DECISION ANALYTICAL ECONOMIC MODELLING WITHIN A
BAYESIAN FRAMEWORK

Nicola J. Cooper, Alex J. Sutton and Keith R. Abrams
Department of Epidemiology and Public Health, University of Leicester

TECHNICAL REPORT 02-03
Department of Epidemiology and Public Health,
University of Leicester, UK

Correspondence address:
Dr. Nicola J. Cooper, Department of Epidemiology and Public Health,
University of Leicester, 22-28 Princess Road West, Leicester, LE1 6TP
e-mail: njc21@le.ac.uk
ABSTRACT

Economic evaluation of health care interventions based on decision analytic modelling can generate valuable information for health policy decision-makers. However, the usefulness of the results obtained depends on the quality of the data input into the model; that is, the accuracy of the estimates for the costs, effectiveness and transition probabilities between the different health states of the model. The aim of this paper is to demonstrate how the individual components required for decision analytical modelling (i.e. systematic review incorporating meta-analyses, estimation of transition probabilities, evaluation of the model and sensitivity analysis) may be addressed simultaneously in one coherent Bayesian model evaluated using Markov Chain Monte Carlo simulation implemented in the specialist Bayesian statistics software WinBUGS.

The approach described is applied to two illustrative examples:

1) The prophylactic use of antibiotics for caesarean section patients; and

2) The use of taxanes for the second-line treatment of advanced breast cancer.

The advantages of using the Bayesian statistical approach outlined compared to the conventional classical approaches to decision analysis include the ability to: (i) perform all necessary analyses, including all intermediate analyses (e.g. meta-analyses) required to derive model parameters, in a single coherent model; (ii) incorporate expert opinion either directly or regarding the relative credibility of different data sources; (iii) use the actual posterior distributions for parameters of interest (opposed to making distributional assumptions necessary for the classical formulation); and (iv) incorporate uncertainty for all model parameters.

KEYWORDS: Decision model, Markov model, Bayesian methods, Meta-analysis; Caesarean section; Breast cancer; Taxanes; Markov Chain Monte Carlo.
1. INTRODUCTION

Decision-analytical models are widely used in economic evaluations of health care interventions with the objective of providing information to allow scarce health care resources to be allocated efficiently\textsuperscript{1,2}. Such models have a range of uses\textsuperscript{3} including the synthesis of data from a variety of sources (often using meta-analysis) to produce the cost or cost-effectiveness results of interest. They are often used to evaluate the complex process usually associated with the implementation of health care interventions\textsuperscript{4}. Examples of instances when decision modelling techniques may be of value include the extrapolation of primary data beyond the endpoint of a trial or to make comparisons between treatments for which no ‘head-to-head’ trials exist\textsuperscript{5,6}.

Decision trees, such as the one depicted in Figure 1, provide a simple way to structure problems of decision making under uncertainty\textsuperscript{7} whilst describing the major factors involved. More complicated decision trees can be represented in the form of Markov models (an example is given in Figure 4). Such models provide a technique for analysing events that are repeatable (e.g. relapses of a chronic disease such as multiple sclerosis, arthritis and asthma), or events that play out over an extended period of time (e.g. the progression of cancer)\textsuperscript{7,8}. To evaluate decision-analytical models, estimates need to be acquired for the costs and health outcomes of the various pathways through the model together with the probability of their occurrence. It should not be forgotten that the usefulness of the results obtained from such models depends on the source and quality of the estimates input into the model\textsuperscript{7}.

Decision-analytical models are sometimes based on primary data collection, but more often rely on published or other secondary sources for cost and effectiveness information\textsuperscript{9}. Systematic review methods are a formal and replicable approach to identifying and
summarising existing evidence\textsuperscript{10}. Where data permits, quantitative synthesis of the evidence, often referred to as meta-analysis\textsuperscript{11}, can be conducted within a systematic review. The use of systematic methods for evidence synthesis are desirable for the evaluation of health care to be truly evidence-based and hence information for decision models should be based on such rigorous methods. However, very little has been written on the methods of systematic reviews (including meta-analysis) to be used for the synthesis of evidence for an economic decision\textsuperscript{9}. It is currently unclear what sources of evidence should be included in systematic reviews informing decision models, e.g. should both RCTs and observational studies be included in the same analysis\textsuperscript{12}. This is a particularly pertinent issue in economic decision modelling since in addition to clinical effectiveness, estimates of costs and probabilities are required.

Probabilistic decision models used in economic evaluation are almost exclusively analysed using classical statistical approaches (two of the rare exceptions are Parmigiani et al. (1997)\textsuperscript{13} and Fryback et al. (2001)\textsuperscript{14}). Such models place probability distributions on parameters where there is uncertainty in their true value. These can be derived from the results of individual studies, or, more desirably, the results of systematic reviews. Parametric distributional assumptions are necessary when specifying parameter uncertainty, but occasions do exist when such assumptions may be inappropriate, such as when events are rare. The Bayesian analyses described herein relax the need for some of these distributional assumptions. Further, by combining the synthesis and decision process into one coherent model, the Bayesian approach described here incorporates uncertainty in incidental model parameters, that need estimating but are not of direct interest in the decision model, that is often ignored in classical analyses. (The between study variance parameters in meta-analyses are examples of these, as will be shown later.) An additional advantage of the Bayesian
approach is that the correlation between parameters induced by the fact that the same data sources e.g. systematic review may be used to propagate different parts of the model is automatically accounted for.

Fryback et al. (2000) outlined how simple probabilistic decision analytical models may be evaluated using Bayesian methods. The aim of this paper is to extend this work by describing a method whereby the whole process (i.e. systematic review incorporating meta-analyses, estimation of transition probabilities, evaluation of model and sensitivity analysis) may be combined into a single coherent Bayesian model. The ease of applying such a method is demonstrated through the use of two illustrative examples:

1) The prophylactic use of antibiotics for caesarean section patients to reduce the incidence of wound infections; and

2) The use of taxanes for the second-line treatment of advanced breast cancer compared to conventional treatment.

Example one illustrates the process of inputting the pooled estimates obtained from a systematic review (i.e. meta-analyses), together with their associated uncertainty, directly into a probabilistic decision analytical model. Example two illustrates the situation whereby the pooled estimates obtained from a systematic review, together with their associated uncertainty, are initially converted to transition probabilities and then applied to a probabilistic Markov decision model. The incorporation of subjective/expert prior beliefs is also illustrated in the latter example.

The paper is organised with section 2 introducing the rationale for Bayesian statistics and section 3 providing an outline of the methods used to construct and evaluate decision models. Sections 4 and 5 show how the Bayesian methods outlined in section 3 can be applied in
practice through the use of the two illustrative examples. Finally, section 6 discusses some of
the issues that have been raised and outlines specific areas for further research.
2. BAYESIAN METHODS

Bayesian methods can be considered as an alternative to the classical approach to statistical inference. Such methods are becoming more frequently used largely due to the increased feasibility of their implementation made possible by the advances made in computer power, and the development of user-friendly software.\(^{15}\)

A key difference between the two approaches is that Bayesian methods allow the incorporation of information external to that included directly in the model; for example, the expert beliefs of clinicians on the likely treatment effect of a new intervention (see section 5). Such information is specified in a prior distribution, \(p(\theta)\), and is combined with the study data, \(y\), in the form of the likelihood \(p(y|\theta)\). The posterior inferential result, \(p(\theta|y)\), is then obtained as an update of the prior distribution due to the observation of the data. That is,

\[
p(\theta|y) \propto p(\theta)p(y|\theta)
\]

This is known as Bayes’ Theorem\(^{16}\). The incorporation of subjective/informative a priori beliefs is not a requirement of a Bayesian analysis, since ‘vague’ or non-informative prior distributions can be used which convey very little or no information relative to the data. In the main analyses presented here all prior distributions placed on model parameters are intended to be ‘vague’; however, sensitivity analyses are used to check that this is indeed the case and that the results are stable across a range of different prior distributions. The Bayesian graphical modelling approach adopted here holds an appeal since it provides a flexible modelling framework, allowing us to venture beyond the confines of analyses provided in standard statistical packages and account fully for all forms of model estimation uncertainty\(^{17}\). Further advantages of the Bayesian over the classical approach include i) all analyses follow directly from the posterior distribution (no separate theories of estimation, testing, multiple comparisons, etc. are needed), ii) posterior distributions permit inference of functions of
parameters (e.g. tail probabilities of parameters), which may be more difficult in the classical framework, iii) the ability to make predictions (in this context to future patients) via the predictive distribution, and iv) the direct interpretability of constructed credible intervals (CrI) (these are analogous to classical confidence intervals but are more intuitively interpretable i.e. a 95% CrI ranging from 1.0 to 2.0 for a model parameter implies that the true value for that parameters lies between 1.0 and 2.0 with 95% probability). For a much fuller introductory exposition of the Bayesian approach to statistics see Lee 1997.

For certain simple analyses the posterior distributions (on which inferences are usually made) are analytically tractable, meaning that they can be written down in standard statistical notation. When this is not possible a Bayesian solution can still be obtained through the use of simulation methods, such as Markov Chain Monte Carlo (MCMC) methods. (Hence asymptotic or approximate methods often relied upon using classical inference are not necessary.) Within the broad range of MCMC simulation methods, one method Gibbs Sampling has been increasingly used in applied Bayesian analyses. The appeal of Gibbs sampling is that it can be used to estimate marginal posterior distributions by drawing sample values randomly from the full conditional distributions of each parameter conditional on all other parameters and the data. A full description of this approach is available elsewhere. Until recently, sampling from one conditional distribution, especially when not of closed form, would have taken a considerable amount of computer programming. Fortunately, the necessary computation routines are now freely available in the software package WinBUGS (website: http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml), which makes Bayesian methods available for more routine use by applied researchers and only requires the actual model to be specified (i.e. not the full conditional distribution). It is worth noting at this point, that distributions are defined in WinBUGS in terms of their mean and precision, where
precision is equal to the reciprocal of the variance. While MCMC methods offer great flexibility, they cannot simply be used as a ‘black box’. Convergence of the simulations, achieved when the sampler is truly sampling from the true conditional posterior distributions required, needs to be established. In order to achieve this, the sampler needs to be ‘burnt-in’, which means running and discarding an initial number of samples before retaining samples from which inferences and estimation can be based. This procedure is necessary otherwise biased estimates may be produced. Several methods have been described to assist with this assessment of convergence and should be used routinely as part of the sensitivity analysis. A further check to ascertain whether the sampler is moving around the whole parameters space is to the start the MCMC chains with different extreme starting values.

For a much more thorough description of the use of Bayesian methods in Health Technology Assessment (HTA) see Spiegelhalter et al. (1999) and references therein.
3. BAYESIAN DECISION ANALYTICAL MODELLING IN ECONOMIC

EVALUATION

The application of Bayesian methods to economic evaluations is a relatively new area of research. In particular, Fryback et al applied Bayesian methods to evaluate a decision analytical model to assess the cost-effectiveness of tissue plasminogen activator compared to streptokinase in acute myocardial ischemia. In this section, we build on their work by considering the whole process associated with decision analytical models. The process is divided into four stages:

1) Systematic review incorporating meta-analyses

2) Estimation of transition probabilities

3) Sensitivity analyses (for data and model specification)

4) Evaluation of the model

3.1 Systematic review incorporating meta-analyses

As discussed in the introduction, decision analytical cost-effectiveness models are usually based on cost and effectiveness information extracted from published or other secondary sources and relevant information should be identified and summarised using formal and replicable systematic review methods. Meta-analysis, which is concerned with the analysis of the data extracted from published or other secondary sources, uses quantitative methods to estimate overall measures of association or effect and provides a framework to assess the sensitivity of the results to possible threats to their validity such as publication bias and study quality. A large literature exists on the methods for meta-analysis generally (see Sutton et al 2000 and references therein) and specifically from a Bayesian context. In all meta-analyses it is important to investigate the potential sources of between-study heterogeneity (i.e. systematic differences between the trial results included in a meta-analysis).
differences may be due to patient-level covariates or trial-level characteristics. The underlying risk of the event occurring (which is a measure of patients underlying risk of the event in question) for all patients in each trial is one potential source of between-study heterogeneity\textsuperscript{28,29}. Investigating this relationship between the treatment effect and underlying risk is important as the existence of such a relationship will affect decisions about which patients should be treated and may identify patient groups in which the treatment is not cost-effective or even harmful. These relationships may be explored by using analysis techniques such as meta-regression\textsuperscript{11}.

### 3.2 Estimation of transition probabilities

The pooled estimates obtained from the meta-analyses can sometimes be applied to the decision analytical model directly (e.g. relative risk of wound infection in example 1), or, as is more often the case, need to be transformed into the necessary format required for the model. For instance, event rates will often need converting into transition probabilities\textsuperscript{30}. This transformation is required and outlined in example 2. Note that there will usually be uncertainty associated with the estimation of transition probabilities. If these probabilities have been derived through transformations of other parameters, themselves estimated with uncertainty, the resulting uncertainty of the transition probability is automatically incorporated into the model. If not using MCMC methodology, the delta method\textsuperscript{31}, which is an asymptotic approximation, would need to be used in order to calculate the variance of the transformed parameter(s)

### 3.3 Sensitivity analyses (of model inputs and model specification)

An essential tool for studying and validating the behaviour of any model is a sensitivity analysis. Such an analysis is used to assess the robustness of the results to specific methods
used and assumptions made; that is, the more the result obtained is materially unchanged by sensible sensitivity analyses, the more confident we can be in the final results of the model. Sensitivity analysis in standard decision modelling has been discussed in detail elsewhere\textsuperscript{3,6}. In addition, assessment of the meta-analysis components of our model, such as the impact of study selection criteria on the results, should be addressed. For example, the impact of variability in the study populations, interventions administered, outcome measure definitions, and study quality could all have an influence on the model parameter estimates\textsuperscript{11}. The fact that all analyses are conducted in one model specification facilitates such sensitivity analyses, as changes in intermediate analyses (e.g. meta-analyses) automatically propagate throughout the model making assessment of the impact on the overall results immediate and transparent.

There are additional issues, which require exploration, when using Bayesian methods\textsuperscript{32}. Perhaps the most important is the sensitivity of the results on the prior distributions placed on the model parameters; it is always wise to check the results over a plausible range of prior beliefs. Particularly critical is the assessment of prior distributions placed on variance components (such as the between-study heterogeneity parameter in a random effects meta-analysis model (see section 4)) since it has been shown that placing truly non-informative prior distributions on such parameters is difficult\textsuperscript{33}. An assessment of this kind is demonstrated for the example given in section 4. In addition, when using MCMC methods to evaluate the model, very different starting values for the simulation parameters should be applied to check convergence of the MCMC sampler has been achieved and that all features of the joint posterior distribution have been identified (e.g. multiple modes)\textsuperscript{20}. As an aid to good practice when reporting Bayesian methods, a set of guidelines, known as Bayeswatch, have been developed\textsuperscript{32}. 
3.4 Evaluation of the model

To evaluate probabilistic decision models, simulation methods are adopted. Such analytical methods provide a means to imitate a real-life system, especially when other analyses are too mathematically complex or too difficult to reproduce. Without using simulation, a decision model will only reveal a single outcome, which is likely to be either the most likely or average scenario. In a classical framework Monte Carlo simulation methods are often adopted for model evaluation. Similarly, Bayesian decision models evaluated within WinBUGS use Markov Chain Monte Carlo (MCMC) simulation methods as described in section 2. It is important to note that in our Bayesian implementation all preliminary analyses, such as meta-analyses and transformation of variables, are also evaluated using MCMC methods within the single model framework.
4. ILLUSTRATIVE EXAMPLE 1: Cost implications of using prophylactic antibiotics in caesarean section to reduce the incidence of wound infections.

There is evidence to suggest that the incidence of wound infection following a caesarean section can be significantly reduced by a very short prophylactic course of antibiotics at the time of the operation\textsuperscript{34}. Women who experience wound infection have a longer postnatal hospital stay, and require more intensive nursing care, together with antibiotic therapy and more laboratory tests than would normally be the case after a caesarean section. Therefore, in this particular example, we are interested in the cost implication of introducing prophylactic antibiotics at a local maternity hospital. Based on a hypothetical scenario, previously considered elsewhere\textsuperscript{35}, it is assumed that the local maternity hospital performed on average 750 caesarean sections in 1997 of which 60 (8\%) experienced a postoperative wound infection. Further, it is assumed that i) the cost of a course of prophylactic antibiotics, including consumables, is fixed at £10.00; ii) a consultant’s time to administer the treatment ranges, on average, from 4 to 7 minutes at a cost of £1 per minute\textsuperscript{36}; iii) the mean length of stay in hospital for a woman who has a caesarean section and develops a wound infection is 8.8 days (standard error (s.e.) 0.55)\textsuperscript{37} at a cost of £262 per day (1998/9); and iv) the mean length of stay for those women who experience no infection, is 6.7 days (s.e. 0.33) \textsuperscript{37} at a cost of £173 per day (1998/9).

The decision tree for this problem is outlined in Figure 1 and details of its implementation and evaluation are outlined in detail below.

4.1 Systematic review incorporating meta-analyses

A systematic review of the effectiveness of prophylactic antibiotic treatment on infectious complications in women undergoing caesarean delivery has recently been published as a
Cochrane Review\textsuperscript{34}. The review identified 61 studies, which compared the incidence of wound infection in women undergoing caesarean delivery who were administered prophylactic antibiotics and a placebo group using meta-analysis. The relevant data are presented in Table 1. A re-analysis of these data using a Bayesian meta-analysis model, on the relative risk (RR) scale (defined as the ratio of the risk of events in each arm)\textsuperscript{38}, is described below. Note although meta-analysis of binary outcomes (such as the occurrence of wound infection) could be analysed on other scales, such as the odds ratio or risk difference\textsuperscript{38}, the RR is used here since this is the outcome required to derive the transition probability of interest (i.e. probability of wound infection given prophylactic antibiotics) as described in section 4.2. It may seem like unnecessary effort to re-analyse the data since the original paper provided a pooled meta-analysis result, however, the Bayesian re-analysis below offers several advantages (as discussed above) that include incorporating the uncertainty associated with the between-study heterogeneity parameter, assessment of the importance of baseline risk on treatment effect, and allows the application of sensitivity analysis. A further issue is the presence of sparse data. As Table 1 indicates, since the occurrence of wound infection is relatively rare and many of the studies are small, there are numerous studies in which zero infections were observed in one or both arms of the trials. Such studies pose a problem since the RR is undefined. In a classical meta-analysis those studies in which there are zero events in both arms would be excluded for the analysis, whilst a continuity correction factor would be required to permit inclusion of those studies in which there are zero events in one arm\textsuperscript{39}. Fortunately, a Bayesian meta-analysis method, that circumvents the need for continuity corrections, by modelling directly the event rates in each arm using binomial distributions, has been developed\textsuperscript{40}. This also avoids the assumption of normality of the effect measure in each trial (necessary in the classical analysis), which may be inappropriate when some of the trials for inclusion in the meta-analysis are small, or observed risks are close to 0 or 1.
The Bayesian random effects meta-analysis model specification used is outlined below.

\[
\begin{align*}
    r_i^c & \sim \text{Binomial}(n_i^c, p_i^c) & r_i^t & \sim \text{Binomial}(n_i^t, p_i^t) & i = 1, \ldots, 61 \\
    \mu_i & = \log( p_i^c ) & \log( p_i^t ) & = \mu_i + \text{min}( \delta_i, -\log( p_i^c ) ) \\
    \delta_i & \sim \text{Normal}(\bar{\delta}, \tau_2^2) & \text{log}(\text{RR}_{\text{antibiotics}}) & = \exp(\bar{\delta}) \\

    p_i^c & \sim \text{Uniform}(0,1) & \bar{\delta} & \sim \text{InverseGamma}(0.001,0.001) \\

\end{align*}
\]

where for the \(i\)th trial, \(r_i^c\) out of \(n_i^c\) have a wound infection in the placebo group and \(r_i^t\) out of \(n_i^t\) have a wound infection in the prophylactic antibiotics group; \(p_i^c\) and \(p_i^t\) are the estimated infection rates in the placebo and prophylactic groups respectively; \(\mu_i\) is the natural logarithm of the event rate in the placebo group; \(\delta_i\) is the estimated \(\log(\text{RR}_{\text{antibiotics}})\). \(\bar{\delta}\) is the pooled \(\log(\text{RR}_{\text{antibiotics}})\) and \(\tau_2^2\) is the between-study variance parameter often referred to as a heterogeneity parameter as it estimates how much variation exists between the results of the different studies. The last line of the equation 2 specifies reasonably ‘vague’ prior distributions for all unknown parameters in the model. Note that the Inverse-Gamma(0.001,0.001) prior distribution is popular choice for a vague prior distribution on a variance parameter\(^{41}\). It can be seen that the WinBUGS specification follows directly from the algebraic exposition given above. Unfortunately the model specification is rather non-intuitive, but a full account of this method and Bayesian meta-analyses models more generally implemented within WinBUGS are available elsewhere\(^{40,42}\). Importantly, this model can be extended to examine whether the treatment effect varies with patients’ underlying risk of wound infection as discussed in section 3.1; explicitly this involves replacing

\[
\delta_i \sim \text{Normal}(\bar{\delta}, \tau^2)
\]

with
where $\gamma$ is the regression slope associated with the influence of underlying baseline risk on the treatment effect, $\bar{\mu}$ is the mean of the natural logarithm of the event rate in the placebo group and $\delta_i^*$ is the estimated log($RR_{antibiotics}$) adjusted for regression to the mean. A vague prior distribution is specified for $\gamma$. This modification is discussed fully by Warn et al. 2002\textsuperscript{40}.

### 4.2 Estimation of transition probabilities

To obtain the expected probability of wound infection in the local hospital if the new intervention is introduced ($p_2$), the pooled $RR_{antibiotics}$ can be applied to local hospital data on the proportion of wound infections in women undergoing caesarean delivery without the new intervention (i.e. without prophylactic antibiotics, $p_1$) using the following formula:

$$p_2 = p_1 \times RR_{antibiotics}$$

This equation expressing the relationship between model parameters can be written directly into the WinBUGS program. This is particularly appealing since the uncertainty in $p_2$ does not have to be calculated explicitly as discussed in section 3.2.

### 4.3 Sensitivity analyses (for data and model specification)

An assessment of the robustness of the results to several assumptions and modelling issues follows.

**Assessment of the impact of model assumptions:** Model assumptions, such as inclusion/exclusion thresholds based on study quality, can be addressed within the same model as the main analysis provided the data is inputted appropriately. In the original meta-analysis\textsuperscript{34}, the quality of the studies is assessed using the standard Cochrane criteria of
adequacy of allocation concealment\textsuperscript{43}. Using this quality assessment, which grades from high quality studies (A) through to poor quality studies (D), as indicated in Table 1, the affect of study methodological quality on the overall results of the analysis is investigated.

This list of factors is certainly not exhaustive but illustrates some of the principles required to address common assumptions in an evaluation such as the one presented here. Further sensitivity analyses, which could be valuable, include investigating the impact of the ‘type’ of caesarean section (elective or non-elective, Table 1) and/or the assumptions necessary to estimate resource use costs.

Assessment of the influence of prior distributions: It is intended that non-informative prior distributions be used in this analysis. Specifying such distributions for variance component parameters, such as $\tau^2$, is difficult as discussed in section 3.3. For this reason, it is always advisable to evaluate the model for a range of prior distributions in a sensitivity analysis to confirm these parameters are not exerting undue influence on the overall conclusions\textsuperscript{33}. This can be accomplished within a single model specification by first replicating the data $j$ times to allow $j+1$ different prior distributions to be fitted. In this example, three differently shaped prior distributions for $\tau^2$ are used to check the sensitivity of the results to the choice of prior distribution for this parameter:

1) Inverse-Gamma $(0.001, 0.001)$ distribution on $\tau^2$.

2) Normal $(0, 1.0^6)$ distribution truncated at zero (i.e. a half-normal distribution) on $\tau$.

3) Uniform $(0, 50)$ distribution on $\tau$.

Note, informative prior distributions, could also be used in a sensitivity analysis. These could be derived from different peoples’ beliefs or derived empirically from previous analyses (e.g. a prior for $\tau^2$ could be derived having examined estimates for this parameter from previous
meta-analyses. This would inform whether model conclusions are robust over a range of prior beliefs currently held.

4.4 Evaluation of the model

Following preliminary test runs, it was decided to use an initial run of 5,000 iterations as a ‘burn in’, in order to achieve convergence (these values were discarded) with inferences based on a further 20,000 sample iterations. As a further step to ensure model convergence has been achieved, two WinBUGS runs with dispersed starting values (i.e. defining two sets of very different initial values and running two chains simultaneously) were carried out. If the two chains give very similar results, which they do, this helps to confirm the convergence of the MCMC sampler.

4.5 Model Results

Primary analysis: In the underlying risk analysis, there was no evidence of an association between treatment effect and underlying risk as can be observed from Figure 2. That is, the regression slope coefficient (i.e. the parameter of primary interest) is \( -0.03 \) (SD=0.22) with a 95% credible interval ranging from \(-0.45\) to \(0.44\).

The results from the model, together with the resulting posterior distributions, for each of the model parameters and the cost outcome are presented in Table 2. The pooled relative risk of a wound infection is \(0.29\) (95% Credible Interval \(0.20\) to \(0.39\)); that is the risk of a wound infection in the prophylactic antibiotic treatment group is about 29% of that in the standard care (no prophylactic antibiotics) group. The cost implication of using prophylactic antibiotics to reduce the incidence of wound infection in women undergoing caesarean delivery is a mean cost reduction of £29.37 (95% CI £16.79 to £44.33). Indeed, the
probability that the cost reduction is less than zero is negligible as all 20,000 sampled values estimated the reduction as being positive. As discussed in section 2, the ability to make probability statements such as this is another advantage of the Bayesian approach.

*Sensitivity analysis:* To assess the influence of prior distributions for the variance component parameters, $\tau^2$, three different prior distributions were applied (as described in section 4.3). As can be observed from Figure 3, the results of the model are found to be robust over this array of prior distributions.

Removing the three quality graded C studies from the meta-analysis, to assess the impact of study quality on the model results, has a minimal effect on the overall outcome (i.e. mean cost reduction of £29.00 (95% Credible Interval £16.54 to £43.71)).

In this second illustrative example we focus on the cost-effectiveness of docetaxel (a type of taxane) as a second-line treatment for metastatic breast cancer compared to conventional treatment (assumed here to be doxorubicin). To date these two treatments have only been compared directly in one published clinical trial\textsuperscript{45}. The model implemented in this illustrative example, is a simplified version of a four-stage Markov model details of which are reported elsewhere\textsuperscript{46}. The model has been simplified for illustrative purposes and considers only the short-term effectiveness of taxane use during the treatment cycles ignoring adverse events and long-term follow-up once treatment has stopped. As can be seen in Figure 4, the model consists of four health states:

i) **Response** - complete and partial (reduced by at least 50% in size) tumour disappearance – absorbing state (i.e. once entered, it is never left);

ii) **Stable** – no change in disease state;

iii) **Progressive** – tumour growth or spread to other sites; and

iv) **Death** – absorbing state (i.e. once entered, it is never left).

One cycle of the model represents 3 weeks to coincide with the chemotherapy treatment intervals. It is assumed that individuals are only allowed a maximum of 7 cycles of treatment due to problems with cumulative toxicities associated with the chemotherapy agents. Individuals are assumed to discontinue treatment before 7 cycles if they move into the ‘progressive’ disease state. The model is run for 7 treatment cycles only.

The type of analysis undertaken is a ‘cohort analysis’\textsuperscript{8}, which simulates the prognosis of a hypothetical cohort of 1,000 individuals on doxorubicin treatment and 1,000 individuals on docetaxel treatment. The Markov process commences with the entire cohort starting
treatment. With each cycle, the individuals are redistributed among the 2 transient (stable and progressive) and 2 absorbing disease states (response and death) according to the associated transition probabilities. This results in a new allocation of the cohort among the various health states for the subsequent cycle. Notice that individuals do not skip health states or move backwards (Figure 4); for example, transitions from ‘response to stable’ and ‘progressive to stable’ are not permitted. Ideally, the model terminates when the entire cohort is absorbed into the two absorbing health states; however, in this short-term model, focusing on the 7 treatment cycles only, this may not be the case.

The costs of the two treatments were obtained from the British National Formulary (BNF)\textsuperscript{47} as £1,355 and £440 for one treatment session on docetaxel and doxorubicin respectively. In this analysis the average cost of progressive disease is assumed to be £365 to cover palliative care (based on expert opinion). Obviously, this estimate will have uncertainty associated with it that should be incorporated into the analysis; although easy to do in principle, for the purposes of this analysis (and to keep it simple) it is ignored. Utility data was extracted from published literature and uniform distributions assumed over the range of estimates obtained.

5.1 Systematic review incorporating meta-analysis
A systematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast was recently published by Lister-Sharp et al. (2000)\textsuperscript{48}. The review identified four randomised controlled phase III trials of docetaxel use; compared to doxorubicin\textsuperscript{45}, mitomycin plus Vinblastine\textsuperscript{49}, methotrexate\textsuperscript{50} plus 5-fluorouracil and 5-fluorouracil plus navelbine\textsuperscript{51} respectively for the treatment of advanced breast cancer. For the purposes of this Markov model, it was necessary, given the available data, to extract effectiveness data on docetaxel from one set of published studies and the effectiveness of
doxorubicin from a different set of published studies: thus breaking randomisation. This is in contrast to example 1 where all the studies, included in the meta-analysis, compared the treatment of interest (i.e. prophylactic antibiotics) with placebo. The issue of breaking randomisation will be discussed further in section 6.

Non-comparative meta-analyses are carried out on the response rate, progression-free survival time and overall survival time for both docetaxel and doxorubicin. For the response rate, a binary variable, the meta-analyses are performed on the log odds scale as this should provide a measure which is approximately normally distributed. The Bayesian random effects meta-analysis model specification used is outlined below.

\[
\begin{align*}
  r_i &\sim Binomial(n_i, p_i) \quad \text{logit} (p_i) \sim Normal(\bar{A}, \tau^2) \quad i = 1, \ldots, m \\
  \bar{A} &\sim Normal(0,0.00001) \quad \bar{\sigma}^2 \sim InverseGamma(0.001,0.001)
\end{align*}
\]

where for the \(i\)th trial, \(r_i\) out of \(n_i\) have the event of interest; \(p_i\) is the estimate for the probability of the event occurring; \(\bar{A}\) is the pooled log odds and \(\tau^2\) is the estimate of between-study heterogeneity. Vague prior distributions are specified for all unknown parameters in the model (last line of equation 5). For the time variables the median estimates and their standard errors are used in the meta-analyses. The Bayesian random effects meta-analysis model specification used is as follows:

\[
\begin{align*}
  y_i &\sim Normal(\theta_i, \sigma_i^2) \quad \theta_i \sim Normal(\bar{A}, \tau^2) \quad i = 1, \ldots, m \\
  \bar{A} &\sim Normal(0,0.00001) \quad \bar{\sigma}^2 \sim InverseGamma(0.001,0.001)
\end{align*}
\]

where \(y_i\) is the effect size estimate of interest from the \(i\)th of \(m\) studies being combined; \(\sigma_i^2\) is the estimate of the within-study variance; \(\theta_i\) is the true underlying effect size for the \(i\)th study, \(\bar{A}\) is the overall pooled effect size estimate; and \(\bar{\sigma}^2\) is the estimate of between-study
heterogeneity. As in equations 2 and 5, the last line of equation 6 specifies reasonably vague prior distributions for all unknown parameters of the model.

5.2 Estimation of transition probabilities

For the continuous variable, that is the median time to an event, the transition probability is estimated via rates as outlined by Miller and Homan (1994)\textsuperscript{30}. For example, the pooled median progression-free survival time for the doxorubicin group is 22 weeks (Table 3), therefore the transition probability can be calculated as:

\[
P = 1 - e^{-R} = 0.090
\]

where \( R = \frac{-\ln(0.5)}{(22/3)} = 0.091 \) \[8\]

For binary data, that is the probability that an individual will transit from one state to another within a specified time period, the transition probability is calculated as follows. For example, the pooled response rate for doxorubicin is 34\% (Table 3) over the 7 treatment cycles; therefore the transition probability of moving into the ‘respond’ health state in any cycle of the model is given by:

\[
P = 1 - [1 - 0.34]^{1/7} = 0.06
\]

5.3 Sensitivity analyses (for data and model specification)

Assessment of the impact of model assumptions: As demonstrated in example 1, sensitivity analysis can be addressed within the same model as the primary analysis. It is possible that if docetaxel is accepted as the preferred treatment then the cost of the treatment may be reduced or a generic version(s) of treatment may be introduced at a cheaper cost. To analyse the impact of such a cost reduction of docetaxel in the future, 50\% lower treatment costs for the docetaxel group are compared to no change in the cost of standard care.
Prior distribution: As shown in illustrative example one, different non-informative prior distributions can be applied to check the sensitivity of the results to the choice of prior distribution. However, for simplicity, this has not been applied here.

5.3 Evaluation of the model

Following preliminary test runs, it was decided to use an initial run of 5,000 iterations as a ‘burn in’\(^1\), in order to reach convergence (these values were discarded), with inferences based on a further 20,000 sample iterations. As for example 1, two WinBUGS runs with very dispersed starting values were applied to check model convergence. The 2 chains gave very similar results thus implying convergence of the MCMC sampler has been achieved.

5.5 Model results

Primary analysis: The pooled estimates from the meta-analyses are given in Table 3 and Table 4 displays all the transition probabilities applied to the model. The overall results from the 20,000 iterations for the two approaches are displayed in Figure 5 in terms of the mean incremental costs plotted against the mean incremental utilities. The main observation from the plot is that docetaxel costs significantly more than doxorubicin (i.e. all points are above the x-axis) but there is little evidence that docetaxel is more effective than doxorubicin (i.e. the points span fairly evenly either side of the y-axis). Table 5 summarises this information by showing the overall mean incremental costs and incremental utilities for the different methods calculated across all 20,000 iterations.

This information can also be displayed as an acceptability curve\(^5\), which shows the probability that docetaxel is cost-effective compared to doxorubicin for given values of the ceiling ratio (i.e. the amount the decision maker is willing to pay for an additional QALY).
The actual value decision makers are willing to pay is usually not known, therefore points are plotted for a range of different values. This probability statement is an additional advantage of the Bayesian approach which allows us to say that at, say, £50,000 per additional QALY gained 29 out of the 20,000 iterations would show docetaxel to be cost-effective compared to doxorubicin (Figure 6). As can be seen from the Figure 6, the probability that docetaxel is cost-effective never rises above approximately 0.4.

Sensitivity analysis: The impact of reducing the cost of docetaxel by 50% reduced the mean cost for the docetaxel group to approximately £5,385 (95% CI £5,163 to £5,773) leading to a reduction in the incremental cost between the two groups of £1,714 (95% CI £1,382 to £2,123). However, the probability that docetaxel is cost-effective compared to doxorubicin remains below approximately 0.45 (Figure 6).

5.6 MODEL EXTENSION – Use of Informative prior distributions elicited from expert beliefs.

The primary analysis used vague prior distributions due to the small amount of data-based information currently available regarding the effectiveness of docetaxel on patients treated for advanced breast cancer. However, a previous study had elicited the beliefs of 20 oncologists worldwide regarding the ‘response rate’ for both standard and docetaxel treatment using a ‘Trial roulette’ elicitation approach which has been used in other cancer settings. Each oncologist was individually and independently asked to provide a histogram that represented his or her beliefs regarding the ‘response rate’ for standard and docetaxel treatment separately. To utilise this information in our model, their individual beliefs were then combined using a linear opinion pooling method to produce an overall histogram for the ‘response rates’ (Figure 7a(i) and b(i)).
Before the ‘response rate’ prior distributions for doxorubicin and docetaxel treatment are integrated into the model, a logit transformation is required since it is logit(response rate) that is being modelled. The logit (response rate) is then assumed to follow a normal distribution with the best fitting distributions to the transformed histograms having means –0.95 and –0.56, and standard deviations 0.92 and 0.87 for doxorubicin and docetaxel treatment respectively (Figure 7a(ii) and b(ii)). When converted back onto the proportion scale this translates to response rates of approximately 31% (s.d. 14.8) and 38% (s.d. 17.3) for doxorubicin and docetaxel treatment respectively. Hence, on average the oncologists believed the response rate would be approximately 7% greater for docetaxel compared to doxorubicin, however the large standard deviations for both response rates indicate considerable uncertainty in these prior beliefs. For further discussion on the pooling of distributions obtained from multiple experts see Genest and Zidek57.

Therefore, the prior distributions for the response rate for each treatment become:

(i) Docetaxel \[ \mu \sim N(-0.56, 0.87), \] and

(ii) Doxorubicin \[ \mu \sim N(-0.95, 0.92), \]

The results of this model are presented in Table 5. The main observation is that the point estimates are very similar to those obtained from the primary analysis but the 95% credible intervals are narrower owing to the additional information incorporated into the analysis on response rate.
6. DISCUSSION, CONCLUSIONS AND FURTHER WORK

This paper has illustrated how decision analytical models may be implemented within a Bayesian framework using Gibbs sampling MCMC methods in the software package WinBUGS. This environment permits great flexibility in model specification and provides the ability to integrate the synthesis and decision process in a single coherent model. As has been demonstrated, other advantages of using a Bayesian approach compared to the conventional approaches to decision analysis include: (i) the incorporation of greater parameter uncertainty by allowing for the fact that both the overall population effect and between-study precision in the meta-analysis have both been estimated by the data; (ii) incorporation of expert opinion either directly, as demonstrated here, or regarding the relative credibility of different data sources; (iii) the use of actual posterior distributions from meta-analyses avoiding the need to make distributional assumptions (such as normality), necessary for the classical analysis. However, distributional assumptions, such as normality of the log-odds from individual studies, have been made here for both types of analyses, although recent developments in Bayesian methods allow these to be relaxed, at the expense of complicating the model; (iv) full allowance is made for potential inter-relationships between all parameters in both the Markov model and the meta-analysis models used to propagate it; (v) the ability to make direct probability statements and thus direct answers to the question of interest (e.g. Bayesian meta-analysis can give a probability that the effect is above (or below) a particular value); and (vi) a framework to assess the importance of baseline risk on the treatment effect and to permit the application of sensitivity analyses.

Despite the advantages listed above, problems remain not addressed in the analyses presented herein. Often the format by which published data are presented is restrictive for the purposes of decision modelling. For example, in example 2, the median survival time and its standard
error were used as an approximation to the mean and its standard error due to limitations in
the published literature. Neither the Bayesian nor the classical approach accounts for the
added uncertainty inherent in making such approximations. Also, uncertainty in the
structural form of the model is not accounted for in the above analyses. For a Bayesian
method that deals with this issue see Draper 1995\textsuperscript{58}

A further limitation of decision analyses generally is that when estimating treatment
effectiveness, although data from RCTs are used, the data is used in such away that
randomisation is often “broken” to derive estimates for the model. This is true for example 2
where interest lies in comparing docetaxel with doxorubicin. Since the head-to-head trial
data is extremely limited (only one trial) it was necessary to obtain estimates of treatment
effectiveness from trials comparing these two compounds to other comparators. In doing so
data from only one arm of the comparative studies are used, hence the benefits of
randomisation are lost. Although not possible in this example, where treatments of interest
(say A and B) have been compared to a further common comparator (say C – e.g. an
established standard treatment or placebo) Bayesian models have been developed specifically
allowing indirect comparisons to be estimated (i.e. derive A v B from A v C and B v C
comparisons) removing the necessity to synthesise individual arms thus breaking
randomisation\textsuperscript{44}. Further, when direct randomised evidence of treatment effect is limited,
there may be advantages in extending the data synthesis to include data from non-randomised
sources. Bayesian methods have recently been developed which allow synthesis of evidence
from different study types while acknowledging the heterogeneity between sources of
evidence\textsuperscript{12,59}. 
A threat to the validity of any evidence synthesis is the potential of publication bias\(^{60}\). It is well established that statistically significant or positive findings are more likely to be published. A meta-analysis based on a biased sample of the total literature will produce biased answers. It has recently been acknowledged that economic evaluations based on such syntheses will themselves be biased\(^{61}\). Methods, both Bayesian and classical, exist which assess for the presence and even adjust for publication bias in a meta-analysis\(^{62}\). The simplest of these is visual assessment via a funnel plot of precision (1/SE) against treatment effect (Figure 8 – funnel plot for example 1). There is no evidence of publication bias if the funnel plot is symmetric around the pooled effect size. Some of the methods proposed are Bayesian and could in principle be included. A further consideration is that if different types of evidence are synthesised to inform a decision model, then these different sources could be affected by differential amounts of publication bias\(^{63}\).

A desirable goal for health evaluation generally is to individualise treatment regimes, identifying which patients are likely to benefit from treatment and those that are not. One possible step to achieving such a goal is to evaluate a decision model for subgroups of patients separately\(^{64}\). For example, considering studies of elective and non-elective caesarean sections separately in example one. However, well-documented problems exist when examining patient subgroups, not least due to issues of multiple testing and hence raised levels of false positive results. Bayesian methods to specifically address this issue are being developed\(^{65}\) which could be employed in decision model based evaluations.

The implementation of decision theoretic methodology such as the use of the expected value of information\(^{66,67}\) could be incorporated into the framework suggested above to identify the cost-utility of conducting future research, and in particular that which would reduce uncertainty.
regarding specific parameters/inputs in the Markov model. This is the subject of ongoing research.

In both examples, the posterior distribution for each parameter obtained from the meta-analyses is used to inform the probabilistic decision model. However, in some cases the predictive distribution may be more appropriate. For example, when inferences are made at the individual unit rather than population level. Using the predictive distribution can considerably increase the uncertainty associated with the outcome of interest. In example one, we are interested in the implication of introducing prophylactic antibiotics for caesarean section patients for an individual hospital. If the predictive distribution for RR had been used, then the cost difference between the treatment and placebo group would no longer be statistically significant (i.e. 95% CI £50.65 to £18.47 compared to £16.79 to £44.33).

The creation of structures in the U.K. (i.e. NICE) and elsewhere to facilitate evidence-based health policy decision making has highlighted the role that systematic reviews and decision models have to play. However, in highlighting their role, numerous methodological challenges have been identified, not least of which is their integration. Whilst we feel the adoption of a Bayesian approach provides a flexible and coherent framework within which to explore and address many of these challenges, it is by no means a panacea and requires careful and critical application, whilst development of these methods will be facilitated by multidisciplinary collaboration between economists, operational researchers and statisticians.
ACKNOWLEDGEMENTS

The authors would like to thank Michael Borowitz, Ruth Brown and John Hutton of MedTap International for their assistance with the elicitation process in example 2, and John Hutton for advise regarding earlier economic models. The authors would also like to thank Albert Bengiat, University Hospitals of Leicester (UHL) NHS Trust, and Kate Smith, Clatterbridge Centre for Oncology (CCO) NHS Trust, for clinical advise regarding development of the taxane model described in example 2 and finally David R. Jones, Paul Lambert and David Turner, Department of Epidemiology and Public Health, University of Leicester, for interesting and valuable discussions.

NJC is funded by University Hospitals of Leicester (UHL) NHS Trust. KRA has acted as a paid consultant to MEDTAP International on a project commissioned by Aventis (formerly Rhône-Poulenc Rorer) who manufacture Taxotere®
REFERENCES


Public Health, University of Leicester 2002.


FIGURES

Figure 1: Illustrative example 1: Prophylactic antibiotic use in caesarean section
Figure 2: Underlying risk – Plot of treatment effect against control group risk.
Figure 3: Sensitivity analysis – using alternative ‘vague’ prior distributions for $\tau^2$ (the between-study variance parameter)

- Gamma(0.001,0.001) distribution on $\tau^2$
- Normal(0,1.0^{-6}) distribution truncated at zero on $\tau$
- Uniform(0,50) distribution on $\tau$

Cost difference (treatment minus standard care)
Figure 4: Illustrative example 2 – Taxane use in second-line treatment of breast cancer
Figure 5: Cost-effectiveness plane – Results of the 20,000 iterations with ‘vague’ prior distributions

Bayesian (MCMC) Simulations

Incremental cost

Incremental utility

Doxorubicin dominates

Docetaxel more effective but more costly

Docetaxel less costly but less effective

Docetaxel dominates
Figure 6: Cost-effectiveness acceptability curve for example 2 (i.e. probability docetaxel cost-effective compared to doxorubicin).
Figure 7: Prior beliefs for response rate

a) Docetaxel

i) Aggregated histogram derived from prior beliefs of 20 oncologists

ii) Histogram transformed onto the logit scale with best fitting normal distribution used as prior distribution for the model

Response rate (docetaxel)

logit (Response rate for docetaxel)
b) Doxorubicin

i) Aggregated histogram derived from prior beliefs of 20 oncologists

ii) Histogram transformed onto the logit scale with best fitting normal distribution used as prior distribution for the model superimposed

Response rate for doxorubicin

logit (Response rate for doxorubicin)
Figure 8: Funnel plot to assess publication bias for studies reporting infection rates between antibiotics & standard care in example 1.
## TABLES

### Table 1: Prophylactic antibiotics in caesarean section

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment</th>
<th>Placebo</th>
<th>Grade</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adeleye et al 1981</td>
<td>11/58</td>
<td>14/48</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Bibi et al 1994</td>
<td>4/133</td>
<td>28/136</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Chan et al 1989</td>
<td>27/299</td>
<td>12/101</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Conover et al 1984</td>
<td>2/68</td>
<td>1/56</td>
<td>A</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Cormier et al 1989</td>
<td>5/55</td>
<td>8/55</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Dashow et al 1986</td>
<td>3/100</td>
<td>0/33</td>
<td>A</td>
<td>Elective</td>
</tr>
<tr>
<td>Dashow et al 1986</td>
<td>4/183</td>
<td>3/44</td>
<td>A</td>
<td>Non-elective</td>
</tr>
<tr>
<td>De Boer et al 1989</td>
<td>1/11</td>
<td>5/17</td>
<td>B</td>
<td>Elective</td>
</tr>
<tr>
<td>De Boer et al 1989</td>
<td>10/80</td>
<td>21/74</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Dillon et al 1981</td>
<td>0/46</td>
<td>4/55</td>
<td>A</td>
<td>Both</td>
</tr>
<tr>
<td>Duff et al 1980</td>
<td>0/26</td>
<td>1/31</td>
<td>A</td>
<td>Both</td>
</tr>
<tr>
<td>Duff et al 1982</td>
<td>0/42</td>
<td>0/40</td>
<td>B</td>
<td>Elect</td>
</tr>
<tr>
<td>Elliot et al 1986</td>
<td>0/119</td>
<td>1/39</td>
<td>B</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Engel et al 1986</td>
<td>1/50</td>
<td>9/50</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Fugere et al 1983</td>
<td>2/60</td>
<td>6/30</td>
<td>B</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Gall 1979</td>
<td>1/46</td>
<td>1/49</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Gerstner et al 1980</td>
<td>3/53</td>
<td>9/50</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Gibbs et al 1972</td>
<td>0/33</td>
<td>4/28</td>
<td>C</td>
<td>Both</td>
</tr>
<tr>
<td>Gibbs et al 1973</td>
<td>0/34</td>
<td>6/34</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Gibbs et al 1981</td>
<td>0/50</td>
<td>2/50</td>
<td>B</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Gordon et al 1979</td>
<td>0/78</td>
<td>1/36</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Hager et al 1983</td>
<td>1/43</td>
<td>1/47</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Hagglund et al 1989</td>
<td>0/80</td>
<td>3/80</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Harger et al 1981</td>
<td>2/196</td>
<td>14/190</td>
<td>B</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Hawrylyshyn et al 1983</td>
<td>2/124</td>
<td>2/58</td>
<td>A</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Ismail et al 1990</td>
<td>2/74</td>
<td>8/78</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Karhunen et al 1985</td>
<td>2/75</td>
<td>9/77</td>
<td>A</td>
<td>Both</td>
</tr>
<tr>
<td>Kreutner et al 1978</td>
<td>0/48</td>
<td>2/49</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Kristensen et al 1990</td>
<td>0/102</td>
<td>1/99</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Lapas et al 1989</td>
<td>1/50</td>
<td>10/50</td>
<td>B</td>
<td>Both</td>
</tr>
<tr>
<td>Leonetti et al 1989</td>
<td>0/100</td>
<td>1/50</td>
<td>B</td>
<td>Non-elective</td>
</tr>
<tr>
<td>Levin et al 1983</td>
<td>0/85</td>
<td>3/43</td>
<td>A</td>
<td>Both</td>
</tr>
<tr>
<td>Lewis et al 1990</td>
<td>1/36</td>
<td>1/25</td>
<td>B</td>
<td>Elective</td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Outcome</td>
<td>Type</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Lewis et al 1990</td>
<td>2/76, 4/75</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Mahomed et al 1988</td>
<td>12/115, 15/117</td>
<td>A</td>
<td>Elective</td>
<td></td>
</tr>
<tr>
<td>Mallaret et al 1990</td>
<td>6/136, 16/130</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>McCowan et al 1980</td>
<td>9/35, 7/38</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Miller et al 1968</td>
<td>13/150, 23/150</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Moodley et al 1981</td>
<td>2/40, 4/20</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Moro et al 1974</td>
<td>0/74, 2/74</td>
<td>A</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Padilla et al 1983</td>
<td>0/34, 5/37</td>
<td>A</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Phelan et al 1979</td>
<td>2/61, 2/61</td>
<td>A</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Polk et al 1982</td>
<td>3/146, 9/132</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Rehu et al 1980</td>
<td>4/88, 4/40</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Roex et al 1986</td>
<td>1/64, 7/65</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Ross et al 1984</td>
<td>7/57, 7/58</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Rothbard et al 1975</td>
<td>0/16, 1/16</td>
<td>C</td>
<td>Elective</td>
<td></td>
</tr>
<tr>
<td>Rothbard et al 1975</td>
<td>2/31, 6/37</td>
<td>C</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Ruiz-Moreno et al 1991</td>
<td>1/50, 4/50</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Saltzman et al 1985</td>
<td>1/50, 2/49</td>
<td>A</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Schedvins et al 1986</td>
<td>2/26, 0/27</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Stage et al 1983</td>
<td>3/133, 12/66</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Stiver et al 1983</td>
<td>6/244, 17/117</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Tully et al 1983</td>
<td>1/52, 2/61</td>
<td>A</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Tzingounis et al 1982</td>
<td>2/46, 4/50</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Weissberg et al 1971</td>
<td>0/40, 3/40</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Wong et al 1978</td>
<td>2/48, 3/45</td>
<td>A</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Work et al 1977</td>
<td>3/40, 1/40</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td>Yip et al 1997</td>
<td>1/160, 1/160</td>
<td>B</td>
<td>Both</td>
<td></td>
</tr>
<tr>
<td>Young et al 1983</td>
<td>1/50, 4/50</td>
<td>B</td>
<td>Non-elective</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean (95% Credible Interval)</td>
<td>Posterior distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------</td>
<td>------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relative Risk, $RR_{antibiotics}$</strong></td>
<td>0.29 (0.20 to 0.39)</td>
<td>![Graph of RR_{antibiotics}] sample: 20000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prob(wound infection/placebo), $p_1$</strong></td>
<td>0.08 (0.06 to 0.10)</td>
<td>![Graph of p_1] sample: 20000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prob(wound infection/antibiotics), $p_2$</strong></td>
<td>0.02 (0.02 to 0.03)</td>
<td>![Graph of p_2] sample: 20000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cost difference (antibiotics minus standard)</strong></td>
<td>£29.37 (£16.79 to £44.33)</td>
<td>![Graph of cost difference] sample: 20000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between study variance, $\tau^2$</strong></td>
<td>0.36 (0.07 to 0.80)</td>
<td>![Graph of \tau^2] sample: 20000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Pooled results from Meta-analyses for example 2: Taxane use in second-line treatment of breast cancer.

<table>
<thead>
<tr>
<th></th>
<th>Docetaxel</th>
<th>Doxorubicin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time in weeks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% Credible Interval)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median progression-free time</td>
<td>24 (16 to 33)</td>
<td>22 (15 to 30)</td>
</tr>
<tr>
<td>Median overall survival time</td>
<td>53 (35 to 75)</td>
<td>61 (35 to 109)</td>
</tr>
<tr>
<td><strong>Probabilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% Credible Interval)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response rate</td>
<td>0.43 (0.29 to 0.59)</td>
<td>0.35 (0.16 to 0.61)</td>
</tr>
</tbody>
</table>
Table 4: Transition probabilities for example 2: Taxane use in second-line treatment of breast cancer.

<table>
<thead>
<tr>
<th>State Transition</th>
<th>Docetaxel Mean (95% CrI)</th>
<th>Doxorubicin Mean (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Stable’ to ‘Response’</td>
<td>0.08 (0.05 to 0.12)</td>
<td>0.06 (0.02 to 0.17)</td>
</tr>
<tr>
<td>‘Stable’ to ‘Progressive’</td>
<td>0.09 (0.04 to 0.20)</td>
<td>0.10 (0.02 to 0.17)</td>
</tr>
<tr>
<td>‘Progressive’ to ‘Death’</td>
<td>0.09 (0.04 to 0.20)</td>
<td>0.07 (0.02 to 0.17)</td>
</tr>
</tbody>
</table>
Table 5: Mean Incremental Costs and Utilities for example 2: Taxane use in second-line treatment of breast cancer.

<table>
<thead>
<tr>
<th></th>
<th>Mean (95% Credible Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incremental cost, $\Delta C$</td>
</tr>
<tr>
<td></td>
<td>(Intercepts vertical axis at $-\Delta C$)</td>
</tr>
<tr>
<td>Bayesian (MCMC) model</td>
<td>£6,923 (£6,553 to £7,512)</td>
</tr>
<tr>
<td>Bayesian (MCMC) with IP model</td>
<td>£6,926 (£6,568 to £7,501)</td>
</tr>
</tbody>
</table>

$\frac{\Delta C}{\Delta E} =$ Incremental cost effectiveness ratio (ICER) = Intercept of horizontal axis.

IP = Informative prior distributions for ‘Response rate’