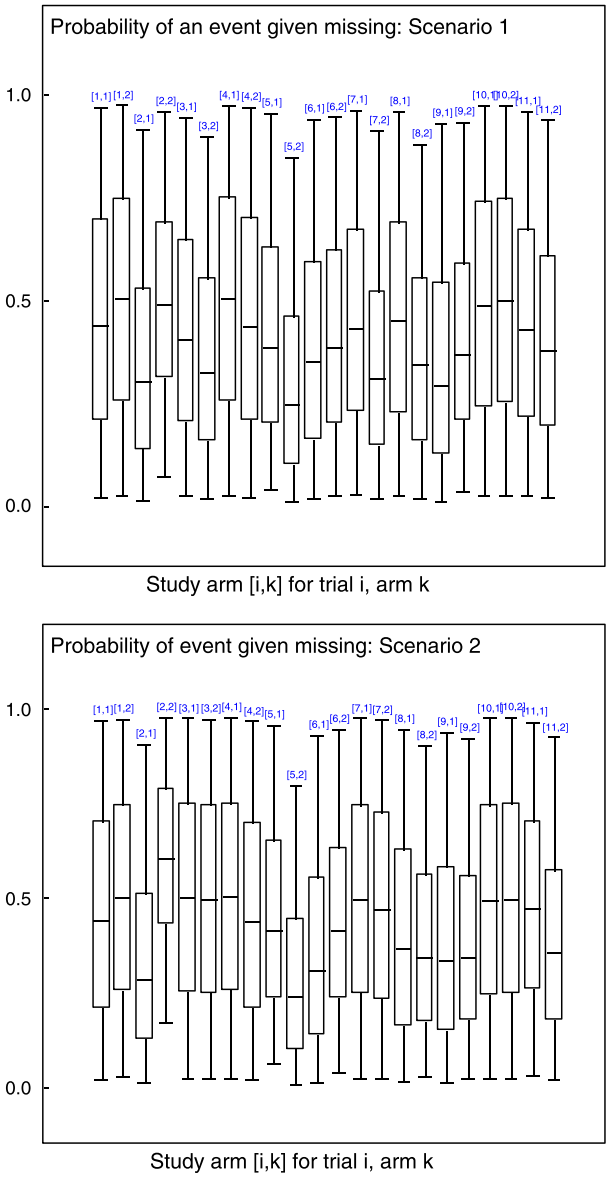

Figure 2. Box-plots of posterior distribution for $\pi_{i,k}^{\text{miss}}$ for each study arm $[i,k]$ for trial i, arm k. Plotted for scenarios 1 and 2, respectively. Boxes represent inter-quartile range, the whiskers represent the 95% CI and line within the box represents the median.

ing data model (Equations 2–6), where we put priors on the probability of an event, given missing, $\pi_{i,k}^{\text{miss}} \sim \text{beta}(1, 1)$.

4.2. Results from artificial data scenarios

4.2.1. Scenario 1: Missingness not associated with outcome or treatment arm. According to the posterior mean residual deviance, there is little difference between the model fit of the complete case model and the Beta(1,1) missing data model (Table II). In terms of treatment effect, the odds ratio is the same in the two models (median OR 0.80 in both the complete case analysis and in the missing data model), as we would expect because data were assumed MAR. However, the CIs around the treatment effect estimate from the missing data model are wider (Table II), reflecting the uncertainty induced by missing data. Figure C.1 in Appendix C shows the posterior mean and 95% CIs for the probability of missing, $q_{i,k}$, which is close to 0 in studies with no missing data (studies 1, 4 and 10), and approximately 0.2 in the other studies, as expected because missingness was introduced by removing approximately 20% of the observed outcomes evenly across arms. Figure 2 shows box-plots of the estimated posterior distribution for $\pi_{i,k}^{\text{miss}}$. It can be seen that whilst there is a large degree of uncertainty in these parameters, in some cases, the posterior distribution has moved away from the prior (Beta(1,1) centred on 0.5), showing that the model has 'learnt' about the p(success | missing) parameter (Figure 2), albeit only weakly.

4.2.2. Scenario 2: Missingness associated with arm and outcome. According to the posterior mean residual deviance, the missing data model is a better fit than the complete case model (Table II). In terms of treatment effect, the odds ratio is different between the two models (median OR 0.74 in the complete case analysis and 0.80 in the missing data model), which together with the model fit results, suggests that the missing data model allows a bias adjustment that explains the lack of fit seen in the complete case model. As in scenario 1, the CI around the treatment effect estimate is wider in the missing data model, reflecting the uncertainty induced by missing data. As with scenario 1, the proportion of missing data is well estimated (Figure C.1 in Appendix C). It can be seen from the box-plots of the estimated posterior distributions for $\pi_{i,k}^{\text{miss}}$ shown in Figure 2 that although there is a large degree of uncertainty in these parameters, there are some study arms where the posterior has moved away from the prior. This suggests that the model has 'learnt' about the $\pi_{i,k}^{\text{miss}}$ parameters, producing a bias-adjusted estimated of treatment effect that accounts for the uncertainty introduced by missing data.

5. Application to schizophrenia example

5.1. Prior distributions for $\pi_{i,k}^{miss}$

We begin by considering the case where there is no prior information on the missingness pattern, and so we use independent Beta(1,1) priors for each $\pi_{i,k}^{\text{miss}}$ as the base case. We then illustrate the impact of using more informative priors. The priors chosen are just illustrative; however, in practise, they would be based on clinical expert opinion. We first explore the impact of increasing the certainty in the prior to Beta(5,5), and then of changing the central estimate by using a Beta(3,7). We then look at the case where there are two different priors for the different arms Beta(3,7) for arm 1 (placebo) and Beta(5,5) for arm 2 (haloperidol). Finally, we use an informative N(0,1) prior on the log(*IMOR*), to illustrate the impact of a prior belief of MAR (log(*IMOR*) = 0). For comparison, we also present the standard sensitivity analyses proposed for missing data: complete case, best-case and worst-case scenarios and assuming all missing data are failures.

All other prior distributions are as described in Section 3.4.

5.2. Results

For each model, we used at least a 20 000 iteration burn-in period with at least 60 000 sampling iterations. The posterior mean residual deviance [29, 31] suggests that a random effects model is a better fit to the data than a fixed effects model in all cases (Table III). Within the different fixed effects models, all of the missing data models seem to be a better fit than the standard complete case analysis (Table III). Trial 2, which is one of the trials with the highest proportion of missing data, has the greatest improvement in model fit, with each arm contributing about 4 to the residual deviance in the complete case model, and only about 1 in the missing data model. Table I shows that the majority of the trials have only a small amount of missing data. This is likely to be the reason why the estimate of treatment effect

Table III. Model fit results from applying the standard models and our missing data model to the schizophrenia meta-analysis example.

		Fixed effect models Posterior mean residual deviance		Random effects models Posterior mean residual deviance	
		Event data (given observed)[a]	Number missing[a]	Event data (given observed)[a]	Number missing[a]
Standard models	Complete cases	62.9	NA	32.6	NA
	All missing are failures	60.4	NA	33.5	NA
	Best-case scenario	66.1	NA	32.3	NA
	Worst-case scenario	115.4	NA	33.0	NA
Missing data framework models priors	$\pi_{i,k}^{\text{miss}} = 0$	57.7	48.9	33.5	46.8
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(1,1)$	46.8	49.4	30.7	46.2
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(5,5)$	50.6	49.5	30.4	46.1
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(3,7)$	53.0	49.1	31.2	46.3
	$\pi_{i,1}^{\text{miss}} \sim \text{Beta}(3,7)$ and $\pi_{i,2}^{\text{miss}} \sim \text{Beta}(5,5)$	52.7	47.8	31.1	45.9
	$\log(IMOR_{i,k}) = \log\left(\varphi_{i,k}\right) \sim \text{Normal}(0,1)$	56.5	48.0	32.7	47.6

[a]Compare with 34 data points. Values larger than this are indicative of lack of fit.

Table IV. Treatment effect estimates results from applying the standard models and our missing data model to the schizophrenia meta-analysis example.

		Fixed effect models Posterior median (95%CI)	Random effects models Posterior median (95%CI)	
		Odds ratio	Odds ratio	Between studies standard deviation
Standard models	Complete cases	3.75 (2.66, 5.37)	8.51 (3.63, 30.66)	1.41 (0.70, 2.80)
	All missing are failures	3.65 (2.65, 5.09)	7.60 (3.48, 26.80)	1.27 (0.58, 2.66)
	Best-case scenario	6.16 (4.40, 8.74)	12.92 (5.42, 48.38)	1.44 (0.72, 2.86)
	Worst-case scenario	1.69 (1.26, 2.29)	4.39 (1.60, 14.96)	1.81 (1.11, 3.15)
Missing data framework models priors	$\pi_{i,k}^{\text{miss}} = 0$	3.59 (2.63, 4.97)	7.06 (3.38, 21.63)	1.19 (0.56, 2.41)
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(1,1)$	3.76 (2.67, 5.37)	5.80 (2.96, 15.34)	1.01 (0.37, 2.12)
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(5,5)$	3.35 (2.44, 4.62)	5.41 (2.83, 14.05)	1.05 (0.46, 2.12)
	$\pi_{i,k}^{\text{miss}} \sim \text{Beta}(3,7)$	3.46 (2.52, 4.77)	5.92 (3.02, 15.71)	1.08 (0.50, 2.15)
	$\pi_{i,1}^{\text{miss}} \sim \text{Beta}(3,7)$ and $\pi_{i,2}^{\text{miss}} \sim \text{Beta}(5,5)$	3.89 (2.83, 5.37)	6.53 (3.37, 17.64)	1.05 (0.48, 2.16)
	$\log(IMOR_{i,k}) = \log\left(\varphi_{i,k}\right) \sim \text{Normal}(0,1)$	4.09 (2.86, 5.90)	8.60 (3.64, 35.41)	1.39 (0.66, 3.00)

is so similar between the complete case model and the Beta(1,1) missing data model (Table IV). The box-plot of the posterior distribution for p(success | missing) for trial 2 with a high amount of missingness illustrates that the model is 'learning' about p(success | missing) as the posterior differs from the Beta(1,1) prior (Figure 3). The missing data model is providing a bias-adjusted estimate, but because of the small amount of missing data, this estimate is very similar to that produced by the complete case analysis.

Within the random effects models, there is little difference in fit between any of the models (Table III). These results were expected as random effects models are flexible models that generally fit very well. However, in the random effects missing data models, the between-study heterogeneity estimates are reduced compared with the standard models, with the exception of the log(*IMOR*) model (Table IV). There are two possible explanations for this. Firstly, increased uncertainty in the study-specific treatment effect estimates is consistent with lower heterogeneity, and secondly, adjusting for the missingness mechanism allows for some of the heterogeneity to be explained. In the random effects missing data models, there was also a reduction in the estimate of treatment effect, with the exception of the log(*IMOR*) model (Table IV). These results are likely because there is only a small amount of missing data, but proportionately, there was more missing data in the control arms of the trials. Increasing the certainty in the prior, by using a Beta(5,5) distribution, resulted in a reduction in the estimate of treatment effect in both the fixed and random effects models. Similarly, changing the central estimate of the prior, using a Beta(3,7) distribution, reduced the estimate of treatment effect

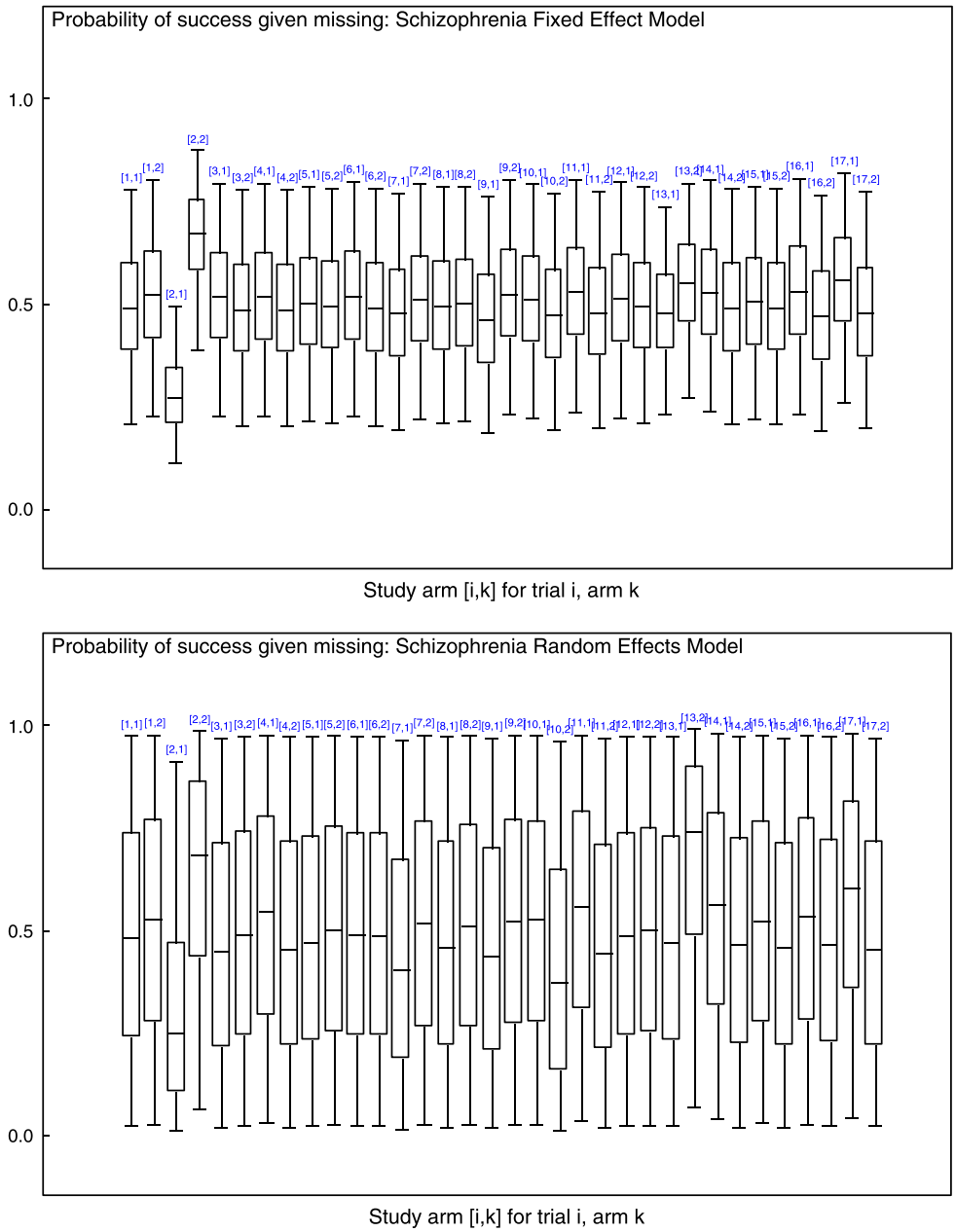

Figure 3. Box-plots of posterior distribution for $\pi_{i,k}^{\text{miss}}$ for each study arm $[i, k]$ for trial i, arm k in the schizophrenia meta-analysis [19]. Plotted for fixed and random effect models where a missing data model is used with Beta(1,1) priors for $\pi_{i,k}^{\text{miss}}$. Boxes represent inter-quartile range, the whiskers represent the 95% CI and line within the box represents the median.

(Table IV). The log(*IMOR*) prior we used was informative and compatible with a belief of MAR. The results from this prior are therefore in line with those from a complete case analysis (Table IV). The meta-analysis results were sensitive to the choice of prior placed on the heterogeneity parameter; however, the same pattern in the results was found between the complete case and missing data models (results not shown).

Setting $\pi_{i,k}^{\text{miss}}$ equal to zero with no uncertainty makes the missing data model similar to an 'all missing are failures' analysis. These two models make the same underlying assumption but are not exactly the same. One fixes the observations, stating that all missing outcomes were failures, whereas the other fixes the parameter that generates the data stating that the probability of success in the missing individuals is zero. The two models produce similar results, although the between-trial standard deviation is reduced when accounting for the missingness parameter in the missing data model.

6. Discussion

We have proposed and illustrated a conceptually and computationally simple Bayesian framework for handling missing binary outcome data in pairwise meta-analysis. Our model provides an estimate of treatment effect that accounts for uncertainty due to missing outcomes, to provide information for clinicians and local and national health policy makers. The missingness parameter can be parameterised in a variety of formats, including the IMOR, and can be viewed as a one-stage version of the model presented by White *et al.* [5]. This differs from the model presented in White *et al.,* [12] which is a one-stage, hierarchical selection model, rather than the pattern-mixture model proposed here.

Our framework does, however, require prior information on the missingness parameter. In the absence of any such prior information, we recommend using a Beta(1,1) prior on the probability of an event given missingness. Note that by using such a flat prior, the missingness parameters may only be weakly identified. However, in the absence of prior information, the uncertainty introduced by missing data will be reflected in the relative effect estimate. If prior information is available, we recommend that it is used. In particular, if the outcome is a rare event, then implicitly, there is prior information that the probability of an event is low regardless of whether the outcome is missing or not, and this information should be incorporated.

A difficulty that arises in all methods that utilise informative prior beliefs is how these beliefs should be obtained [32]. If the aim is to simply reflect the uncertainty around missingness, then a flat, uninformative prior will achieve this. One option for eliciting informative prior distributions for unknown outcome parameters is to ask experts to assign a total weight of 100 over several categories of potential options for the unknown parameter and then to model this distribution of weights [12]. A similar approach was used by Turner *et al.* to elicit priors based on asking experts to mark a range of believable values on a scale [33], whilst Mavridis *et al.* suggest a method based on eliciting the lowest and highest values an expert would conceivably expect [34]. White *et al.* [5] suggest that a more realistic prior about the missingness might be obtained by providing an expert with a plausible estimate for the probability of success given a subject was observed (perhaps based on a previous study) and from that, asking their opinion on the probability of success given a subject was missing. From this, the IMOR can then be calculated [5]. Other potential options for $\theta_{i,k}$ in our framework (i.e. probability success ratio and probability success difference) could also be elicited this way. Ideally, prior beliefs about unknown parameters should be elicited from experts within the appropriate field. It could be argued that in this context, the investigators of each individual trial may be in the best position to inform the prior distributions required for our models; however, priors elicited from individuals who were involved in the design, running or analysis of the trials may not be objective. It may be more appropriate to elicit prior distributions based on the beliefs of the clinical members of the meta-analysis team who were not involved personally with any of the trials being used in the analysis. To date, the only study we are aware of that has elicited priors to inform missingness parameters in a meta-analysis is the work of White *et al.* [9], although there is work on eliciting opinion for individual trials [12, 35, 36]. There is clearly a need for further empirical elicitation studies, to explore the value of this approach in different clinical areas (where reasons for missingness may be very different) and where the proportion of missingness varies (there may be less value in elicitation if there is high power within the meta-analysis to estimate the missingness pattern). The papers identified in the systematic review may describe methods to adjust for missing data, which may indicate whether there is likely to be informative missingness, and the potential direction of any bias and thus providing further prior information. A recent review of how missing data was handled in RCTs published in top medical journals [37] noted that about 86% of the included studies stated a reason for missingness, with some RCTs reporting specific explanations for missingness. These may be useful to inform priors. Note that, as the posterior distribution for $\theta_{i,k}$ will in many cases be heavily dependent on the prior assigned to it, it is essential that sensitivity analyses are carried out varying this prior distribution. In addition, the ability of the model to learn about the missingness parameter will diminish with increasing heterogeneity, although the missingness mechanism may potentially explain heterogeneity, as seen in the artificial example.

Another issue that can occur in all meta-analyses is poor reporting of trials. Trials may not specifically mention missing data, but this does not necessarily mean that outcomes were observed for all participants; there may just have been poor reporting. In some cases, even after close inspection of the text, tables and CONSORT diagram, it may still be unclear whether there are missing data. Because we estimate the probability of a subject being missing, $q_{i,k}$, for all studies, our model will estimate $q_{i,k}$ as non-zero even when there are no missing data in that trial, and there will be uncertainty in this estimate (Figure C.1

Table V. Examples of prior structures for the missingness parameter.

	Model structure for $\theta_{i,k}$	Arm-specific	Trial-specific	Example
1	Independent	No	No	$\theta_{i,k} \sim f(\cdot)$
2	Independent	Yes	No	$\theta_{i,k} \sim f_k(\cdot)$
3	Independent	No	Yes	$\theta_{i,k} \sim f_i(\cdot)$
4	Independent	Yes	Yes	$\theta_{i,k} \sim f_{i,k}(\cdot)$
5	Hierarchical	No	No	$\theta_{i,k} \sim f(\boldsymbol{\eta})\,,\ \boldsymbol{\eta} \sim g(\cdot)$
6	Hierarchical	Yes	No	$\theta_{i,k} \sim f(\boldsymbol{\eta}_k)\,,\ \boldsymbol{\eta}_k \sim g_k(\cdot)$
7	Hierarchical	No	Yes	$\theta_{i,k} \sim f(\boldsymbol{\eta}_i)\,,\ \boldsymbol{\eta}_i \sim g_i(\cdot)$
8	Identical	No	No	$\theta_{i,k} = \eta\,,\ \eta \sim f(\cdot)$
9	Identical	Yes	No	$\theta_{i,k} = \eta_k\,,\ \eta_k \sim f_k(\cdot)$
10	Identical	No	Yes	$\theta_{i,k} = \eta_i\,,\ \eta_i \sim f_i(\cdot)$
11	Identical	Yes	Yes	$\theta_{i,k} = \eta_{i,k}\,,\ \eta_{i,k} \sim f_{i,k}(\cdot)$

Given prior information on the missingness mechanisms, combinations of these prior structures can be chosen for subsets of the trials included in the meta-analysis.
f and g represent generic distributions with given parameters values or depending on a vector of parameters η or scalar η.

in Appendix C). In some applications, this may be appropriate, as some trials may not report missing observations when there are in fact some. Again, clinical input can be helpful to assess whether there are likely to be unreported missing observations. The Cochrane risk of bias tool [1] includes an assessment of risk of bias due to missing data. In theory, these assessments could be incorporated into the modelling framework either as covariates or to inform priors; however, this is a notoriously difficult dimension of risk of bias to assess.

The Bayesian approach we have taken is essential in order to be able to specify prior distributions for the missingness parameters and also has the advantage of easy extension from pairwise to network meta-analysis [14]. In network meta-analysis, because of the consistency assumption, there is the potential to be able to estimate the missingness parameter with more certainty, although the power to do so may still be low [16, 38]. This power could be increased by making some of the structural assumptions set out in Table V. This is an area for further work. The Bayesian approach also allows for the results from the meta-analysis to be easily integrated into decision models [39, 40]. In cost-effectiveness analyses, estimating the uncertainty in the optimal decision is vital to inform policy decisions. Our framework provides a treatment effect estimate that reflects the uncertainty introduced by missing data, can correct for bias induced by the missing data and can be viewed as a probabilistic sensitivity analysis.

This paper has restricted to the case where the outcome of interest is binary. Dealing with missing data where there are continuous outcome measures is more complex. Different studies may report results using different imputation methods (e.g. last observation carried forwards, baseline observation carried forwards or multiple imputation techniques), which raise challenges for evidence synthesis. Recent work [41] has extended the concept of the IMOR to continuous outcomes, defining the informative missingness difference of means and the informative missingness ratio of means, and proposed a pattern-mixture model for pairwise and network meta-analysis, using a two-stage estimation procedure. Extending our framework together with the ideas from Mavridis *et al.* [41] is an exciting area for further work.

In conclusion, the framework outlined here provides a simple method for the best utilisation of evidence produced from RCTs that accounts for any uncertainty induced by missing data, allows for the incorporation of prior information on a variety of different missingness metrics and fits easily into the wider evidence synthesis model for medical decision making [42].

Appendix A: The linking equation when the missingness parameter is the IMOR

By definition [17], the IMOR, which we shall denote $\varphi_{i,k}$, is

$$IMOR_{i,k} = \varphi_{i,k} = \frac{\left(\pi_{i,k}^{\text{miss}} \Big/ \left(1 - \pi_{i,k}^{\text{miss}}\right)\right)}{\left(\pi_{i,k}^{\text{obs}} \Big/ \left(1 - \pi_{i,k}^{\text{obs}}\right)\right)} \tag{A.1}$$

If $\varphi_{i,k} = 1$, then $\pi_{i,k}^{\text{obs}} = \pi_{i,k}^{\text{miss}}$, so that Equation (6) gives the linking equation $\pi_{i,k}^{\text{obs}} = \pi_{i,k}^{\text{all}}$. Otherwise, for $\varphi_{i,k} \neq 1$, rearranging (A.1) and substituting into Equation (6) gives the linking equation:

$$\pi_{i,k}^{\text{obs}} = \frac{-\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right) \pm \sqrt{\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right)^2 - 4\pi_{i,k}^{\text{all}}(1-q_{i,k})(1-\varphi_{i,k})}}{2(1-q_{i,k})(1-\varphi_{i,k})} \tag{A.2}$$

However, we need to identify which is the correct root to take. If we define (dropping the *ik* subscript for convenience)

$$\begin{aligned} A &= 1 - (\pi^{\text{all}} - q)(\varphi - 1) \\ B &= 2(1-q)(1-\varphi) \\ &\text{so that} \\ \pi^{\text{obs}} &= \frac{A\sqrt{A^2 - 2B\pi^{\text{all}}}}{B} \end{aligned} \tag{A.3}$$

then very careful consideration of the following scenarios shows that:

(1) If $A < 0$ and $\varphi < 1$ then $\pi^{\text{all}} < 0$, which is impossible.
(2) If $A > 0$ and $\varphi < 1$ then the negative root lies on the interval [0,1], whereas the positive root is always greater than 1. π^{obs} must lie between 0 and 1, and so we take the negative root.
(3) If $A < 0$, $\varphi > 1$ ($\Rightarrow B < 0$), then the negative root lies on the interval [0,1], whereas the positive root is always less than 0. We therefore take the negative root.
(4) If $A > 0$, $\varphi_{ik} > 1$ ($\Rightarrow B < 0$), then the negative root lies on the interval [0,1], whereas the positive root is always less than 0. We therefore take the negative root.

In all cases, we therefore take the negative root, giving the linking equation:

$$\pi_{i,k}^{\text{obs}} = \frac{-\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right) - \sqrt{\left(\left(q_{i,k} - \pi_{i,k}^{\text{all}}\right)(1-\varphi_{i,k}) - 1\right)^2 - 4\pi_{i,k}^{\text{all}}(1-q_{i,k})(1-\varphi_{i,k})}}{2(1-q_{i,k})(1-\varphi_{i,k})} \tag{A.4}$$

Appendix B: WinBUGS code

```
# Binomial likelihood, logit link, pairwise meta-analysis (2 treatments)

# Random effects model

Model                          # Program starts
{
for(i in 1:ns)                 # Loop through studies
{
delta[i,1] <- 0                # Treatment effect is zero for control arm
mu[i] ~ dnorm(0,.0001)         # Vague priors for all trial baselines
for(k in 1:2)                  # Loop through arms
{
r[i,k] ~ dbin(pi_obs[i,k], c[i,k])          # Binomial likelihood for the number of observed events
m[i,k] ~ dbin(q[i,k], n[i,k])               # Binomial likelihood for the number of missing observations
logit(pi_all[i,k]) <- mu[i] + delta[i,k]    # Model for linear predictor
pi_obs[i,k] <- max(0, min(1,(pi_all[i,k] - (pi_miss[i,k] * q[i,k])) / (1 - q[i,k])))          # Linking equation*
q[i,k] ~ dunif(0,1)            # Uninformative prior for the probability of being missing
pi_miss[i,k] ~ dbeta(1,1)      # Beta(1,1) prior distribution for p(success|missing)
rhat[i,k] <- pi_obs[i,k] * c[i,k]           # Expected value of the numerators (observed data)
dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))          # Deviance contribution (observed data)
+ (c[i,k]-r[i,k]) * (log(c[i,k]-r[i,k]) - log(c[i,k]-rhat[i,k])))
m0[i,k] <- m[i,k]+0.01*equals(m[i,k],0)
rhat_miss[i,k] <- q[i,k] * n[i,k]           # Expected value of the numerators (missing data)
dev_miss[i,k] <- 2 * (m0[i,k] * (log(m0[i,k])-log(rhat_miss[i,k]))    # Deviance contribution (missing data)
+ (n[i,k]-m0[i,k]) * (log(n[i,k]-m0[i,k]) - log(n[i,k]-rhat_miss[i,k])))
}
```

```
delta[i,2] ~ dnorm(d[2],prec)          # Trial specific LOR distributions
resdev[i] <- sum(dev[i,])              # Summed residual deviance contribution for this trial (observed data)
resdev_miss[i] <- sum(dev_miss[i,])          # Summed residual deviance contribution for this trial (missing data)
}
totresdev <- sum(resdev[])             # Total residual deviance (observed data)
totresdev_miss <- sum(resdev_miss[])         # Total residual deviance (missing data)
d[1] <- 0                              # Treatment effect is zero for reference treatment
d[2] ~ dnorm(0,.0001)                  # Vague prior for treatment effect
tau ~ dunif(0,5)                       # Vague prior for between trial SD
tau.sq <- tau*tau                      # Between trial variance = between trial SD^2
prec <- 1/(tau.sq)                     # Between trial precision = (1/between trial variance)
OR <- exp(d[2])                        # Convert to odds ratio scale = exponential(treatment effect)
}                                      # Program ends

#Data

list(ns=17)

r[,1]   m[,1]   n[,1]   c[,1]   r[,2]   m[,2]   n[,2]   c[,2]
18      0       51      51      25      2       52      50
20      34      68      34      29      22      69      47
2       1       31      30      12      1       30      29
0       0       12      12      3       0       12      12
3       0       22      22      10      0       21      21
1       0       15      15      11      0       19      19
4       1       26      25      7       1       26      25
3       0       13      13      8       0       17      17
14      2       66      64      19      2       66      64
0       0       10      10      1       1       10      9
0       0       13      13      11      3       37      34
2       0       11      11      20      0       29      29
7       18      29      11      17      11      29      18
0       1       14      13      4       0       14      14
0       1       8       7       2       0       16      16
1       0       12      12      11      0       12      12
0       1       30      29      9       1       30      29

END
```

*Linking equations for alternative models with other missingness parameters:

IMOR

```
pi_obs[i,k] <- max(0, min(1,(  (-((q[i,k] - pi_all[i,k])*(1 - phi[i,k]) - 1) - sqrt((pow(((q[i,k] - pi_all[i,k])*(1 - phi[i,k]) -
1),2)) - ((4*pi_all[i,k])*(1 - q[i,k])*(1 - phi[i,k])))  )  /  (2*(1 - q[i,k])*(1- phi[i,k]))   )))
```

Response probability ratio

```
a[i,k] <- omega[i,k]*(pi_all[i,k] / (1 - pi_all[i,k]))

pi_obs[i,k] <- max(0, min(1, (a[i,k]/ (1 + a[i,k]))  ))
```

Success probability ratio

```
a[i,k] <- (pi_all[i,k]) / (1 - ((q[i,k])*(1 - rho[i,k])))

pi_obs[i,k] <- max(0, min(1, (a[i,k])))
```

Appendix C Artificial data scenarios:

Original dataset

Trial	Arm 1		Arm 2	
i	No. of events $r_{i,1}$ (%)	Total no. randomised $n_{i,1}$	No. of events $r_{i,2}$ (%)	Total no. randomised $n_{i,2}$
1	3 (5.5%)	55	1 (1.8%)	55
2	10 (10.6%)	94	3 (3.2%)	95
3	40 (7.0%)	573	32 (5.7%)	565
4	2 (3.3%)	61	3 (4.8%)	62
5	16 (3.8%)	419	20 (4.8%)	421
6	5 (7.2%)	69	3 (4.2%)	71
7	5 (6.7%)	75	5 (6.7%)	75
8	59 (7.5%)	782	52 (6.6%)	790
9	5 (6.2%)	81	2 (2.5%)	81
10	16 (7.1%)	226	12 (5.3%)	225
11	8 (12.1%)	66	6 (8.5%)	71

Scenario 1 dataset: Missingness not associated with outcome or treatment arm

	Arm1						Arm2					
Trial	No. of events $r_{i,1}$	No. missing $m_{i,1}$	No. of complete cases $c_{i,1}$	Total no. randomised $n_{i,1}$	% events in obs	% events in miss	No. of events $r_{i,2}$	No. missing $m_{i,2}$	No. of complete cases $c_{i,2}$	Total no. randomised $n_{i,2}$	% events in obs	% events in miss
1	3	0	55	55	5.5		1	0	55	55	1.8	
2	8	19	75	94	10.7	10.5	2	19	76	95	2.6	5.3
3	32	115	458	573	7.0	7.0	26	113	452	565	5.8	5.3
4	2	0	61	61	3.3		3	0	62	62	4.8	
5	13	84	335	419	3.9	3.6	16	84	337	421	4.7	4.8
6	4	14	55	69	7.3	7.1	2	14	57	71	3.5	7.1
7	4	15	60	75	6.7	6.7	4	15	60	75	6. 7	6. 7
8	47	156	626	782	7.5	7.7	42	158	632	790	6.6	6.3
9	4	16	65	81	6.2	6.3	1	16	65	81	1.5	6.3
10	16	0	226	226	7.1		12	0	225	225	5.3	
11	6	13	53	66	11.3	15.4	5	14	57	71	8.8	7.1

Scenario 2 dataset: Missingness associated with arm and outcome

	Arm1						Arm2					
Trial	No. of events $r_{i,1}$	No. missing $m_{i,1}$	No. of complete cases $c_{i,1}$	Total no. randomised $n_{i,1}$	% events in obs	% events in miss	No. of events $r_{i,2}$	No. missing $m_{i,2}$	No. of complete cases $c_{i,2}$	Total no. randomised $n_{i,2}$	% events in observed	% events in missing
1	3	0	55	55	5.5		1	0	55	55	1.8	
2	10	19	75	94	13.3	0	1	19	76	95	1.3	10.5
3	40	0	573	573	7.0		32	0	565	565	5. 7	
4	2	0	61	61	3.3		3	0	62	62	4.8	
5	10	84	335	419	3.0	7.1	16	84	337	421	4.7	4.8
6	4	14	55	69	7.3	7.1	1	14	57	71	1.8	14.3
7	5	0	75	75	6. 7		5	0	75	75	6. 7	
8	59	156	626	782	9.4	0	42	158	632	790	6.6	6.3
9	4	16	65	81	6.2	6.3	2	16	65	81	3.1	0
10	16	0	226	226	7.1		12	0	225	225	5.3	
11	6	13	53	66	11.3	15.4	6	14	57	71	10.5	0

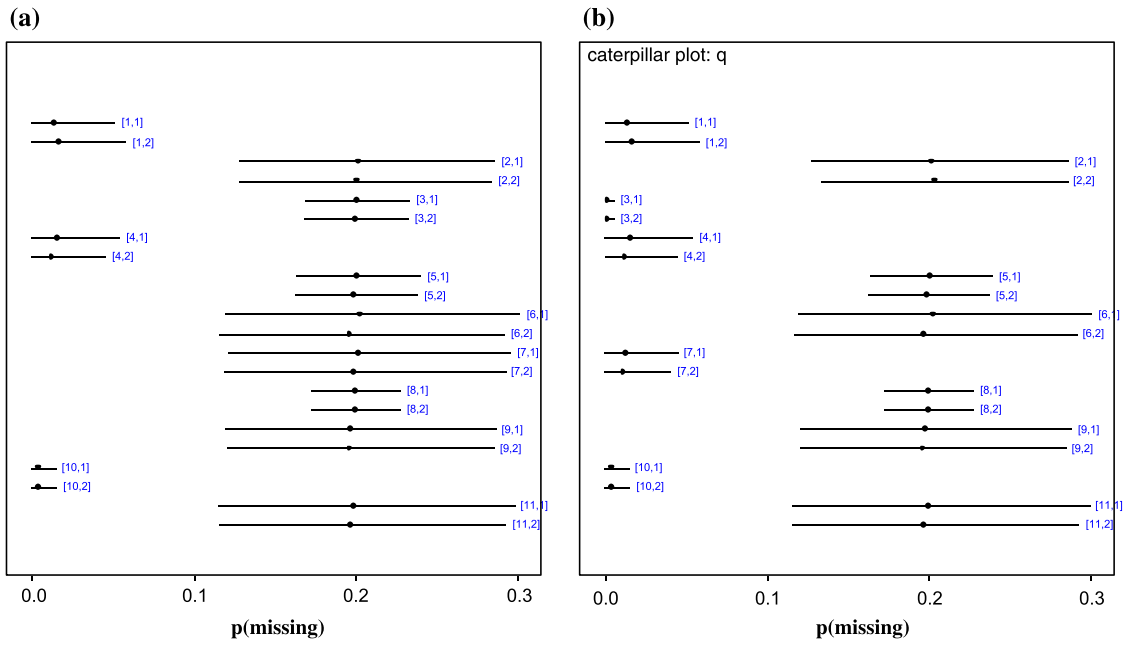

Figure C.1. Caterpillar plots showing posterior mean (dot) and 95% CIs (line) of p(missing) for each trial arm in the artificial data scenarios. Indices $[i, k]$ represent arm k ($k = 1$ = placebo; $k = 2$ = active treatment) of trial i. (a) Scenario 1 – trials 2, 3, 5, 6, 7, 8, 9 and 11 had 20% of outcomes removed (from the observed events) from each arm. (b) Scenario 2 – trials 2, 5, 6, 8, 9 and 11 had 20% of outcomes removed from each arm.

Acknowledgements

N. L. T. was supported by an NIHR Research Methods Fellowship; S. D., A. E. A. and N. J. W. were supported by the National Institute of Heath and Care Excellence (NICE), through the NICE Clinical Guidelines Technical Support Unit, University of Bristol. N. J. W. was also supported by an MRC Methodology Research Fellowship and the MRC ConDuCT Hub for Trials Methodology Research. We thank Dr. Vanessa Didelez, University of Bristol, for help on the interpretation of directed acyclic graphs.

References

1. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008]*, The Cochrane Collaboration. Wiley: Chichester, 2008.
2. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *British Medical Journal* 1996; **312**:71–72.
3. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 2: a generalised linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials*, 2011. (Available from http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm.) [Accessed on 12 March 2015].
4. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, Als-Nielsen B, Balk EM, Gluud C, Gluud LL, Ioannidis JPA, Schulz KF, Beynon R, Welton NJ, Wood L, Moher D, Deeks JJ, Sterne JAC. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Annals of Internal Medicine* 2012; **157**:429–438.
5. White I, Higgins J, Wood AM. Allowing for uncertainty due to missing data in meta-analysis – Part 1: two-stage methods. *Statistics in Medicine* 2008; **27**:711–727.
6. Little RJA, Rubin DB. *Statistical Analysis with Missing Data* 2nd ed. Wiley: New Jersey, 2002.
7. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*, 2013. (Available from http://www.nice.org.uk/article/pmg9/chapter/Foreword). [Accessed on 12 March 2015].
8. National Institute for Health and Clinical Excellence. *The guidelines manual, (November 2012)*, 2012. (Available from: http://publications.nice.org.uk/the-guidelines-manual-pmg6.) [Accessed on 12 March 2015].
9. Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *Journal of Biopharmaceutical Statistics* 2013; **23**:1352–1371.
10. Mallinckrodt CH, Lin Q, Lipkovich I, Molenberghs G. A structured approach to choosing estimands and estimators in longitudinal clinical trials. *Pharmaceutical Statistics* 2012; **11**:456–461.
11. Panel on handling missing data in clinical trials National Research Council. *The prevention and treatment of missing data in clinical trials*. The National Academies Press, 2010.

12. White IR, Carpenter J, Evans S, Schroter S. Eliciting and using expert opinions about dropout bias in randomized controlled trials. *Clinical Trials* 2007; **4**:125–139.
13. Mavridis D, Salanti G. A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research* 2013; **22**:133–158.
14. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence Synthesis for Decision Making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Medical Decision Making* 2013; **33**:607–617.
15. White IR, Wood A, Welton NJ, Ades AE, Higgins JPT. Allowing for uncertainty due to missing data in meta-analysis – Part 2: hierarchical models. *Statistics in Medicine* 2008; **27**:728–745.
16. Spineli LM, Higgins JPT, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. *Clinical Trials* 2013; **10**:378–388.
17. Higgins JPT, White IR, Wood A. Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clinical Trials* 2008; **5**:225–239.
18. Magder LS. Simple approaches to assess the possible impact of missing outcome information on estimates of risk ratios, odds ratios, and risk differences. *Controlled Clinical Trials* 2003; **24**:411–421.
19. Joy CB, Adams CE, Lawrie SM. Haloperidol versus placebo for schizophrenia. *Cochrane Database of Systematic Reviews* 2006; CD003082.
20. Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in Medicine* 1995; **14**:2685–2699.
21. Welton NJ, Sutton AJ, Cooper NJ, Abrams KR, Ades AE. *Evidence Synthesis for Decision Making in Healthcare*. Wiley: Chichester, 2012.
22. Little RJA. Pattern-mixture models for multivariate incomplete data. *Journal Of The American Statistical Association* 1993; **88**:125–134.
23. Daniels MJ, Hogan JW. *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman & Hall/CRC: Boca Raton, 2008.
24. Mason A, Richardson S, Plewis I, Best N. Strategy for modelling nonrandom missing data mechanisms in observational studies using Bayesian methods. *Journal of Official Statistics* 2012; **28**:279–302.
25. Mallinckrodt CH, Lin Q, Molenberghs M. A structured framework for assessing sensitivity to missing data assumptions in longitudinal clinical trials. *Pharmaceutical Statistics* 2012; **12**:1–6.
26. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis* 2nd ed. Chapman & Hall: London, 2004.
27. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JPT. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology* 2012; **41**:818–827.
28. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000; **10**:325–337.
29. 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**:583–616.
30. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; **331**:897–900.
31. Dempster AP. The direct use of likelihood for significance testing. *Statistics and Computing* 1997; **7**:247–252.
32. O'Hagan A, Buck CE, Alireza Daneshkhah J, Eiser R, Garthwaite PH, Jenkinson DJ, Oakley JE, Rakow T. *Uncertain judgements: eliciting experts' probabilities*. Wiley: Chichester, 2006.
33. Turner RM, Spiegelhalter DJ, Smith GCS, Thompson SG. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society (A)* 2009; **172**:21–47.
34. Mavridis D, Sutton A, Cipriani A, Salanti G. A fully Bayesian application of the Copas selection model for publication bias extended to network meta-analysis. *Statistics in Medicine* 2013; **32**:51–66.
35. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. *Journal Of The American Statistical Association* 2005; **100**:680–701.
36. Kadane JB, Wolfson LJ. Experiences in elicitation. *Journal of the Royal Statistical Society: Series D (The Statistician)* 1998; **47**:3–19.
37. Bell ML, Fiero M, Horton NJ, Hsu C-H. Handling missing data in RCTs: a review of the top medical journals. *BMC Medical Research Methodology* 2014; **14**:118.
38. Dias S, Welton NJ, Marinho VCC, Salanti G, Higgins JPT, Ades AE. Estimation and adjustment of bias in randomised evidence by using mixed treatment comparison meta-analysis. *Journal of the Royal Statistical Society (A)* 2010; **173**:613–629.
39. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: introduction. *Medical Decision Making* 2013; **33**:597–606.
40. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. *Medical Decision Making* 2013; **33**:671–678.
41. Mavridis D, White IR, Higgins JPT, Cipriani A, Salanti G. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. *Statistics in Medicine* 2014; **34**:721–741.
42. Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision making: current approaches. *Journal of the Royal Statistical Society (A)* 2006; **169**:5–35.

