Bayesian cost-effectiveness analysis using individual-level data is sensitive to the choice of Uniform priors on the standard deviations for costs in Log-Normal models

Xiaoxiao Ling[1,2], Andrea Gabrio[3], Gianluca Baio[1]

1 Department of Statistical Science, University College London, UK

2 Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

3 Department of methodology and statistics, Faculty of Health Medicine and Life Science, Maastricht University, Maastricht, Limburg, NL

Correspondence: Xiaoxiao Ling, x.ling.17@ucl.ac.uk

Abstract

Background: Bayesian cost-effectiveness analysis (CEA) requires the specification of prior distributions for all parameters to be empirically estimated via Bayes' rule. When costs are modelled via Log-Normal distributions, Uniform prior distributions are commonly applied on the logarithm-scale standard deviations for costs due to its ease of implementation. However, the consequences of placing wide Uniform priors on standard deviations of log costs for the interpretation of original-scale CEA results remain unclear. The purpose of our study is to explore the impact of using Uniform priors for the standard deviations of cost data on CEA conclusions when costs are assumed to be log-normally distributed.

Methods: The analysis has been performed using individual-level cost-utility data from a randomised controlled trial. Costs are initially jointly modelled with QALYs using Log-Normal and Beta distributions respectively. Uniform prior distributions with different upper bounds are applied to log-scale standard deviations in the cost Log-Normal model. We compare the performance of Uniform priors under the Log-Normal distribution with other distributional assumptions for costs. A simulation study has then been conducted to explore the impact of these models and prior choices on cost estimates in CEAs.

Results: Results show that the choice of Uniform priors on standard deviations of log costs in Log-Normal model can substantially induce large fluctuations in cost estimates, and thus potentially affect the final estimates of the intervention being cost-effective compared to other distributional assumptions. This is potentially driven by the occurrence of zero values in cost data.

Conclusion: Bayesian CEAs may be sensitive to the choice of upper bounds of the Uniform priors for the standard deviations of log costs in Log-Normal models, particularly when data contain zero values. Our results suggest that caution should be made when Uniform distributions with large upper bounds are used.

Keywords

Bayesian method, Cost-effectiveness analysis, Uniform prior, Log-Normal distribution

Key Points

- Bayesian cost-effectiveness analyses are sensitive to the choice of upper bounds of the Uniform prior distributions on log cost standard deviations when costs data are assumed to be log-normally distributed.

- Caution should be given to the potential consequences of applying Uniform distributions on cost standard deviations as minimally informative priors based on both distributional model assumptions and actual characteristics of the data.

- It is more conservative to choose models that are more robust to the upper bounds of Uniform prior distributions and is also necessary to consider alternative priors to assess the robustness of posterior results.

1 Introduction

Bayesian analysis has become increasingly popular in cost-effectiveness analysis (CEA), over the years [1–4]. One of the reasons behind its popularity stems from its ability to incorporate external evidence or elicited expert opinions into the analysis in a principled way, through prior distributions. However, in cases where genuine information is lacking, "minimally informative" prior distributions are often used, at least as a starting point [5,6].

While common in Bayesian CEAs [7], the implementation of models based on scarce contextual information in real applications can be challenging. Given that CEA data are usually skewed, simple assumptions such as a Normal distribution to describe sampling variability are less likely to hold. Alternative choices (e.g. modelling sampling variability using a Gamma distribution to cater for skewness and positivity of the individual costs) are often specified using classical parameters whose ``physical'' interpretations, while meaningful, may be less immediately intuitive to health economists. For instance, unlike the Normal distribution which is naturally parameterised by a population mean and standard deviation, a Gamma distribution is usually indexed by a *rate* and a *shape* parameter (rather than its mean and standard deviation), which are difficult to understand in a health economics context and complicate the natural interpretation and elicitation of prior beliefs [8].

The Uniform distribution on a suitable large and positive range is often considered as the default choice for a minimally informative prior concerning the standard deviation of costs [9–11]. However, slight differences exist in the specification of the Uniform prior under different distributional assumptions in the analysis of cost data. Uniform priors tend to be placed on the original scale of standard deviations or their functions when costs are assumed to follow a Normal or Gamma distribution [8–10,12], while it is straightforward to directly assign priors on standard deviations of log costs when costs are modelled through a Log-Normal distribution [13].

In the general Bayesian statistical literature, it has been already recognised that placing vague Uniform priors on log-scale standard deviations in Log-Normal models is often a sensitive choice, which exerts a potentially notable influence on posterior results [14,15]. Alternative options of minimally informative priors for variance or standard deviation have been suggested in the wider statistical literature. For example, Gelman recommended the half-t family of prior distributions on the original-scale standard deviation [16], but the article primarily focused on hierarchical Normal models, which may not always align with CEA data. Under the choice of a Log-Normal distribution, Zellner proposed the use of diffuse priors [17], although the method encounters the difficulty of summarising the posterior distribution of the mean on its original scale. Fabrizi and Trivisano proposed a generalised inverse Gaussian prior on the log-scale variance of the Log-Normal distribution, allowing for finite posterior mean estimates [18]. Yet, the application of these priors from statistical literature to CEAs poses challenges in interpretation since they are less intuitive compared with a Uniform prior distribution.

The objective of this study is to investigate how setting Uniform priors for the standard deviations of a cost Log-Normal distribution may affect the overall CEAs. Our focus is different from identifying the "optimal" prior from a statistical viewpoint, but is more specialised in the context of health economic evaluations. So far, there has been little discussion regarding the choice of prior distributions in CEAs. It remains poorly understood to what extent the application of Uniform distributions as minimally informative priors for standard deviations of log costs in Log-Normal models can function to easily convey specific beliefs about cost standard deviations on their natural scale.

The study will explore the impact of using Uniform prior distributions for the log-scale standard deviations of cost data on cost-effectiveness results when costs are assumed to be log-normally distributed using a real case study. The Log-Normal model will be compared to models with other distributional assumptions in terms of their CEA results. Based on these, recommendations will be formulated regarding the use of Uniform prior distributions for cost standard deviations in Bayesian CEAs.

2 Materials and Methods

2.1 Case Study

The ORBIT trial is a randomised controlled trial that compares the cost-effectiveness of the intervention, online-delivered therapist-supported exposure and response prevention, with the control, online education, among children having Tourette syndrome or chronic tics disorder from a health care perspective over 6 months [19]. A total of 224 patients had been randomly assigned to the two treatment arms: 112 were randomised to the intervention and 112 to the control. Cost and utility data were collected at baseline, 3 months and 6 months. The total costs and quality-adjusted life years (QALYs) [20] were estimated at 6 months using the costs and utility data at each time point.

Complete cases, defined as patients with fully observed health care costs and utility scores at the three data collection time points, are used to streamline the analysis (71 in the control and 62 in the intervention). The decision is made because we will concentrate on providing a comprehensive insight into how the choice of Uniform priors with various upper bounds can influence cost-effectiveness conclusions and consider that a complete case analysis is adequate to serve this purpose.

2.2 Statistical models

Analyses are performed within a Bayesian framework using a joint modelling approach [21]. The joint distribution of costs (c) and QALYs (e), denoted as $p(c, e)$, is parameterised by the product of a conditional distribution of costs given QALYs, $p(c \mid e)$, and a marginal distribution of QALYs, $p(e)$, to capture the correlation between the outcomes. Apart from the Log-Normal distribution, Normal and Gamma distributions are considered for cost data to illustrate the extent to which different distributional assumptions can influence the impact of Uniform priors for cost standard deviations on cost-effectiveness results. Normal and Beta distributions are applied to QALYs: the former is chosen following the original health economic analysis of the trial, while the latter is used to reflect the potential skewness of QALYs.

The following notation is used: let c_{it} and e_{it} denote total costs and QALYs for patient i in each treatment arm t respectively. The treatment indicator t is dropped to simplify model specification.

2.2.1 Normal model

The first model, assuming that both costs and QALYs are normally distributed, is undertaken to mimic the assumptions of routine CEAs. It can be thought of as a Bayesian equivalent to *seemingly unrelated regression* (SUR) [22] but re-expressed by using the aforementioned product formulation of the conditional distribution of costs given QALYs and the distribution of QALYs.

We model costs conditional on QALYs as:

$$
\begin{aligned}
c_i \mid e_i &\sim \text{Normal}(\mu_{ci}, \sigma_c^2) \\
\mu_{ci} &= \alpha_0 + \alpha_1(c_{0i} - \bar{c}_0) + \alpha_2(e_i - \bar{e}) + \alpha_3 s_i
\end{aligned}
$$

where μ_{ci} and σ_c are the (individual-level) mean and the (population-level) standard deviation of the costs. Covariates are selected based on the original health economics analysis: $(c_{0i} - \bar{c}_0)$ is the centred baseline costs, $(e_i - \bar{e})$ is the centered QALYs while s_i is the site at which individual i has been treated. α_0 is the intercept, α_1 represents the impact of baseline costs on total costs, α_2 quantifies the association between costs and QALYs while α_3 represents the effects of site.

QALYs are modelled as:

$$e_i \sim \text{Normal}(\mu_{ei}, \sigma_e^2)$$
$$\mu_{ei} = \beta_0 + \beta_1(u_{0i} - \bar{u}_0) + \beta_2 s_i$$

where the parameters μ_{ei} and σ_e are the (individual-level) mean and the (population-level) standard deviation for QALY. In line with the original economic evaluation, centred version of utility values at baseline $(u_{0i} - \bar{u}_0)$ and site s_i are included as covariates in the QALY model.

Suitable priors are specified for parameters to complete the model (see Appendix A for details and Appendix D for prior sensitivity assessment): vague priors are assigned to the regression coefficients and Uniform priors are assumed for the standard deviations.

2.2.2 Gamma model for cost, Beta model for QALY

Since cost data are often right-skewed and non-negative, the cost component of the model above is expanded by assuming a Gamma distribution to reflect these features [23–25]. To facilitate prior specification and improve interpretability, the Gamma distribution for the costs given QALYs is re-parameterised by defining the shape parameter $\zeta_{ci} = \mu_{ci}\tau_{ci}$ and the rate parameter $\tau_{ci} = \mu_{ci}/\sigma_c^2$, where μ_{ci} and σ_c describe the (individual-level) mean and the (population-level) standard deviation of the costs, respectively. This parameterisation allows priors to be placed directly on more intuitive parameters such as cost standard deviation and avoids the difficulty of formalising straightforward prior knowledge for the shape and rate parameters. A log link function is used to connect the conditional mean costs given QALYs and relevant covariates. Since some participants accrued zero costs in the trial, we re-scale the observed values for the costs by adding a small constant $\epsilon = 1$ to the original data [8]. We note that modifying the data is necessary only to fit Gamma distributions to costs and has minimal impact on the conclusions under the overall scope of the study.

The costs conditional on the QALYs are modelled as:

$$c_i \mid e_i \sim \text{Gamma}(\mu_{ci}\tau_{ci}, \tau_{ci})$$
$$\log(\mu_{ci}) = \alpha_0 + \alpha_1(c_{0i} - \bar{c}_0) + \alpha_2(e_i - \bar{e}) + \alpha_3 s_i$$

The QALYs' empirical distributions also show some degree of skewness, and are theoretically restricted within the range $(0, 0.5)$ in the ORBIT trial. Following the recommendation of previous research [26], a Beta distribution is fitted to the QALYs to capture the skewness of the outcome. The distribution of the QALYs is parameterised by two shape parameters $\kappa_{ei} = \mu_{ei}\phi_{ei}$ and $\gamma_{ei} = (1 - \mu_{ei})\phi_{ei}$, where μ_{ei} represents the (individual-level) mean while ϕ_{ei} is an (individual-level) scale parameter, defined as:

$$\phi_{ei} = \frac{(1 - \mu_{ei})\mu_{ei}}{\sigma_e^2} - 1$$

Such parameterisation allows to construct the priors directly on the standard deviations of the QALYs and brings more intuitive understanding of the model. A logit link function is applied for the generalised linear model.

The model for the QALYs can be expressed as:

$$e_i \sim \text{Beta}(\mu_{ei}\phi_{ei}, (1 - \mu_{ei})\phi_{ei})$$
$$\text{logit}(\mu_{ei}) = \beta_0 + \beta_1(u_{0i} - \bar{u}_0) + \beta_2 s_i$$

The model is completed by placing priors on the regression coefficients and the standard deviations: $\text{Normal}(0, 100^2)$ for the coefficients in the cost model, $\text{Normal}(0, 2^2)$ for the coefficients in the QALY model given the use of logit link; wide Uniform priors have been given to cost standard deviations – they have been selected to

convey minimally information to the analysis and to ensure consistency with the Normal distribution choices for model comparisons throughout this study; although Uniform priors have been placed on the standard deviations of the QALYs, they are restricted using a Uniform$\left(0, \sqrt{\mu_e(1 - \mu_e)}\right)$ due to the property of Beta distribution [4].

2.2.3 Log-Normal model for cost, Beta model for QALY

The model above is modified by assuming that the cost data are log-normally distributed, while maintaining the Beta distribution for the QALYs.

The cost Log-Normal model can be written as:

$$
\begin{aligned}
c_i \mid e_i \quad &\sim \text{Log-Normal}(\nu_{ci}, \delta_c^2) \\
\nu_{ci} \quad &= \alpha_0 + \alpha_1(c_{0i} - \bar{c}_0) + \alpha_2(e_i - \bar{e}) + \alpha_3 s_i
\end{aligned}
$$

It should be noted the ν_{ci} and δ_c in the Log-Normal model are now (individual-level) mean and (population-level) standard deviation of costs on the log scale respectively. To obtain the mean (μ_{ci}) and standard deviation (σ_{ci}) of costs on their original scale for each individual i, the following expressions could be used:

$$
\begin{aligned}
\mu_{ci} \quad &= \exp\left(\nu_{ci} + \frac{\delta_c^2}{2}\right) \\
\sigma_{ci} \quad &= \sqrt{[\exp(\delta_c^2) - 1]\exp(2\nu_{ci} + \delta_c^2)}
\end{aligned}
$$

Similar to the Gamma distribution, the Log-Normal distribution requires the data to be positive, necessitating the addition of a small value ($\epsilon = 1$) to the original cost data. Vague Normal prior distributions are assigned to regression coefficient parameters.

Uniform priors with different upper bounds are considered for the standard deviations of log costs in Log-Normal distribution (see Appendix A) to assess their impact on the cost-effectiveness results. Specifically, to ensure the comparability across models with different distributional assumptions, the Uniform prior distributions on the log-cost standard deviation in the Log-Normal models are intended to lead to reasonable and comparable ranges to those on the original cost scale in other models. The exact upper bound values to be considered in the Log-Normal model may depend on the mean of the data, given the mathematical properties of the Log-Normal distribution. When the actual mean costs are large, a Uniform prior distribution with a large upper bound on log-cost standard deviation may produce unrealistic results. An illustration of the potential impact of the Uniform priors for log-scale standard deviation on original-scale standard deviation is provided in Appendix B. Therefore, Uniform prior distributions with narrower and reasonable ranges have been chosen for the Log-Normal models, compared to those in the Normal and Gamma models.

2.3 Implementation

All models are fitted in `JAGS` [27], a program for Bayesian inference based on Markov Chain Monte Carlo (MCMC) via the `R2jags` package in `R` version 4.0.3 [28]. We run two chains, each with 52,000 iterations, and discarded the first 2,000 iterations from each chain as a burn-in phase. A thinning rate of 10 has also been applied to reduce autocorrelation, leading to a total sample of 10,000 iterations for inference. Convergence is assessed through potential scale reduction statistics, $\hat{R}$ [29], and a visual inspection on trace plots. To overcome the difficulty in model convergence arising from the differing scales of costs and QALYs, the original cost data have been scaled down by a factor of 10, but the resulting inferences are robust to the choice of the scaling factor. Model fit is

assessed through posterior predictive checks (see Appendix C) and compared by deviance information criterion (DIC) [30].

Given the main focus of our study, cost estimates are initially reported for the CEA models to identify how much the marginal mean costs can be affected by different upper bounds of the Uniform distributions. Mean costs with different distributional assumptions are also compared to investigate whether the influence brought by the Uniform prior distributions on cost standard deviations is sensitive to the choice of cost distributions. After that, the impact of using Uniform prior distributions on cost-effectiveness results is explored using cost-effectiveness plane [31] and cost-effectiveness acceptability curve (CEAC) [32].

3 Results

3.1 ORBIT trial data

Both costs and QALYs show some degree of skewness (Fig. 1). Most of the cost data are below £2,000 with 3 participants as an exception in each treatment arm. A comparable proportion of participants with zero health care costs in both treatment arms are observed: 5/62 (8.06%) participants accrue zero costs in the intervention while 6/71 (8.45%) participants in the control. QALYs mostly lie between 0.3 and 0.5 with no participant having a perfect health status of 0.5 by the end of the 6-month follow-up.
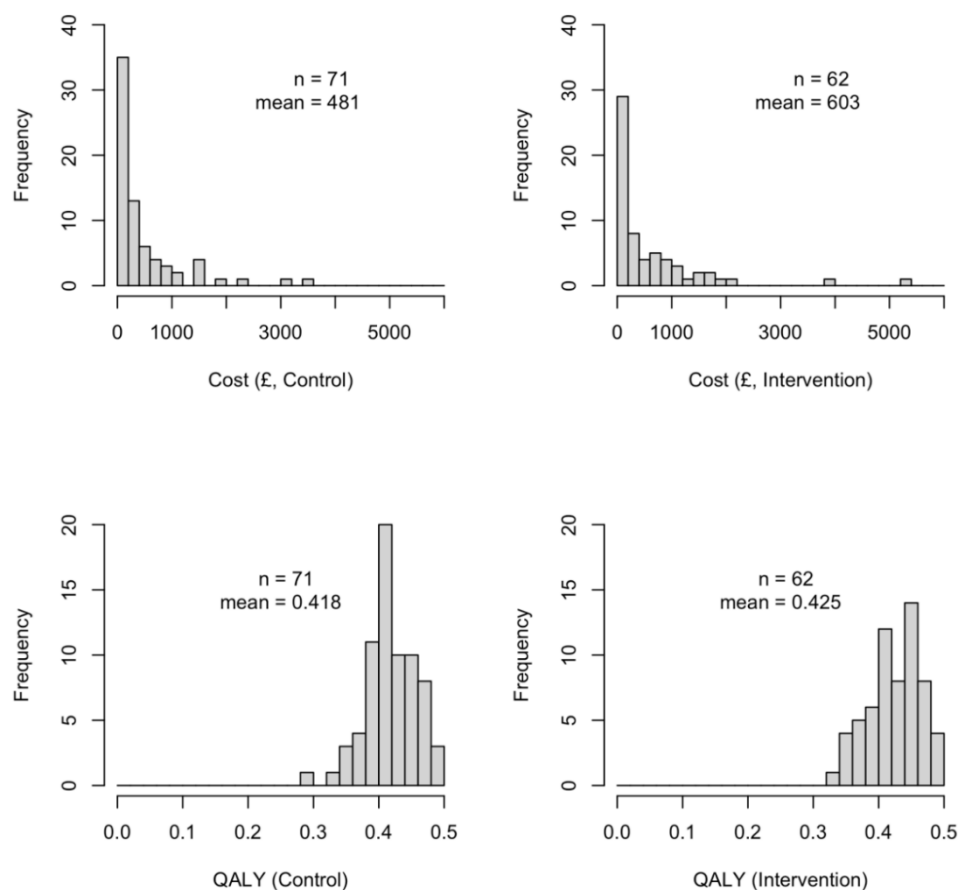

Fig. 1 Distribution of trial cost data by treatment arms. QALY=Quality-Adjusted Life-Years.

3.2 Cost analysis

The specification of a Uniform distribution as the prior on standard deviation for log costs in the Log-Normal model influences results when zero values exist in the cost data (Table 1). Compared to the Normal and Gamma models, the model fit and cost estimates from the Log-Normal models are more sensitive to the choice of upper bounds for the Uniform priors. The Log-Normal models with the two lowest upper bound values for the Uniform prior present higher DIC values than the others, suggesting a poorer fit to the data, which has also been confirmed via posterior predictive check (see Appendix C). Improvements in model fit can be observed when the upper bound of the Uniform prior distribution increases. Although the DICs of the Log-Normal models with Uniform(0,2) and Uniform(0,3) are similar, the two models produce different scales of cost estimates: the incremental mean costs of the intervention compared to the control are £344 (95% credible intervals: -£2,258, £3,256) with Uniform(0,2) but they jump to £1,301 (95% CIs: -£4,169, £8,699) when Uniform(0,3) is applied. The Normal models do not fit the data well, as their DIC values are higher than those of the Gamma models and most Log-Normal models. In contrast, the Gamma models show the best fit to the data, as confirmed by both DIC and posterior predictive checks.

Table 1 Marginal mean and incremental mean cost estimates, with 95% credible intervals in parentheses, for models with different Uniform prior distributions on cost standard deviations. Note: Costs are measured using British pound (£). DIC = Deviance Information Criteria; Con = Control; Int = Intervention.

Prior	DIC	Costs (Con)		Costs (Int)		Incremental Costs	
Normal model							
Uniform(0,1000)	948.8	476	(320, 632)	598	(421, 775)	123	(-102, 366)
Uniform(0,10000)	948.8	476	(320, 632)	598	(421, 775)	123	(-102, 366)
Gamma model							
Uniform(0,1000)	718.7	537	(388, 692)	709	(451, 970)	171	(-128, 484)
Uniform(0,10000)	718.4	535	(393, 697)	710	(469, 982)	174	(-131, 480)
Log-Normal model							
Uniform(0,3)	744.4	2,095	(497, 4,761)	3,396	(620, 8,279)	1,301	(-4,169, 8,699)
Uniform(0,2)	740.1	1,610	(489, 3,266)	1,954	(651, 3,868)	344	(-2,258, 3,256)
Uniform(0,1)	919.3	412	(260, 585)	468	(287, 667)	56	(-208, 335)
Uniform(0,0.8)	1126.9	337	(241, 453)	386	(265, 518)	49	(-127, 215)

Mean cost estimates by treatment arms, incremental mean costs and their 95% CIs are graphically compared across models with different distributional assumptions and prior distributions (Fig. 2). Unlike with the Normal and Gamma model, the Log-Normal model exhibits large variations in cost estimates based on the choice of the Uniform priors for standard deviations of log costs. The use of Uniform(0,0.8) and Uniform(0,1) — the two Uniform priors associated with a poorer fit to the data — results in comparable cost estimates to those produced from the Normal and Gamma models. In contrast, the other two Uniform priors, which have shown a better model fit, can inflate the mean cost estimates in the two treatment arms (and thus incremental mean costs), even when the upper bound of the

Uniform prior distribution has been limited to 2. The Normal and Gamma models are more robust to the choice of upper bounds of the Uniform priors compared to the Log-Normal model, producing similar results across the different upper bound values explored.
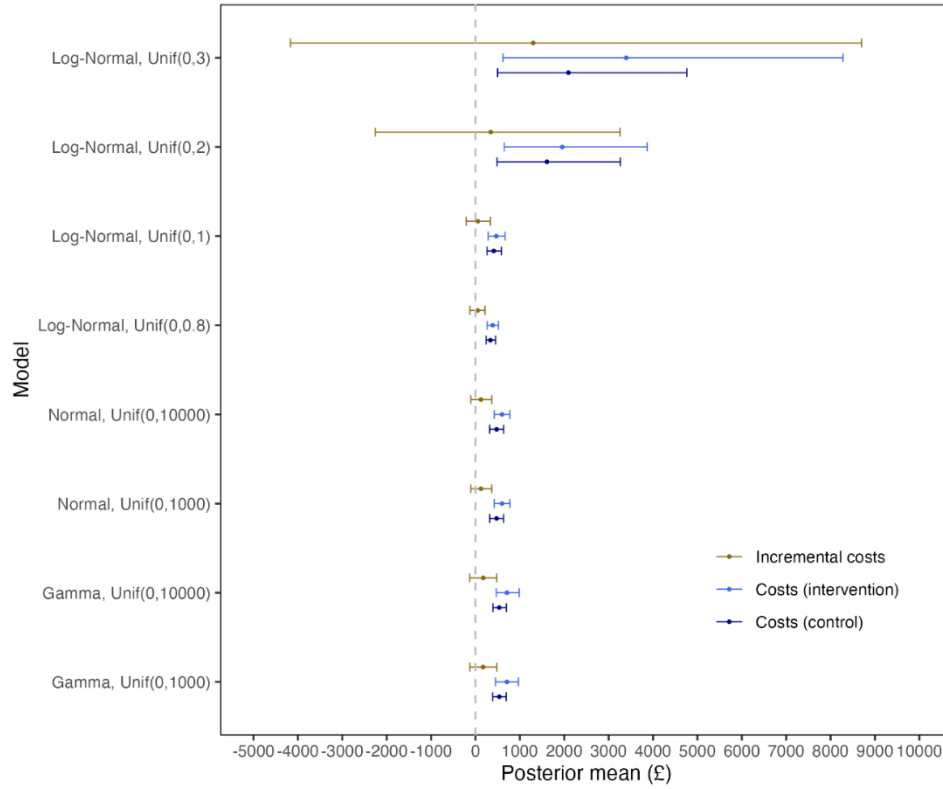

Fig. 2 Marginal mean cost by treatment arms, incremental mean costs and their 95% credible intervals of Normal, Gamma and Log-Normal models with different Uniform prior distributions on cost standard deviations.

3.3 Cost-effectiveness results

Fig. 3 shows the cost-effectiveness planes of models based on different distributional assumptions and varying Uniform priors for cost standard deviations. Here we specifically focus on the cost-effectiveness planes for the two Uniform distributions with a better fit to the data — Uniform(0,2) and Uniform(0,3) — as priors on the standard deviations of log costs in Log-Normal models. The decision is made because the Uniform(0,0.8) and Uniform(0,1) (though their cost-effectiveness planes show similar patterns to Normal and Gamma models) appear too restrictive for the data based on both their DIC values and posterior predictive checks.

The cost-effectiveness planes illustrate how sensitive the cost-effectiveness results are to the choice of models and upper bounds of the Uniform prior distributions for cost standard deviations. The pairs of incremental costs and QALYs, and the associated incremental cost-effectiveness ratio (ICER), predominantly fall below the willingness-to-pay threshold of £30,000 in the Normal and Gamma models, supporting the cost-effectiveness of the intervention. Conversely, the Log-Normal models present much higher incremental costs and ICERs compared to models with other distributional assumptions.

The sensitivity of the cost-effectiveness results to the upper bounds of the Uniform priors may be affected by model distributional assumptions. As shown in Fig. 3: panels (a) and (b) for the Normal model, and panels (c) and (d) for

the Gamma model, the two models are robust to the upper bounds of the Uniform prior distributions for their cost standard deviations. In contrast, the results of the Log-Normal models can be greatly influenced by the Uniform prior distributions used for the standard deviations of log costs, as illustrated in Fig. 3 panels (e) and (f). An increase in the upper bounds of the Uniform priors can bring more extreme values and substantial uncertainty to the incremental mean costs, thus becoming less likely to conclude that intervention is cost-effective.
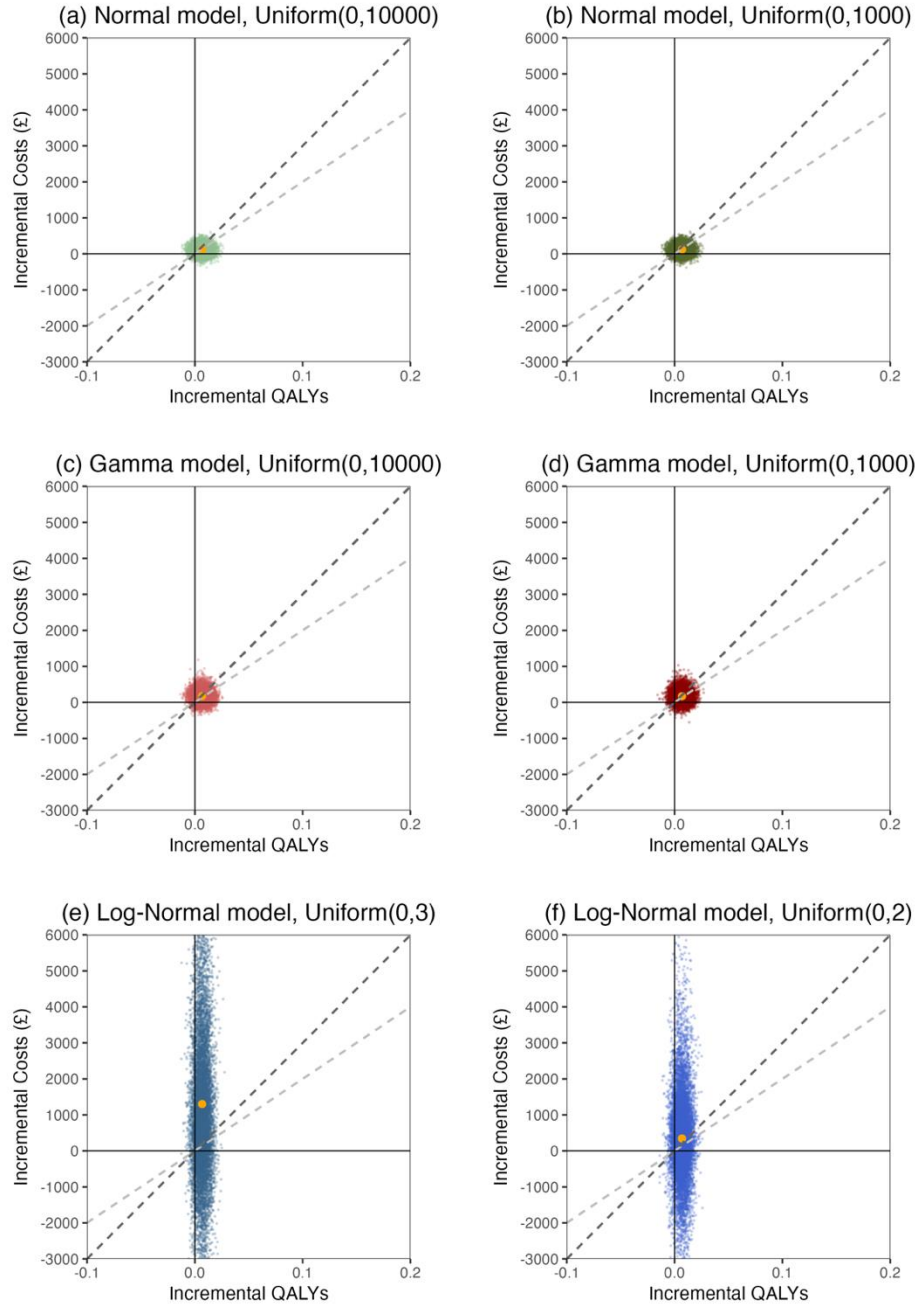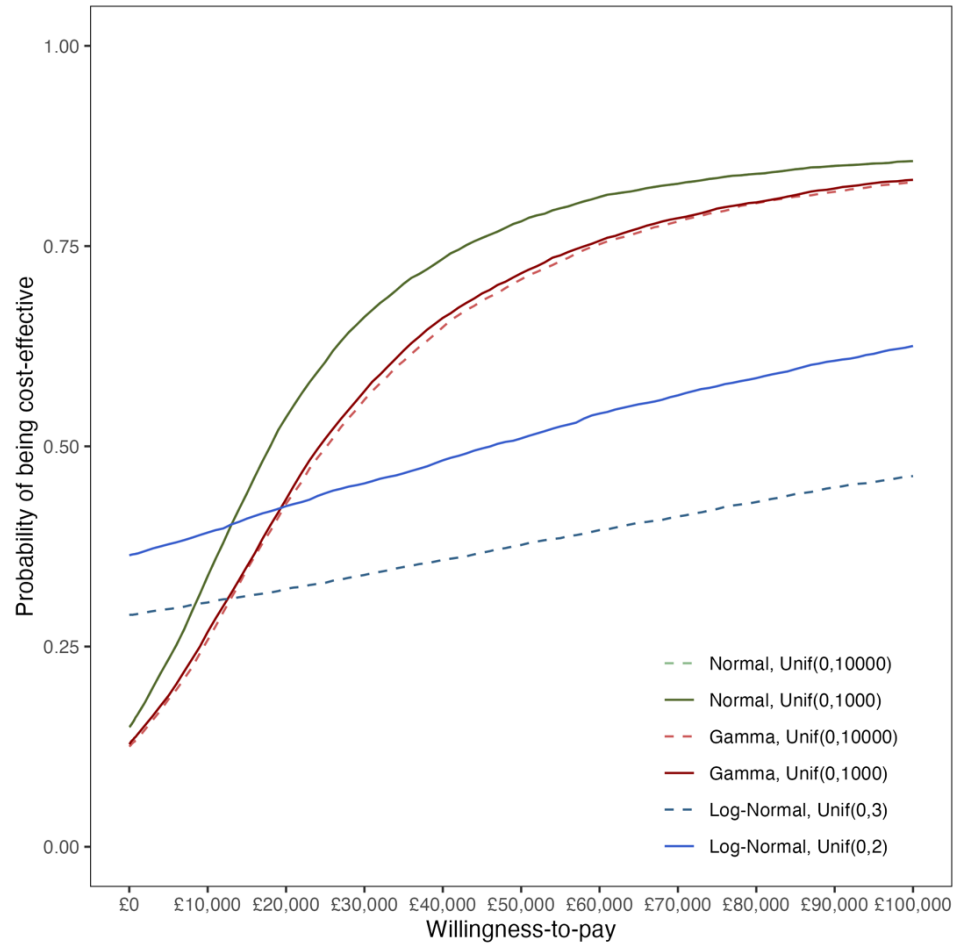

Fig. 3 Cost-effectiveness planes across Normal, Gamma and Log-Normal models with Uniform priors for cost standard deviations. The grey dashed line represents a willingness-to-pay threshold at £20,000 while the black dashed line represents a threshold at £30,000. The orange dot in each sub-figure denotes the incremental cost-

effectiveness ratio (ICER) of each model. The Uniform priors are placed on standard deviation of original costs in Normal and Gamma models while allocated to standard deviations of log costs in Log-Normal models.

The CEACs show that the intervention is more likely to be cost-effective under the Normal and Gamma models compared to the Log-Normal models when a willingness-to-pay threshold reaches £30,000 or above (Fig. 4). When Uniform prior distributions are used for the log cost standard deviations in the Log-Normal models, the uncertainty in incremental mean costs reduces the probability of the intervention being cost-effective. As the upper bound of the Uniform prior distribution decreases to 2, the probability that the intervention is more cost-effective than the control increases.
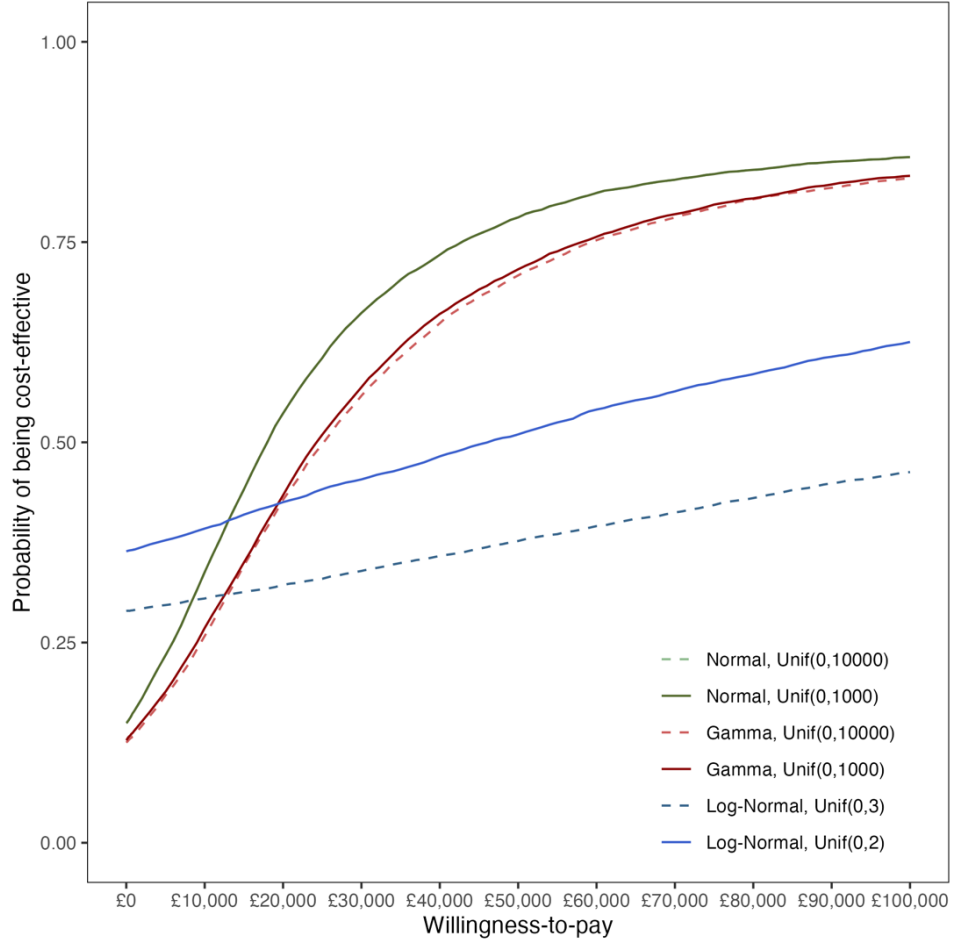
Fig. 4 Cost-effectiveness acceptability curves across Normal, Gamma and Log-Normal models with different Uniform prior distributions. The solid lines indicate results using Uniform priors with narrower ranges, i.e. Uniform(0,1000) for the Normal and Gamma models and Uniform(0,2) for the Log-Normal model; while the dashed lines represent results based on Uniform priors with wider ranges, i.e., Uniform(0,10000) for the Normal and Gamma models and Uniform(0,3) for the Log-Normal model. The green lines show CEACs from the Normal models; the red lines are for the Gamma models, and the blue lines are for the Log-Normal models.

4 Simulations

4.1 Simulation Overview

Building upon the findings of the case study, a simulation study has been designed to further evaluate the performance of the Bayesian CEA models – i.e. Normal, Gamma and Log-Normal models with different Uniform prior distributions on cost standard deviations. We investigate the impact of these models and prior choices on cost estimates in CEAs, including the mean cost estimates for each treatment arm and their incremental costs. The aim of this simulation study is to explore how sensitive the Bayesian CEA cost models are to the choice of priors for cost standard deviations and how this sensitivity might vary across different scenarios.

The simulation settings have been designed to reflect the common challenges that might be encountered in real-world CEA applications, motivated by our case study. We consider a cost-effectiveness analysis alongside a simple,

six-month two-arm RCT when generating individual-level cost-effectiveness data. In this simulation study, we assume that baseline utilities and QALYs are normally distributed, however, such choice does not influence the conclusions of the entire study as we focus on the impact of prior distributions on cost estimates.

4.2   Data Generating Process

Cost data are often skewed, and prone to zero values, indicating no use of health care services. The findings from the case study suggest that the occurrence of zero values in cost data may lead to poor performance of the Log-Normal cost model; however, the impact of the proportion of zeros on its relative performance remains unclear. Additionally, Bayesian analyses are known to be sensitive to the prior specification when sample sizes are small [33]. Therefore, a total of eight scenarios are defined based on the three dimensions on cost – data skewness, proportion of zeros and sample size:

- Data skewness: To account for different levels of skewness, the positive values in cost data are assumed to follow Gamma, and Log-Normal distributions, respectively.

- Proportion of zero values: Scenarios with relatively low proportions of zeros (varying between 0% and 10%) are considered to mimic real-world situations with either no or mild occurrences of zeros in cost data. For cases with higher proportions of zero values, a two-part model is recommended [34].

- Sample size: The total sample size is set to either 200 or 2000, representing 100 and 1000 participants per treatment arm, respectively, under the assumption of a 1:1 allocation rate between intervention and control groups. The sample size of 200 reflects the situation typified by the ORBIT trial, while the larger sample size of 2000 represents a typical high-sample-size scenario for trial-based CEAs.

Initially, QALYs are generated from a Normal distribution, and then cost data are simulated from a two-part model to determine the presence of zero values. Further details of the data generating process are provided in Appendix E due to space limitations.

We conduct 1000 simulations for each scenario and fit three CEA models with different Uniform priors using a Bayesian approach to align with the case study and allow a fair comparison within the simulation setting. Again, a constant value ($\epsilon = 1$) will be added to the Gamma and Log-Normal models in scenarios with zeros. The QALY model component in the Bivariate Normal model has been retained consistently for the Gamma and Log-Normal models.

The focus of this simulation is the estimation of mean cost per treatment arm and incremental costs. To assess the relative performance in terms of treatment-specific mean costs and incremental costs across the Normal, Gamma and Log-Normal models with different Uniform priors on cost standard deviations, we compare their bias, empirical standard error (Empirical SE) and root mean squared error (RMSE) in each simulated scenario. The definitions of these performance measures are provided in the Appendix E.

The Log-Normal distribution with Uniform priors whose ranges are more limited, i.e. Uniform (0,1) and a Uniform(0,0.8), cannot converge when cost data with a large sample size of 2000 follow a Gamma distribution and present no zeros. Data re-scaling has not been considered for this scenario as it may distort the original relationship between the data scale and the choice of prior, which does not align with the initial purpose of this study. Therefore, the Log-Normal models with these two Uniform priors will not be reported in this scenario.

4.3   Simulation Results

Figure 5 graphically summarises the performance of the assessed models across different distributional assumptions and Uniform prior distributions for cost standard deviations when cost data are log-normally distributed. As expected, when cost data follow a Log-Normal distribution, the occurrence of zeros strongly undermines the performance of the Log-Normal models, regardless of whether the sample size (as shown in Fig 5, Column A and C). Under these scenarios, the Log-Normal model is highly sensitive to the upper bound values of the Uniform priors for cost standard deviation, influencing the mean cost estimates in terms of more bias in each treatment arm.

In scenarios without zeros, the Log-Normal model still exhibits sensitivity to the prior selection. Specifically, estimates from Log-Normal models under wider Uniform priors – Uniform$(0,2)$ and Uniform$(0,3)$ – are less biased than those using narrower priors – i.e. Uniform$(0,1)$ and Uniform$(0,0.8)$. This finding suggests that the narrower Uniform prior may be overly restrictive for these scenarios. However, the Log-Normal models show robustness when the two wider Uniform prior distributions are used.

The Normal and Gamma models seem to present higher degree of robustness to their prior specifications compared to the Log-Normal models. When there is no zero in the data, the Normal models have greater bias and RMSE than the Log-Normal models with wider Uniform priors, irrespective of sample size. The Normal model performance improves compared to the Log-Normal models with narrower Uniform priors when the sample size is large. Surprisingly, the Gamma models achieve comparable performance to that of the Log-Normal models with wider Uniform prior distributions, even when data are log-normally distributed and are fully positive.
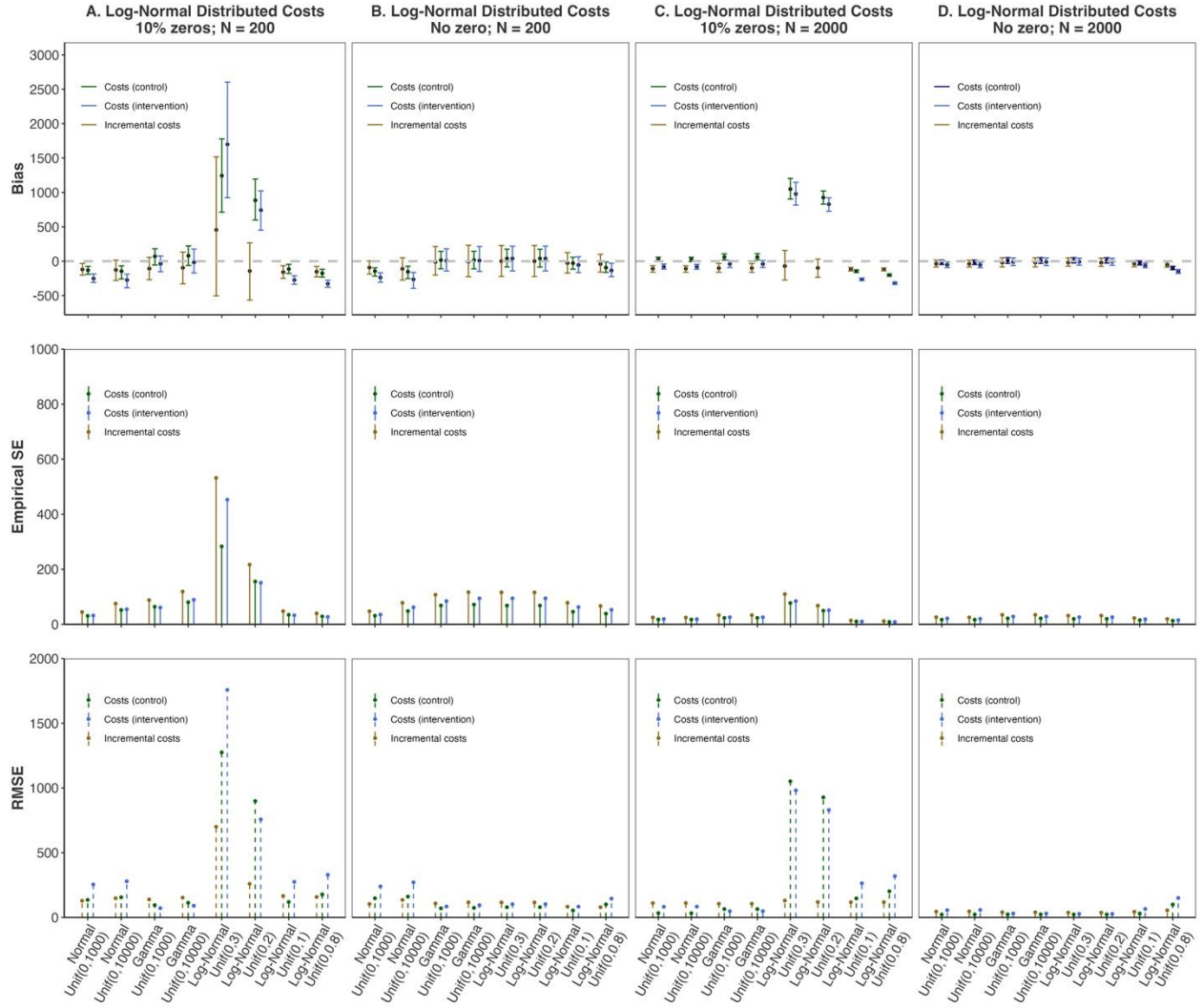
Fig. 5 Relative performance of the Normal, Gamma, and Log-Normal models with different Uniform prior distributions when cost data are log-normally distributed. Each sub-figure reports the performance of the models in terms of their bias (with its 95% credible intervals), empirical SE, and RMSE arranged from top to bottom, for the estimation of mean costs per treatment arm and incremental costs. The column A of sub-figures presents model performance under scenarios with 10% zero values and a small sample size of 200, while the column B shows scenarios with no zero values and the small sample size. The remaining two columns (C and D) report model performance for scenarios with a larger sample size of 2000.

	Scenarios			
Models	**10% Zeros,**	**No Zero,**	**10% Zeros,**	**No Zero,**
	N = 200	**N = 200**	**N = 2000**	**N = 2000**
Normal model				
Uniform(0,1000)	-127	-117	-116	-81
Uniform(0,10000)	-140	-134	-117	-82
Gamma model				
Uniform(0,1000)	-91	-80	-120	-66
Uniform(0,10000)	-56	-33	-110	-62
Log-Normal model				
Uniform(0,3)	459	977,821	-146	27,069
Uniform(0,2)	-320	68,408	-676	1,307
Uniform(0,1)	-171	13,904	-250	NA
Uniform(0,0.8)	-162	11,497	-229	NA

Table 2 Bias in incremental costs of Normal, Gamma, and Log-Normal models with different Uniform prior distributions when cost data follow a Gamma distribution.

In this simulation study, when the cost data follow a Gamma distribution, the relative performance of the Log-Normal models is inferior to the other models. Taking the mean bias in incremental costs as an example, as illustrated in Table 2, the Log-Normal models show greater bias and variations to their prior selections compared to the Normal and Gamma models, especially when the data contain no zero values. This is likely due to the Gamma distribution being able to generate small cost values close to zero, without the need to add a small constant to the original data when fitting the Log-Normal models in these scenarios. Given this mismatch between the true underlying cost distribution and the Log-Normal model specification, and considering the presence of near-zero values, the performance of the Log-Normal models deteriorates substantially and becomes extremely sensitive to the choice of priors.

In contrast, both Normal and Gamma models are more robust to the Uniform prior specification. They are generally less biased and more precise than the Log-Normal models across all four Gamma-distributed data scenarios. Full details on model relative performance, including bias in mean costs per treatment arm, empirical SE, and RMSE, are provided in Appendix E, given that they are similar to the findings for bias in incremental costs reported in Table 2.

5   Discussion

This study compares the impact of implementing Uniform prior distributions with different upper bounds on CEA results under a Bayesian modelling framework. These Uniform prior distributions are applied on log-scale cost standard deviations when costs are modelled using a Log-Normal distribution, and on original cost standard deviations when costs are assumed to follow a Normal or Gamma distribution. The analysis is initially based on a real case study with non-negligible proportion of patients having zero costs and then validated via a simulation

approach. Our study shows that Bayesian CEAs can be sensitive to the choice of upper bounds of the Uniform prior distributions for log cost standard deviations when cost data contain zero values and are assumed to be log-normally distributed. Unlike the Log-Normal distribution, the Normal and Gamma distributions have shown robustness to the upper bounds of Uniform distributions as the priors on the standard deviations of costs.

It is common to use Uniform distributions as the priors on standard deviations of cost data in Bayesian CEAs due to their benefits of conveying minimal information through the priors when limited knowledge about these parameters is available. The underlying belief is that given the lack of knowledge on variance parameters, the bounds for "minimally informative" Uniform priors are not expected to influence final conclusions substantially [35]. However, our study indicates the importance of recognising the appropriateness of Uniform prior distributions in applied Bayesian CEAs. The use of such priors can potentially bring "unwanted" information into the posterior inference and lead cost estimates towards unintended values.

One potential reason for the overly large cost estimates found in our study is that the Uniform prior distribution is applied on the standard deviation of costs on the log scale rather than on the original scale in the Log-Normal distribution. Indeed, concerning the mathematical property of the Log-Normal distribution, when a Uniform prior distribution is assigned to the standard deviation of log costs, the magnitude of costs on its original scale can possibly, explosively increase to an extreme high level. In other words, even small upper bounds of Uniform priors can yield large standard deviations on original costs and thus lead to mean cost estimates that are unlikely to occur in reality. Using Uniform distribution as a minimally informative prior for the purpose of letting data dominate posterior inference is clearly not effective in most of the scenarios explored in our analysis.

Another possible explanation is the characteristics of the actual data. Our case study features some typical characteristics of CEA data that may contribute to the poor performance of the Log-Normal model, which potentially drives the sensitivity of cost-effectiveness results to the Uniform prior distributions. These include small sample sizes, potential model misspecification for skewed data, and the existence of zero-cost observations. Our simulation shows that, if the data contain zero, the cost-effectiveness results would remain substantially sensitive to Uniform priors, even when sample size is increased and regardless of whether the data are generated from a Log-Normal or Gamma distribution. This suggests that the sample size and model misspecification are less likely to be the primary driver of the sensitivity observed. Surprisingly, the performance of the Log-Normal model is extremely unreliable when cost data follow a Gamma distribution. This aligns with a previous simulation study that investigates parametric modelling options for costs based on data generated from parametric distributions [25].

The excessively large cost estimates from our empirical analysis of trial data may be related to the existence of several observations with zero costs over time. In line with our case study, a previous study that focuses on Bayesian cost prediction using cohort data with excess zeros has also found that Log-Normal models could overestimate mean observed costs [36]. To date, very few Bayesian CEAs using Log-Normal distribution for costs report whether they encounter the same situation where a non-negligible proportion of zero occurs in the cost data. Our simulation study provides evidence that cost estimates – biased with the presence of zero values – can be unbiased and insensitive to appropriate Uniform prior distributions when all cost values are strictly positive. A related study that explores various modelling approaches to handle zero-cost data via a simulation approach supports our finding, highlighting the poor performance of the Log-Normal model under low, moderate and high proportions of zeros [34].

Based on the aforementioned results, we recommend that the sensitivity of cost-effectiveness results to Uniform prior distributions with different upper bounds is assessed in any CEA involving individual-level zero cost data. A conservative alternative is to rely on models that are more robust to the upper bounds of Uniform prior distributions, for example, Gamma model, or those that can address the zero-cost feature, such as hurdle model (also referred as two-part model) [12,35].

Despite the wide use of Bayesian methods in CEAs, no guidance exists that provides validated advice on prior elicitation from a statistical perspective, nor does it clearly illustrate the possible consequences resulting from the choice of prior distributions. Our study attempts to explore the influence of Uniform distributions as minimally informative prior distributions on standard deviations of costs in a CEA context through both a case study and a simulation approach. Limitations of our analysis include: first, the simulation study has not addressed whether and how much the poor performance of the Log-Normal model, and its sensitivity to the Uniform prior distributions, are influenced to the continuity correction for zero values, i.e. adding a constant of one to the original data. As a result, we could not explore any potential remedies for the issues introduced by CEA data containing zeros. Second, we did not compare different choices of minimally informative prior distributions and thus cannot make evidence-based recommendations for future applications. However, the current study structure – starting with an empirical analysis of real-world trial data and followed by a tailored simulation study – would be sufficient to support our primary goal, i.e. to reveal the potential impact of choosing Uniform prior distributions with large upper bounds as the log cost standard deviations of Log-Normal model in real applications.

6 Conclusion

Bayesian CEAs are sensitive to the choice of upper bounds of the Uniform prior distributions for standard deviations of log costs when they are assumed to follow a Log-Normal distribution. Our results suggest careful consideration should be given to the potential consequences of applying Uniform distributions on cost standard deviations as minimally informative priors based on both distributional model assumptions and actual characteristics of the data. It may be more conservative to consider alternative minimally or weakly informative priors to assess the robustness of posterior results.

Reference

[1] Briggs A. A Bayesian approach to stochastic cost-effectiveness analysis: an illustration and application to blood pressure control in type 2 diabetes. Int J Technol Assess Health Care 2001;17:69–82. https://doi.org/10.1017/S0266462301104071.

[2] Spiegelhalter D, Best N. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Stat Med 2003;22:3687–709. https://doi.org/10.1002/sim.1586.

[3] O'Hagan A, Stevens J. Bayesian methods for design and analysis of cost-effectiveness trials in the evaluation of health care technologies. Stat Methods Med Res 2002;11:469–90. https://doi.org/10.1191/0962280202sm305ra.

[4] Gabrio A, Baio G, Manca A. Bayesian Statistical Economic Evaluation Methods for Health Technology Assessment. Oxf. Res. Encycl. Econ. Finance, Oxford University Press; 2019.

[5] O'Hagan A, Stevens J. A framework for cost-effectiveness analysis from clinical trial data. Health Econ 2001;10:303–15. https://doi.org/10.1002/hec.617.

[6] Baio G. Bayesian methods in health economics. CRC Press; 2012.

[7] McCarron CE, Pullenayegum E, Marshall D, Goeree R, Tarride J-E. Handling uncertainty in economic evaluations of patient level data: A review of the use of Bayesian methods to inform health technology assessments. Int J Technol Assess Health Care 2009;25:546–54. https://doi.org/10.1017/S0266462309990316.

[8] Baio G. Bayesian models for cost-effectiveness analysis in the presence of structural zero costs. Stat Med 2014;33:1900–13. https://doi.org/10.1002/sim.6074.

[9] Manca A, Lambert P, Sculpher M, Rice N. Cost-Effectiveness Analysis Using Data from Multinational Trials: The Use of Bivariate Hierarchical Modeling. Med Decis Making 2007;27:471–90. https://doi.org/10.1177/0272989X07302132.

[10] McCarron CE, Pullenayegum E, Thabane L, Goeree R, Tarride J-E. The Impact of Using Informative Priors in a Bayesian Cost-Effectiveness Analysis: An Application of Endovascular versus Open Surgical Repair for Abdominal Aortic Aneurysms in High-Risk Patients. Med Decis Making 2013;33:437–50. https://doi.org/10.1177/0272989X12458457.

[11] Grieve R, Nixon R, Thompson S. Bayesian Hierarchical Models for Cost-Effectiveness Analyses that Use Data from Cluster Randomized Trials. Med Decis Making 2010;30:163–75. https://doi.org/10.1177/0272989X09341752.

[12] Gabrio A, Mason A, Baio G. A full Bayesian model to handle structural ones and missingness in economic evaluations from individual-level data: Handling structural ones and missingness in economic evaluations. Stat Med 2019;38:1399–420. https://doi.org/10.1002/sim.8045.

[13] Gabrio A, Daniels M, Baio G. A Bayesian parametric approach to handle missing longitudinal outcome data in trial-based health economic evaluations. J R Stat Soc Ser A Stat Soc 2020;183:607–29. https://doi.org/10.1111/rssa.12522.

[14] Lambert P, Sutton A, Burton P, Abrams K, Jones D. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 2005;24:2401–28. https://doi.org/10.1002/sim.2112.

[15] Fabrizi E, Trivisano C. Bayesian Conditional Mean Estimation in Log-Normal Linear Regression Models with Finite Quadratic Expected Loss. Scand J Stat 2016;43:1064–77. https://doi.org/10.1111/sjos.12229.

[16] Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal 2006;1. https://doi.org/10.1214/06-BA117A.

[17] Zellner A. Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression. J Am Stat Assoc 1971;66:327–30.

[18] Fabrizi E, Trivisano C. Bayesian Estimation of Log-Normal Means with Finite Quadratic Expected Loss. Bayesian Anal 2012;7. https://doi.org/10.1214/12-BA733.

[19] Hollis C, Hall CL, Jones R, Marston L, Novere ML, Hunter R, et al. Therapist-supported online remote behavioural intervention for tics in children and adolescents in England (ORBIT): a multicentre, parallel group, single-blind, randomised controlled trial. Lancet Psychiatry 2021;8:871–82. https://doi.org/10.1016/S2215-0366(21)00235-2.

[20] Weinstein M, Torrance G, McGuire A. QALYs: The Basics. Value Health 2009;12:S5–9. https://doi.org/10.1111/j.1524-4733.2009.00515.x.

[21] Thompson S, Nixon R. How Sensitive Are Cost-Effectiveness Analyses to Choice of Parametric Distributions? Med Decis Making 2005;25:416–23. https://doi.org/10.1177/0272989X05276862.

[22] Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Stat Assoc 1962;57:348–68.

[23] Barber J, Thompson S. Multiple regression of cost data: use of generalised linear models. J Health Serv Res Policy 2004;9:197–204. https://doi.org/10.1258/1355819042250249.

[24] Nixon R, Thompson S. Parametric modelling of cost data in medical studies. Stat Med 2004;23:1311–31. https://doi.org/10.1002/sim.1744.

[25] Briggs A, Nixon R, Dixon S, Thompson S. Parametric modelling of cost data: some simulation evidence. Health Econ 2005;14:421–8. https://doi.org/10.1002/hec.941.

[26] Basu A, Manca A. Regression Estimators for Generic Health-Related Quality of Life and Quality-Adjusted Life Years. Med Decis Making 2012;32:56–69. https://doi.org/10.1177/0272989X11416988.

[27] Plummer M. JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling. 3rd Int Workshop Distrib Stat Comput DSC 2003 Vienna Austria 2003;124.

[28] Su Y-S, Yajima M. R2jags: Using R to Run "JAGS" 2021.

[29] Gelman A, Rubin D. Inference from Iterative Simulation Using Multiple Sequences. Stat Sci 1992;7:457–72.

[30] Spiegelhalter D, Best N, Carlin B, van der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol 2002;64:583–639. https://doi.org/10.1111/1467-9868.00353.

[31] Black WC. The CE Plane: A Graphic Representation of Cost-Effectiveness. Med Decis Making 1990;10:212–4. https://doi.org/10.1177/0272989X9001000308.

[32] Van Hout BA, Al MJ, Gordon GS, Rutten FFH. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994;3:309–19. https://doi.org/10.1002/hec.4730030505.

[33] Little R, Rubin D. Statistical analysis with missing data. vol. 793. John Wiley & Sons; 2019.

[34] Mullahy J, Norton EC. Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero*. Oxf Bull Econ Stat 2024;86:417–47. https://doi.org/10.1111/obes.12583.

[35] Neelon B, O'Malley AJ, Smith VA. Modeling zero-modified count and semicontinuous data in health services research Part 1: background and overview: Modeling zero-modified count and semicontinuous data in health services research Part 1: background and overview. Stat Med 2016;35:5070–93. https://doi.org/10.1002/sim.7050.

[36] Cooper NJ, Lambert PC, Abrams KR, Sutton AJ. Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. Health Econ 2007;16:37–56. https://doi.org/10.1002/hec.1141.