Choice of statistical model for cost-effectiveness analysis and covariate adjustment: empirical application of prominent models and assessment of their results

Mantopoulos, Theodoros; Mitchell, Paul; Welton, Nicky J; McManus, Richard; Andronis, Lazaros

Link to publication on Research at Birmingham portal

Choice of statistical model for cost-effectiveness analysis and covariate adjustment: empirical application of prominent models and assessment of their results.

Theodoros Mantopoulos[a] , Paul Mitchell[b], Nicky J. Welton[a], Richard McManus[c], Lazaros Andronis[b,1]

[a] School of Social and Community Medicine, University of Bristol

[b] Health Economics Unit, University of Birmingham

[c] Nuffield Department of Primary Health Care Sciences, University of Oxford

Correspondence to: Health Economics Unit, Public Health Building, University of Birmingham, Birmingham, B15 2TT. Email: l.andronis@bham.ac.uk; Tel: +44 (0) 121 414 3197 Fax: +44 (0) 121 414 8969

Conflicts of interest: The authors are not aware of any financial or personal relationships between themselves and others that might be perceived by others as biasing this work.

Summary

Context: Statistical models employed in analysing patient-level cost and effectiveness data need to be flexible enough to adjust for any imbalanced covariates, account for correlations between key parameters, and accommodate potential skewed distributions of costs and/or effects. We compare prominent statistical models for cost-effectiveness analysis alongside randomised controlled trials (RCTs) and covariate adjustment to assess their performance and accuracy using data from a large RCT.

Methods: Seemingly unrelated regressions, linear regression of net monetary benefits, and Bayesian generalized linear models with various distributional assumptions were used to analyse data from the TASMINH2 trial. Each model adjusted for covariates prognostic of costs and outcomes.

Results: Cost-effectiveness results were notably sensitive to model choice. Models assuming normally distributed costs and effects provided a poor fit to the data, and potentially misleading inference. Allowing for a beta distribution captured the true incremental difference in effects and changed the decision as to which treatment is preferable.

Conclusions: Our findings suggest that Bayesian generalized linear models which allow for non-normality in estimation offer an attractive tool for researchers undertaking cost-effectiveness analyses. The flexibility provided by such methods allows the researcher to analyse patient-level data which are not necessarily normally distributed, while at the same time it enables assessing the effect of various baseline covariates on cost-effectiveness results.

Introduction

Clinical research has been seen as a key activity and a vital means of improving the health of the population. Experimental studies of a particular design—randomised controlled trials (RCTs)—have been described as *"the crown jewel"* of clinical research [1] and are considered as a prime source of input in assessing the effectiveness and cost-effectiveness of competing health care technologies [2]. Much of the rigor in the results of RCTs stems from the design characteristics of such studies. A key design characteristic, randomisation of participants across treatment arms, aims to distribute patients across treatment groups so that relevant patient characteristics are balanced across treatments, with a view to avoiding bias and ensuring that any observed outcomes are due to the assigned treatment [3, 4].

However, flaws in randomisation techniques, small numbers of participants, or simply chance, can lead to imbalances in baseline covariates [2, 8], which will inevitably result in biased effectiveness and cost-effectiveness estimates [9-11]. To safeguard against such bias, researchers often carry out covariate adjustments, which typically account for the confounding effect of baseline imbalance through regression techniques [12].

While such adjustments are commonplace in studies looking into the effectiveness of technologies, they are relatively less customary in cost-effectiveness analyses (CEA), despite the fact that covariate adjustment can reduce variation and give more accurate cost-effectiveness estimates, even when covariates are balanced [11, 13]. In addition to covariate adjustment, CEA analyses need to account for possible correlation between costs and effects [14] as well as the skewed distribution of costs [15].

In recognition of this, Hoch *et al*. [9] suggested the net benefit regression model, which is a linear regression model with net monetary benefits (NMBs) as a dependent variable allowing for the inclusion of covariates. NMBs are defined as a measure that combines costs and health outcomes by transforming health outcomes into monetary units, using as an exchange rate a hypothesised value of the decision makers' willingness to pay (WTP) for a unit of outcome [16]. Willan *et al*. [17] extended this work and considered costs and effects jointly, assuming a bivariate normal distribution, by proposing the use of a system of seemingly unrelated regressions (SUR). In essence, SUR represents a set of regression equations in which the error terms are assumed to be correlated across a set of regression equations [18].

A different route was taken by Nixon and Thompson [13] and Vazquez-Polo *et al*. [19] who, unlike Hoch *et al*. [9] and Willan *et al*. [17], adopted a Bayesian approach to provide covariate-adjusted cost-effectiveness estimates. Nixon and Thompson [13] and Vazquez-Polo *et al.* [19] extended Bayesian methods previously considered in the CEA literature [14, 20-22] to incorporate covariates. A regression model directly on effects and costs was proposed, in which patient characteristics were included as covariates, and effects were assumed to be correlated with costs. The relative advantage of adopting Bayesian methods is that distributions beyond the normal can be assigned to costs and effects resulting in more flexible approaches to estimation.

While a number of approaches have been proposed in the literature, no definite answers exist as to the most appropriate method for modelling cost-effectiveness data. Methods assuming normally distributed costs or effects are widely used in CEA and are advocated by Thompson and Barber and Nixon et al. [23, 24]. Inference is based on the sample means, which can be obtained from linear regression models modelling cost-effectiveness in the scale of interest. According to others, the extreme skewness typically observed in costs and/or effects needs to be acknowledged by either employing data transformations or using generalized linear models (GLMs) that can accommodate distributions more appropriate for skewed data [25, 26, 27]. Methods assuming normally distributed variables typically require relatively large sample sizes, non-extreme skewness and the absence of extreme outliers; unless these conditions are met, methods based on normal distribution are considered inappropriate for modelling cost-effectiveness data [28].

In either case, researchers setting out to analyse cost-effectiveness data will need to select the most appropriate method, as each dataset is unique and different statistical models may perform differently according to the characteristics of the data.

Costs typically exhibit positive skewness, or, in some cases, they may even be multimodal. Distributions typically used to accommodate positively skewed data are the gamma and the log-normal. Effects expressed in terms of quality-adjusted life-years (QALYs) are subject to similar idiosyncrasies. Data are usually truncated at both ends of the distribution (ranging between 0 and 1 when the time horizon is 1 year or less). Also, QALYs exhibit negative skewness with most values lying in the upper end of the measurement scale and some extreme outliers at the lower end of the scale. The beta distribution is a candidate for modelling data in the range (0, 1), while supporting both positive and negative skewed distributions. The gamma distribution can also be used to model effects, though effects are usually negatively skewed while the gamma distribution is

appropriate for modelling positive skewed data. To overcome this problem an alternative is to model 1-effects (conditional on effects ranging between o and 1) with the gamma distribution.

Important aspects in statistical analysis of cost-effectiveness data include selection of relevant covariates for subgroup analyses [29] as well as analyses of cost-effectiveness data which can incorporate covariate adjustment. The first refers to identifying the optimal decision for different subgroups of patients, as cost-effectiveness estimates may vary in these populations due to the presence of treatment-modifying covariates. The second type of analysis refers to performing an adjustment for any imbalanced covariates so that their confounding effect can be accounted for. In this case, the aim is to provide unbiased estimates in order to determine the optimal decision for the whole population. The latter is the focus of this study.

Given the above, this paper aims to compare the observed model fit and cost-effectiveness estimates of three prominent methods for CEA and covariate adjustment: a) ordinary least squares (OLS) regression of NMBs; b) SUR and c) GLMs with interaction between costs and effects for various distributions for costs (normal, gamma, log-normal) and effects (normal, gamma, beta). The choice of NMB and SUR models is based on their popularity and their easy application in standard statistical software. Bayesian GLMs are chosen for their flexibility to accommodate different distributions. Irrespective of the distribution of costs and effects and the correlation between them, all the methods considered would be unbiased if we were to replicate a given study a number of times in a simulation. That is, what would differ between each method is the precision of costs and effects and not their point estimates. In a particular study, we expect on average to obtain relatively similar cost-effectiveness estimates with different levels of precision, according to the fit of each model to the specific data. NMB regression and SUR models are expected to provide precise estimates when the cost-effectiveness data are approximately normally distributed and there are no extreme outliers. In such cases we could directly consider the NMB regression or the SUR, since these models are considerably easier to apply than the Bayesian GLMs. By considering and assessing different families of models and different underlying distributions, the model that best reflects the available data can be chosen. To illustrate the above methods, we used data from a large RCT, the TASMINH2 study, aiming to look into the effect of self-management compared with usual care of hypertension in West Midlands, UK. To make the methods accessible to applied researchers, supplementary online material with the code for fitting the Bayesian models is provided.

The remainder of the paper is structured as follows. The next section outlines the purposes and characteristics of the randomised clinical trial which provided data for this work, and describes the statistical models employed in this analysis. Section 3 reports the results obtained from each of the statistical models under assessment, while section 4 discusses and interprets the results, gives the key strengths and limitations in the analysis and draws conclusions specific to the trial data.

Materials and Methods

Case study: the TASMINH2 trial

The Tele-monitoring And Self-Management in the control of Hypertension (TASMINH2) randomised control trial was carried out to examine whether tele-monitoring and self-management of people with hypertension would lead to lower levels of blood pressure as compared to usual care [30]. The study took place in 24 general practices (GPs) in the West Midlands, United Kingdom and involved patients aged between 35 and 85 years old with blood pressure more than 140/90 mm Hg despite receiving antihypertensive treatment. Patient randomisation to tele-monitoring and self-management or usual care was stratified by GP, with minimisation factors including sex, baseline systolic blood pressure, and presence or absence of diabetes or chronic kidney disease. The trial's main clinical endpoint was change in mean systolic blood pressure between baseline and the two follow-up periods (6 months and 12 months). Further information about the trial can be found elsewhere [30].

Patient level data on resource use and quality of life collected as part of TASMINH2 were analysed to obtain estimates of the cost-effectiveness of each treatment. Five hundred twenty seven (n=527) patients were randomised to self-management (n=263) or usual care (n=264). Of those patients, 47 were excluded from complete case analyses as they did not attend follow-up visits at 6 and 12 months. In the present analysis, a further 17 observations were disregarded due to missing data for important covariates used in the analysis, giving a total number of 463 patients (88% of original sample size; 227 in self-managements arm, 236 in usual care arm). The per patient NHS cost over a 12-month period was estimated as the sum of the cost for medications, training and equipment, inpatient and outpatient care and GP visits. Alongside patient's mean systolic blood pressure (i.e. the principal clinical outcome in the RCT), the study collected patients' responses to EQ-5D-3L [31, 32] a generic measure of preference-based health-related quality of life. EQ-5D scores were

used to calculate QALYs from baseline to 12 months, using the 'area under the curve' (AUC) approach [33]. EQ-5D scores were calculated using the UK tariff [34].

Statistical models

Three different methods for covariate adjustment are considered for the CEA of the TASMINH2 RCT: OLS regression of NMBs, SUR, and generalized linear regression models with interaction between costs and effects, estimated under a Bayesian approach. The following notation is used: let c_i and e_i be the costs and effects for the ith individual.

OLS regression of NMBs

Net benefits can be calculated in order to convert costs and effects to a single variable and then be used in typical regression analyses. Hoch *et al*. [9] proposed a linear regression model, where NMB is the response variable with explanatory variables comprising an indicator for treatment arm plus the covariates of interest. That is:

$$NMB_i = a + \delta t_i + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \tag{1}$$

where, a is an intercept term, t_i a treatment dummy taking the value zero for the standard treatment and the value one for the new treatment, x_{ij} are the p covariates of interest, and ε is a stochastic error term. The regression coefficient δ represents the incremental net benefit (INB) attributable to the new treatment controlling for covariates, for that WTP level. The INB is the difference in the mean NMB of the new treatment and the mean NMB of the standard treatment.

Seemingly unrelated regressions

SUR is a system of different regression equations with error terms that are assumed to be correlated across the equations [35]. Different sets of covariates can be included in each equation, allowing for a more flexible modelling approach to estimation.

$$c_i = a^c + \delta^c t_i + \beta_1^c x_{i1} + \cdots + \beta_p^c x_{ip} + \varepsilon_i^c$$

$$e_i = a^e + \delta^e t_i + \beta_1^e x_{i1} + \cdots + \beta_p^e x_{ip} + \varepsilon_i^e \quad \begin{pmatrix} \varepsilon_i^c \\ \varepsilon_i^e \end{pmatrix} \sim BVN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_c^2 & \rho \sigma_c \sigma_e \\ & \sigma_e^2 \end{pmatrix} \right) \tag{2}$$

where, a is the intercept term in each model, t a treatment dummy taking the value zero for the standard treatment and the value one for the new treatment, x are the p covariates of interest, and ε are the stochastic error terms in each model. The regression coefficient δ^c represents the incremental cost (*IC*) attributable to the new treatment controlling for covariates, and the regression coefficient δ^e represents the incremental effect (*IE*) attributable to new treatment, again, controlling for covariates [17]. The error terms (ε) are assumed to follow a bivariate normal distribution, with mean zero and variances σ_c^2 and σ_e^2, while ρ represents the correlation between costs and effects, conditional on covariates.

Generalized linear regression models with interaction between costs and effects

Cost and effect data are frequently non-normally distributed, which can be estimated using GLMs [35]. A Bayesian approach provides a flexible way to estimate non-normal models. Five models with different underlying distributions for costs and effects are examined to show the iterative process of changing one distribution at a time. In this way, the impact from changing the distribution at either costs or effects is evidenced in the results. The fitted distributions for costs and effects are described next together with the respective models.

   a) *Model with gamma distribution on costs and normal on effects*

Nixon and Thompson [13] and Vasquez-Polo *et al*. [19] described a model for covariate adjustment for CEA of patient level data using normal distribution for effects and gamma distribution for costs with likelihoods

$$e_i \sim Normal(\varphi_{e,i}, \sigma_e^2), \qquad c_i \sim Gamma(a_c, \lambda_{c,i}), \ \lambda_{c,i} = a_c/\varphi_{c,i}$$

where the treatment and covariate effects are linear on the mean effects and mean costs

$$\varphi_{e,i} = a^e + \delta^e t_i + \beta_1^e x_{i1} + \cdots + \beta_p^e x_{ip}$$

$$\varphi_{c,i} = a^c + \delta^c t_i + \beta_1^c x_{i1} + \cdots + \beta_p^c x_{ip} + \beta_{p+1}^c (e_i - \varphi_{e,i}) \tag{3}$$

where, $\beta^e = (\beta_1^e, \ldots, \beta_p^e)$, $\beta^c = (\beta_1^c, \ldots, \beta_{p+1}^c)$ are vectors of unknown coefficients and precision term $\tau_e = 1/\sigma_e^2$ for effects. Correlation between costs and effects is allowed by including the term $\beta_{p+1}^c e_i$ in the above

equation. The subtraction of $\varphi_{e,i}$ from e_i in the above equation is done so that the interpretation of a^c remains the same as the overall mean cost of the control arm of the trial. For the same reason, when covariates $x_1, \dots, x_p$ are continuous, they are centred on their mean by subtracting each covariate value from their overall mean $(x_{ip} - \bar{x}_p)$. In the presence of categorical covariates, dummy variables equal to the number of the categories of each covariate are included in the above model (e.g. 2 dummy variables for a dichotomous covariate). In this case, constraints on the coefficients of each covariate are needed so that their sum over all the trial population is zero and the interpretation of a^c and a^e remains the same [13].

The expected *IE* attributable to the new treatment, controlling for covariates is given by coefficient δ^e and the expected *IC* attributable to the new treatment controlling for covariates is given by

$$IC = \delta^c + \beta_{p+1}^c \delta^e$$

Under a Bayesian estimation framework, the simultaneous prior distribution on coefficients α, β, δ, precision term τ_e and shape parameter α_c must be specified. We use the following prior structure so that the influence of prior distributions on the model estimates is minimal

$$\alpha \sim Normal(0, 100), \beta \sim Normal(0, \ 100), \delta \sim Normal(0, \ 100)$$

$$\tau_e \sim Uniform(0, \ 10), a_c \sim lnNormal(0, 100)$$

b) *Model with normal distribution on costs and beta on effects*

For the model under a beta distribution for effects, a GLM with a logit link function is chosen due to the unit-support of the outcome variable (QALYs). The following likelihoods and equations are applied

$$e_i \sim Beta(a_e, \lambda_{e,i}) \quad \lambda_{e,i} = a_e(1 - \varphi_{e,i})/\varphi_{e,i}$$

$$c_i \sim (\varphi_{c,i}, \sigma_c^2)$$

where the treatment and covariate effects are linear on the log-odds (logit) of the mean effects and mean costs

$$logit(\varphi_{e,i}) = a^e + \delta^e t_i + \beta_1^e x_{i1} + \dots + \beta_p^e x_{ip}$$

$$\varphi_{c,i} = a^c + \delta^c t_i + \beta_1^c x_{i1} + \dots + \beta_p^c x_{ip} + \beta_{p+1}^c (e_i - \varphi_{e,i}) \tag{4}$$

In model (4) the expected *IE* is estimated by

$$IE = \frac{e^{(a^e + \delta^e)}}{1 + e^{(a^e + \delta^e)}} - \frac{e^{a^e}}{1 + e^{a^e}}$$

and the expected *IC* is estimated by

$$IC = \delta^c + \beta^c_{p+1}\left(\frac{e^{(a^e + \delta^e)}}{1 + e^{(a^e + \delta^e)}}\right) - \beta^c_{p+1}\left(\frac{e^{a^e}}{1 + e^{a^e}}\right)$$

With regard to prior distributions, again, log-normal distributions are assigned to shape parameters and normal distributions to regression coefficients and the precision term.

c) *Model with gamma distribution on costs and beta on effects*

The following likelihoods and equations are applied assuming a gamma distribution on costs and a beta on effects

$$e_i \sim Beta\left(a_e, \lambda_{e,i}\right) \quad \lambda_{e,i} = a_e(1 - \varphi_{e,i})/\varphi_{e,i}$$

$$c_i \sim Gamma\left(a_c, \lambda_{c,i}\right) \quad \lambda_{c,i} = a_c/\varphi_{c,i}$$

where the treatment and covariate effects are linear on the log-odds of the mean effects and mean costs:

$$logit(\varphi_{e,i}) = a^e + \delta^e t_i + \beta^e_1 x_{i1} + \cdots + \beta^e_p x_{ip}$$

$$\varphi_{c,i} = a^c + \delta^c t_i + \beta^c_1 x_{i1} + \cdots + \beta^c_p x_{ip} + \beta^c_{p+1}(e_i - \varphi_{e,i}) \tag{5}$$

In model (5), the expected *IE* and *IC* are estimated as in model (4).

d) *Model with gamma distribution on costs and on 1-effects*

To fit a gamma distribution on costs and 1-effects the following likelihood and model is used

$$(1 - e_i) \sim Gamma\left(a_e, \lambda_{e,i}\right) \quad \lambda_{e,i} = a_e/\varphi_{e,i}$$

$$c_i \sim Gamma\left(a_c, \lambda_{c,i}\right) \quad \lambda_{c,i} = a_c/\varphi_{c,i}$$

where the treatment and covariate effects are linear on the mean (1- effects) and mean costs:

$$\varphi_{e,i} = a^e + \delta^e t_i + \beta^e_1 x_{i1} + \cdots + \beta^e_p x_{ip}$$

$$\varphi_{c,i} = a^c + \delta^c t_i + \beta^c_1 x_{i1} + \cdots + \beta^c_p x_{ip} + \beta^c_{p+1}(e_i - \varphi_{e,i}) \tag{6}$$

The expected *IE* is given by $-\delta^e$ (negative because the model is on 1-effects) and the expected *IC* estimate is given as in model (3).

e) _Model with log-normal distribution on costs and beta on effects_

The final model we consider is a log-normal model for costs and a beta model for effects:

$$e_i \sim Beta\big(a_e, \lambda_{e,i}\big) \quad \lambda_{e,i} = a_e(1 - \varphi_{e,i})/\varphi_{e,i}$$

$$\log(c_i) \sim Normal(\varphi_{c,i}, \sigma_c^2)$$

where the treatment and covariate effects are linear on the log-odds of the mean effect and log-mean costs:

$$logit(\varphi_{e,i}) = a^e + \delta^e t_i + \beta_1^e x_{i1} + \cdots + \beta_p^e x_{ip}$$

$$\varphi_{c,i} = a^c + \delta^c t_i + \beta_1^c x_{i1} + \cdots + \beta_p^c x_{ip} + \beta_{p+1}^c (e_i - \varphi_{e,i}) \tag{7}$$

The expected _IE_ is given as in model (4) and the expected _IC_ is given as

$$IC = e^{a^c + \delta^c + \beta_{p+1}^c \left(\frac{e^{(a^e + \delta^e)}}{1 + e^{(a^e + \delta^e)}} \right)} - e^{a^c + \beta_{p+1}^c \left(\frac{e^{a^e}}{1 + e^{a^e}} \right)}$$

Model comparison

Each of the statistical methods described above were applied to adjust for covariates which are prognostic of costs and effects. The comparison of statistical models was aided by obtaining values of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [37]. Lower AIC and BIC values indicate improved model fit and are preferred to higher values. The standard error (SE) of the expected INB at a willingness-to-pay level of £20,000 per additional QALY is also reported.

OLS regression and SUR models were estimated in STATA software, version 12.1 [38], while the Bayesian models were implemented using Markov Chain Monte Carlo (MCMC) methods in WinBUGS software [39]. For the Bayesian models, two parallel chains were used with different starting values. Posterior distributions for the parameters of interest were derived from 60,000 iterations of the Markov chain, after an initial burn-in of 20,000 iterations. History and autocorrelation plots were examined to ensure convergence was achieved.

Results

Descriptive analyses

The distributions of costs for the control and intervention groups are presented in Figures 1.a and 1.b, respectively, together with fitted densities from the normal distribution. It is obvious from Figure 1 that costs are positively skewed (median costs in intervention arm £367, interquartile range £228 to £558; median costs in control arm £229, interquartile range £109 to £467) and that the normal distribution fits poorly the data. Effectiveness data are illustrated in Figure 2. QALYs exhibit negative skewness (median QALYs in intervention arm 0.848, interquartile range 0.739 to 0.962; median QALYs in control arm 0.9194, interquartile range 0.796-1.000). Again, the normal distribution provided a poor fit of the data. A low level of correlation between costs and effects was found in the descriptive analysis (-0.10).

Table 1 reports the balance of baseline characteristics between the two treatment groups, measured as per cent standardised mean differences (SMDs)[1], which is invariant to sample size [41]. There is not a pre-specified level of imbalance that should be a concern, but a SMD of more than 10% is considered to be important [41, 42]. Apart from the SMD, the correlation between each covariate with costs and QALYs for the two treatment groups is reported. As it can be seen, baseline EQ-5D scores were imbalanced (SMD=30.51%), while BMI and coronary kidney disease were slightly imbalanced (SMD=10.83% and SMD=13.86%, respectively). With regard to correlations with endpoints, baseline EQ-5D scores were strongly correlated with QALYs in both treatment groups ($r_{control}$ = 0.77, $r_{intervention}$ = 0.88), while for the other covariates a low level of correlation or no correlation was found with either costs or QALYs.

The findings in Table 1 motivate the need to adjust for the imbalanced and correlated with endpoints covariates. Therefore, baseline EQ-5D is included in the analysis while BMI and coronary kidney disease are not, as despite the slight imbalance between the treatment arms they are not prognostic of costs or outcomes.

Results of statistical models

---

[1] The formula for calculating the SMD for a continuous covariate (x) is: $SMD_x = \frac{\mu_{x1} - \mu_{x2}}{\sqrt{(var_{x1} + var_{x2})/2}}$, where μ_{x1}, μ_{x2} and var_{x1}, var_{x2} are the means and variances for each group [40].

Assessment of the overall cost-effectiveness of the TASMINH2 trial using the different regression methods is provided in Table 2. Generally, each model reports different cost-effectiveness estimates leading to different reimbursement decisions in some cases. Model fit, reported in the AIC and BIC measures, also varies across the methods and so does the precision of the estimates. Models allowing for the skewness in the observations report improved fit and in some cases report more precise estimates.

The NMB regression model achieves the worst model fit (AIC and BIC values of 8488 and 8501, respectively) with the expected INB estimated at 14.1. The largest level of uncertainty is also reported in the NMB model, with the standard error (SE) of the expected INB being equal to 217.1. The SUR model reports a better performance in terms of both model fit and precision (AIC and BIC values of 6718 and 6739, respectively; SE of expected INB equal to 215.8). The expected INB estimate is equal to -5.2, indicating that the standard treatment is cost-effective under the SUR model.

Allowing for a gamma distribution on costs and a normal distribution on effects in the Bayesian GLM resulted in further improve in both the model fit and the precision of the expected INB. The respective AIC and BIC values are 6094 and 6127 while the SE of the expected INB is equal to 203.7, indicating that the gamma distribution is more appropriate for modelling the cost data. The expected INB is equal to 51.1 under this model. The change in the expected INB compared to the SUR model is driven from the change in the expected incremental cost.

Allowing for the skewness in the effect data by using a beta distribution also results in improved model fit. That is, the GLM having a normal distribution on costs and a beta distribution on effects provides better fit of the data; however, the expected INB is less precise compared to the gamma-normal distributed GLM. The AIC and BIC values are 5987 and 6020, respectively and the SE of the expected INB is equal to 212.4. A possible explanation for the increased uncertainty in the expected INB relates to the flexibility of the beta distribution in fitting different types of data. This flexibility in handling QALYs results in improved fit at the expense of reduced precision. Modelling effects with a beta distribution results in negative incremental effect, indicating that the new treatment is less effective than the current one. The expected INB in this model is -282.2.

The same pattern is noticed in the SE of the expected INB in the model having a gamma distribution on costs and a beta on effects. Even though an improved fit of the data is observed compared to all previous models (AIC and BIC values of 5222 and 5255, respectively), the SE of the expected INB is equal to 210.8. That is, less

accurate results are obtained from the models using the beta distribution to model the effect outcome compared to the remaining models for the effect data. The expected INB is estimated at -330.8 due to the negative incremental effect. Applying the beta distribution on effects changes the decision rule and in this case the current treatment dominates the new treatment as it is less costly and more effective.

Allowing for a gamma distribution on both costs and effects results in worst model fit compared to the gamma-beta distributed model (AIC and BIC values of 5461 and 5494, respectively). This indicates that the beta distribution is more appropriate for modelling the effect data. However, the precision of the expected INB is considerably improved (SE equal to 199.0) resulting in the most precise estimates of all the models under consideration. Using a gamma distribution on effects results in an incremental effect of 0.0082 which is similar to the incremental effect estimate obtained using a normal distribution. The expected INB under this model is 94.8.

Finally, allowing for a log-normal distribution on costs and a beta on effects results in the best model fit with AIC and BIC values of 4801 and 4834, respectively. However, the precision of the INB estimate (SE of 214.1) is less than all the remaining Bayesian GLMs. Despite th**is**, the expected INB is more precise than that of the NMB and SUR models. The expected INB under this model is estimated at -378.0, similar to the other two models with beta distributed effects.

Discussion

This study explores the appropriateness of three prominent methods for covariate adjustment in cost-effectiveness analyses: OLS regression of NMBs, SUR, and generalized linear regression with interaction between costs and effects. Each of the methods was applied to patient-level cost and health outcome data from the TASMINH2 trial. Distributions other than the normal (gamma and log-normal on costs, gamma and beta on QALYs) were fitted for the purposes of these analyses. Prognostic factors of costs and QALYs were considered as covariates with only baseline EQ-5D found to be a significant predictor of QALYs.

Findings suggest that cost-effectiveness inferences are sensitive to the statistical model employed, and therefore an assessment of model fit is essential. Despite the different INB estimates obtained from each model, such differences are not statistically significant and they are driven from small differences in the

estimates of incremental QALY. On the basis of the available data from TASMINH2, we found that OLS regression of NMBs gave a poor fit to the data. Without taking into consideration the skewed distributions of costs and effects, the SUR model provided relatively good fit of the trial data. Moreover, considering that its application requires less time and effort than Bayesian GLM models, SUR can be a preferred modelling approach in circumstances where costs and effects are approximately normally distributed.

Distributions of cost data usually exhibit a high degree of skewness and other idiosyncrasies such as non-negative values and heteroskedasticity) [43]. In such occurances, based on the findings from our study, non-normal distributions could be applied to costs as model fit and accuracy is improved. While this finding is in agreement with conclusions in previous studies [13, 26, 44, 45], it is not in line with previous findings stipulating that methods that assume normality are reasonably robust to skewed cost data [17, 46]. In this particular dataset, extreme outliers in costs resulted in poor fit to the normal distribution. On the contrary, the Bayesian GLMs were more robust to these outliers, reporting more precise estimates, in accordance with findings from Cantoni and Ronchetti [47].

Distributions of QALYs can also present the same idiosyncrasies observed in cost data [48]. Therefore, methods that extend beyond the normal distribution could be applied. In this specific example, fitting a gamma distribution to effects improved the goodness of fit of the model compared to a normal distribution; yet, it was not enough to capture the negative difference in mean QALYs between the two groups. The beta distribution provided further improvement in model fit and captured the true incremental difference in QALYs. This finding is in line with previous research suggesting that beta regression models are superior to different regression techniques [48-50]. Overall, the best fit of the TASMINH2 trial dataset was provided by a GLM specified under a Bayesian approach, allowing for a log-normal distribution on costs, a beta distribution on QALYs and controlling for baseline EQ-5D scores.

An unexpected finding was the performance of the Bayesian GLMs in terms of both the fit of the data and the precision of the estimates. Modelling effect data with a beta distribution resulted in considerably improved model fit, but the precision of the expected incremental effect and consequently the expected INB was reduced. A potential explanation on the loss of precision is employing a logit link function to model effects which required back-transforming data to the original scale. The issue of loss of precision from transformation of data has already been considered in terms of the analysis of cost data [28]. However, in a Bayesian MCMC

framework, the transformation is calculated at each iteration of the simulation and this is not expected to lead to less precise estimates, only to slower convergence of the Markov Chain. Having carefully examined convergence in the models under consideration, we believe that it is the flexibility of the beta distribution in modelling QALYs that provides improved fit of the data, while giving less precise estimates. In any case, better model fit of the data does not always imply more precise estimates. For instance, in some circumstances more flexible statistical models may provide better fit and this flexibility in fitting the observed data results in less precise estimates. Random-effect models, typically used in meta-analyses of trial data, are an example of models that can provide a better fit while giving less precise estimates compared to fixed-effects models [51].

Although Bayesian GLMs are extremely flexible in handling different types of datasets their application requires considering some aspects. First, it should be highlighted that such bivariate models are only an approximation of the true joint distribution of costs and effects by recognising the correlation between them. There is no reassurance that the combination of a marginal (effects) with a conditional (costs) model will converge to their true joint distribution, as the properties of the bivariate distributions considered here are not well known. Prior distributions require attention as their impact on the posterior estimates should be minimum, although in large datasets their impact diminishes. Prior distributions should also be chosen so that model parameters do not cause costs or effects to lie outside their appropriate bounds (e.g. become negative). Consideration should also be given on convergence of the posterior estimates, as with model complexity increasing converge may become slower and larger number of iterations may be required. A final consideration regards issues of autocorrelation in the simulation due to high correlation between the intercept and slope parameters. This correlation results in poor mixing of the MCMC chains, which in turn results in lack of convergence [52]. However, centring of effects and covariates at their mean values, as discussed in the Materials and Methods section, would solve such problem. It must be noted that the correct method of model selection from a family of possible models under a Bayesian paradigm is through their associated Bayes factors [53]. Whilst Bayes factors are very helpful in model selection, they are complex to compute, and not available from an MCMC simulation in WinBUGS, which is the most commonly used tool for such analyses. In addition, presenting the Bayes factor of each Bayesian GLM does not allow a comparison with the frequentist models (NMB regression and SUR). Presenting model fit in terms of AIC and BIC together with visual inspection of the data is a common method of model comparison, which can result in robust conclusions.

Covariates can be incorporated in the analysis for the additional reason of assessing the cost-effectiveness of interventions at a more individualised level, by examining whether different subgroups of the population are associated with different cost-effectiveness estimates. This is owing to the fact that effects, or even costs, may be modified by a covariate and as a consequence the choice of an optimal intervention may vary for different values of the covariate [29]. While testing for subgroup effects, especially for subgroups that have not been pre-specified in the trial protocol, is viewed with some suspicion in clinical effectiveness studies [54-57], subgroup analyses have been actively encouraged in cost-effectiveness studies [58]. Such analyses give policy makers the flexibility not only to identify the optimal treatment for the trial population, but also to make more 'individualised' decisions for subgroups of the trial population. The case study in our analysis focuses on identifying the optimal intervention for a population akin to the trial population, nonetheless, results for patient subgroups can also be obtained by extending the models to consider treatment with covariate interaction terms.

Data from the TASMINH2 trial have informed a recent economic evaluation by Kaambwa *et al*. [59]. The authors developed a Markov model to predict the costs and health effects associated with usual care and self-management of hypertension over a 35 year time horizon. The results of this study suggest that self-management is cost-effective for both men and women, with a probability of cost-effectiveness at £20000 per QALY exceeding 0.99 for both genders. While there is disagreement between the very appealing ICER values cited by Kaambwa *et al*. [59] and the results obtained from the present study, there are important differences between these studies which render any comparisons between them potentially misleading. While the analysis by Kaambwa *et al*. [59] makes use of a decision model which is populated by estimates of clinical progression taken from various studies, the present study is based solely on patient-level data from TASMINH2 trial, and relates to a follow-up period of 12 months. It must be noted that, rather than conducting an economic evaluation to highlight differences in the results between dissimilar studies, our study aimed to assess the performance of different statistical approaches.

The above findings are specific to the TASMINH2 RCT and different models and methods may perform better in different situations. To draw solid conclusions on the relative performance of each model, simulated data that compare these models across a range of circumstances that may be faced by researchers should be employed. However, such work is beyond the scope of this paper and could be the objective of a future study.

Future studies should also examine the relative performance of QALYs generated by other measures, including the longer version EQ-5D 5 level that has recently been developed [60].

Our findings illustrate that cost-effectiveness results can be sensitive to the choice of model and distributional assumptions. We would therefore recommend that a wide variety of modelling assumptions are considered, and model fit is thoroughly assessed and taken into account when selecting a model for analysis. This should be coupled with visual inspection of the empirical distribution of the cost and effect observations. In particular, this application has shown that methods based on Bayesian approaches that allow for non-normality in estimation, offer an attractive alternative for cost-effectiveness analyses and covariate adjustment. The flexibility provided by employing such methods allows the researcher to explore different underlying distributions and baseline covariates in order to identify an optimal methodology. On this basis, it is thought that the use of such methods in economic evaluations of healthcare technologies warrants more attention.

References

1. Fuchs, V., R., Garber, A., M.: The new technology assessment. New England Journal of Medicine. **323**(10), 673 (1990)

2. Drummond, M., F., Sculpher, M., J., Torrance, G., W., O'Brien, B., J., Stoddart, G., L.: Methods for the economic evaluation of health care programmes (3rd ed.). Oxford University Press, Oxford (2005)

3. Senn, S., J.: Covariate imbalance and random allocation in clinical trials. Statistics in Medicine. **8**, 467-475 (1989)

4. Moher, D., Hopewell, S., Schulz, K., F., Montori, V., Gøtzsche, P., C., Devereaux, P., J., Elbourne, D., Egger, M., Altman, D., G.: ConSoRT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. British Medical Journal. **340**, 869-897 (2010)

5. Akobeng, A., K.: Understanding randomised controlled trials. Archives of disease in childhood. **90**:, 840-844 (2005)

6. Taves, D., R.: Minimization – New method of assigning patients to treatment and control groups. Clinical pharmacology and therapeutics. **15**(5), 443-453 (1974)

7. Pocock, S., J., Simon, R.: Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics. **31**(1), 103-115 (1975).

8. Senn, S., J.: Testing for baseline balance in clinical trials. Statistics in Medicine. **13**, 1715-1726 (1994)

9. Hoch, J., Briggs, A., Willan, A.: Something old, something new, something borrowed, something blue: A framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economics.* **11**, 415– 430 (2002)

10. Manca, A., Hawkins, N., Sculpher, M.: Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics. **14**, 487–496 (2005)

11. Willan, A., Briggs, A.: Statistical Analysis of Cost-Effectiveness Data. Wiley, Chichester (2006)

12. McNamee, R.: Regression modelling and other methods to control for confounding. Occupational and Environmental Medicine. **62**, 500-506 (2005)

13. Nixon, R., M., Thompson, S., G.: Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. Health Economics. **14**, 1217–1229 (2005)

14. O'Hagan, A., Stevens, J., W.: A framework for cost-effectiveness analysis from clinical trial data. Health Economics. **10**, 303–315 (2001)

15. Briggs, A., Gray, A.: The distribution of health care costs and their statistical analysis for economic evaluation. Journal of Health Services Research and Policy. **3**, 233–245 (1998)

16. Tambour, M., Zethraeus, N., Johannesson, M.: A note on confidence intervals in cost-effectiveness analysis. International Journal of Technology Assessment in Health Care. **14**(3), 467–471 (1998)

17. Willan, A., R., Briggs, A., H., Hoch, J., S.: Regression methods for covariate adjustment and subgroup analysis for non-censored cost–effectiveness data. Health Economics. **13**, 461–475 (2004)

18. Greene, W., H.: (2002) Econometric Analysis (5th ed.). Prentice Hall, New Jersey (2002)

19. Vazquez-Polo, F.J., Negrin, M., Gonzalez, B.: (2005) Using covariates to reduce uncertainty in the economic evaluation of clinical trial data. Health Economics. **14**, 545–557 (2005)

20. Briggs, A., H.: A Bayesian approach to stochastic cost-effectiveness analysis. Health Economics. **8**, 257–261 (1999)

21. Chaloner, K., Rhame, F., S.: Quantifying and documenting prior beliefs in clinical trials. Statistics in Medicine. **20**, 581–600 (2001)

22. O'Hagan, A., Stevens, J., W.: Bayesian assessment of sample size for clinical trials of cost-effectiveness. Medical Decision Making. **21**, 219–230 (2001)

23. Thompson, S., G., Barber, J., A. How should cost data in pragmatic randomised trials be analysed? British Medical Journal **320** (7243): 1197-1200 (2000)

24. Nixon, R., M., Wonderling, D., Grieve, R., D. 2010. Non-parametric methods for cost-effectiveness analysis: the Central Limit theorem and the bootstrap compared. Health Economics **19**(3): 316–333 (2010)

25. Manning, W., G., Mullahy, J. Estimating log models: to transform or not to transform? Journal of Health Economics **20**(4): 461–494 (2001)

26. O'Hagan, A., Stevens, J., W.: Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality? Health Economics. **12,** 33-49 (2003)

27. Basu, A. Extended generalised linear models: simultaneous estimation of flexible link and variance functions. The Stata Journal **5**(4): 501-516 (2005)

28. Mihaylova, B., Briggs, A., O'Hagan, A., Thompson, S., G. Review of statistical methods for analysing healthcare resources and costs. Health Economics, **20**(8): 897-916 (2011)

29. Moreno, E., Giron, F., J., Vazquez-Polo, F., J., Negrin, M., A.: Optimal healthcare decisions: The importance of the covariates in cost-effectiveness analysis. European Journal of Operational Research. **218**, 512-522 (2012)

30. McManus, R., J., Mant, J., Bray, E., P., Holder, R., Jones, M., I., Greenfield, S., Kaambwa, B., Banting, M., Bryan, S., Little, P., Williams, B., Hobbs, F., D., R.: Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. The Lancet. **376**(9736), 163-172 (2010)

31. The EuroQol Group: EuroQol-a new facility for the measurement of health-related quality of life. Health Policy. **16**(3), 199-208 (1990)

32. Brooks, R.: EuroQol: the current state of play. Health Policy. **37**(1), 53-72 (1996)

33. Billingham L., J., Abrams, K., R., Jones, D., R.: Methods for the analysis of quality-of-life and survival data in health technology assessment. Health Technology Assessment. **3**(10) (1999)

34. Dolan, P. Modeling valuations for EuroQol health states. Medical Care. **35**(11), 1095–1108 (1997).

35. Zellner, A.: An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. Journal of the American Statistical Association. **57**, 348–368 (1962)

36. McCullough, P., Nelder, J., A.: Generalized Linear Models (2nd ed.). Chapman & Hall, London (1989)

37. Anderson, D., R. Model based Inference in the Life Sciences. Springer: New York (2008)

38. Stata: Stata Programming Reference Manual, Version 11. StataCorp, Texas (2009)

39. Spiegelhalter, D., J., Thomas, A., Best, N.: WinBUGS, version 1.2. MRC Biostatistics Unit, Cambridge (1999)

40. Gomes, M., Grieve, R., Nixon, R., Edmond S., W., Carpenter, J., Thompson, S., G.: Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. Health Economics. **21**, 1101–1118 (2012)

41. Austin, P., C.: Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statistics in Medicine. **28**, 3083-3107 (2009)

42. Rosenbaum, P., R., Rubin, D., B.: Constructing a control-group using multivariate matched sampling methods that incorporate the propensity score. The American Statistician. **39**, 33–38 (1985)

43. Altman, D., G.: Comparability of randomised groups. Statistician. **34**(1), 125–136 (1985)

44. Zhou, X., H.: Inferences about population means of health care costs. Statistical Methods in Medical Research. **11**, 327-339 (2002)

45. Nixon, R., M., Thompson, S., G.: Parametric modelling of cost data in medical studies. Statistics in Medicine. **23**, 1311–1331 (2004)

46. Nixon, R., M., Wonderling, D., Grieve, R., D.: Non-parametric methods for cost-effectiveness analysis: the central limit theorem and the bootstrap compared. Health Economics. **19**, 316-333 (2010)

47. Cantoni, E., Ronchetti, E. A robust approach for skewed and heavy-tailed outcomes in the analysis of health care expenditures. Journal of Health Economics **25**(2): 198-213 (2006)

48. Basu, A., Manca, A.: Regression Estimators for Generic Health-Related Quality of Life and Quality-Adjusted Life Years. Medical Decision Making. **32**, 56-69 (2012)

49. Kieschnick, R., McCullough, B., D.: Regression analysis of variates observed on (0,1): percentages, proportions and fractions. Statistical Modelling. **3**, 193–213 (2003)

50. Smithson, M., Verkuilen, J.: A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods. **11**(1), 54–71 (2006)

51. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

52. Welton, N., J., Sutton, A., J., Cooper, N., J., Abrams, K., R., Ades, A., E. Evidence Synthesis for Decision Making in Healthcare. John Wiley: Chichester, UK (2012)

53. Bernardo, J., Smith, A., F., M. Bayesian Theory. John Wiley: New York. (1994)

54. European Agency for the Evaluation of Medicinal Products (EMEA)-Committee for Proprietary Medical Products (CPMP): Points to consider on adjustment for baseline covariates. http://www.emea.europa.eu/pdfs/human/ewp/286399en.pdf. (2003). [Accessed 24 August 2013]

55. Cook, T., D., DeMets, D., L.: Introduction to statistical methods for clinical trials. Chapman & Hall/CRC, Boca Raton (2008)

56. Pocock, S., J., Assmann, S., E., Enos, L., E., Kasten, L., E.: Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Statistics in Medicine. **21**, 2917-2930 (2002)

57. Sun, X., Ioannidis, J., P., A., Agoritsas, T., Alba, A., C., Guyatt., G.: How to Use a Subgroup Analysis. Users' Guides to the Medical Literature. JAMA. **311**(4), 405-411 (2014)

58. Sculpher, M., J. Subgroups and Heterogeneity in Cost-Effectiveness Analysis. Pharmacoeconomics. **26**(9), 799-806 (2008)

59. Kaambwa, B., Bryan, S., Jowett, S., Mant, J., Bray, E., P., Hobbs, R., F., D., Holder, R., Jones, M., I., Little, P., Williams, B., McManus, R., J. Telemonitoring and self-management in the control of hypertension (TASMINH2): a cost-effectiveness analysis. European Journal of Preventive Cardiology. 0(00), 1–14 (2013)

60. Herdman, M., Gudex, C., Lloyd, A. Development and preliminary testing of the new five level version of EQ-5D (EQ-5D-5L). Quality of Life Research. **20**(10), 1727–1736 (2011)

Baseline characteristics	Intervention group (n=227)	Control group (n=236)	Standardised difference (%)	Correlation with endpoints	
Baseline EQ-5D, mean (SD)	0.81 (0.21)	0.88 (0.19)	30.51	$r_{cost\,1} = -0.08$	$r_{cost\,0} = -0.04$
				$r_{QALYs\,1} = 0.87$	$r_{QALYs\,0} = 0.77$
Male, n(%)	105 (46.30%)	108 (45.80%)	1.11	$r_{cost\,1} = -0.15$	$r_{cost\,0} = -0.11$
				$r_{QALYs\,1} = -0.07$	$r_{QALYs\,0} = -0.08$
Age (years), mean (SD)	66.10 (8.73)	65.70 (8.86)	4.34	$r_{cost\,1} = -0.01$	$r_{cost\,0} = 0.07$
				$r_{QALYs\,1} = -0.14$	$r_{QALYs\,0} = -0.01$
Baseline SBP*, mean (SD)	151.96 (11.92)	151.66 (11.88)	2.49	$r_{cost\,1} = 0.05$	$r_{cost\,0} = 0.03$
				$r_{QALYs\,1} = 0.04$	$r_{QALYs\,0} = -0.01$
Ethnicity: White, n(%)	218 (96.00%)	228 (96.60%)	3.14	$r_{cost\,1} = -0.04$	$r_{cost\,0} = 0.34$
				$r_{QALYs\,1} = -0.04$	$r_{QALYs\,0} = -0.09$
Height, mean (SD)	1.66 (0.10)	1.66 (0.09)	0.11	$r_{cost\,1} = 0.09$	$r_{cost\,0} = -0.02$
				$r_{QALYs\,1} = -0.01$	$r_{QALYs\,0} = 0.05$
Body mass index, mean (SD)	29.46 (5.68)	30.06 (5.47)	10.83	$r_{cost}\,1 = 0.02$	$r_{cost}\,0 = -0.05$
				$r_{QALYs\,1} = -0.18$	$r_{QALYs\,0} = -0.10$
Marital status: Married (n%)	171 (77.00%)	172 (73.00%)	3.71	$r_{cost\,1} = 0.06$	$r_{cost\,0} = -0.05$
				$r_{QALYs\,1} = -0.08$	$r_{QALYs\,0} = -0.12$
Occupation: Managerial, n(%)	96 (42.30%)	106 (44.90%)	5.81	$r_{cost\,1} = -0.06$	$r_{cost\,0} = 0.06$
				$r_{QALYs\,1} = 0.06$	$r_{QALYs\,0} = 0.02$
IMD** 2007 score, mean (SD)	16.85 (13.38)	17.31 (13.88)	3.39	$r_{cost\,1} = -0.02$	$r_{cost\,0} = 0.01$
				$r_{QALYs\,1} = -0.14$	$r_{QALYs\,0} = -0.11$
Current smoker, n(%)	16 (7.10%)	14 (6.00%)	4.53	$r_{cost\,1} = -0.06$	$r_{cost\,0} = 0.04$
				$r_{QALYs\,1} = 0.03$	$r_{QALYs\,0} = -0.10$
Alcohol intake in last year, n(%)	184 (81.10%)	188 (79.70%)	3.51	$r_{cost\,1} = 0.02$	$r_{cost\,0} = 0.03$
				$r_{QALYs\,1} = 0.29$	$r_{QALYs\,0} = 0.10$
Past medical history:					
Coronary kidney disease, n(%) (CKD)	16 (7.10%)	26 (11.00%)	13.86	$r_{cost\,1} = 0.00$	$r_{cost\,0} = 0.22$
				$r_{QALYs\,1} = 0.04$	$r_{QALYs\,0} = -0.01$
Diabetes, n(%)	18 (7.90%)	16 (6.80%)	4.39	$r_{cost\,1} = 0.01$	$r_{cost\,0} = 0.03$
				$r_{QALYs\,1} = -0.01$	$r_{QALYs\,0} = -0.06$

* Systolic blood pressure; **Index of multiple deprivation

Table II. Estimates and model fit of statistical models applied to TASMINH2 data

N=463	NMB regression	SUR	Bayesian regression models with different underlying distributions (Costs - QALYs)				
			gamma - normal	normal - beta	gamma - beta	gamma - gamma	log-normal - beta
	Mean [SE]	Mean [SE]	Mean [SE]	Mean [SE]	Mean [SE]	Mean [SE]	Mean [SE]
Incremental cost	Not applicable	107.57 [72.36]	68.25 [7.91]	13.39 [9.89]	68.30 [7.82]	68.43 [7.89]	110.90 [29.07]
Incremental QALY	Not applicable	0.0051 [0.0099]	0.0060 [0.0102]	-0.0134 [0.0106]	-0.0131 [0.0105]	0.0082 [0.0099]	-0.0133 [0.0106]
INB at λ = £20,000	14.1 [217.1]	-5.2 [215.8]	51.1 [203.7]	-282.2 [212.4]	-330.8 [210.8]	94.8 [199.0]	-378.0 [214.1]
AIC	8488	6718	6094	5987	5222	5461	4801
BIC	8501	6739	6127	6020	5255	5494	4834

INB: Incremental net benefit; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion;
NMB: Net monetary benefit; SUR: Seemingly unrelated regression; QALYs:
quality-adjusted life years; SE: standard error.

Figure 1.a Costs distribution for Control group
Normal fitted distribution

Figure 1.b Costs distribution for Intervention group
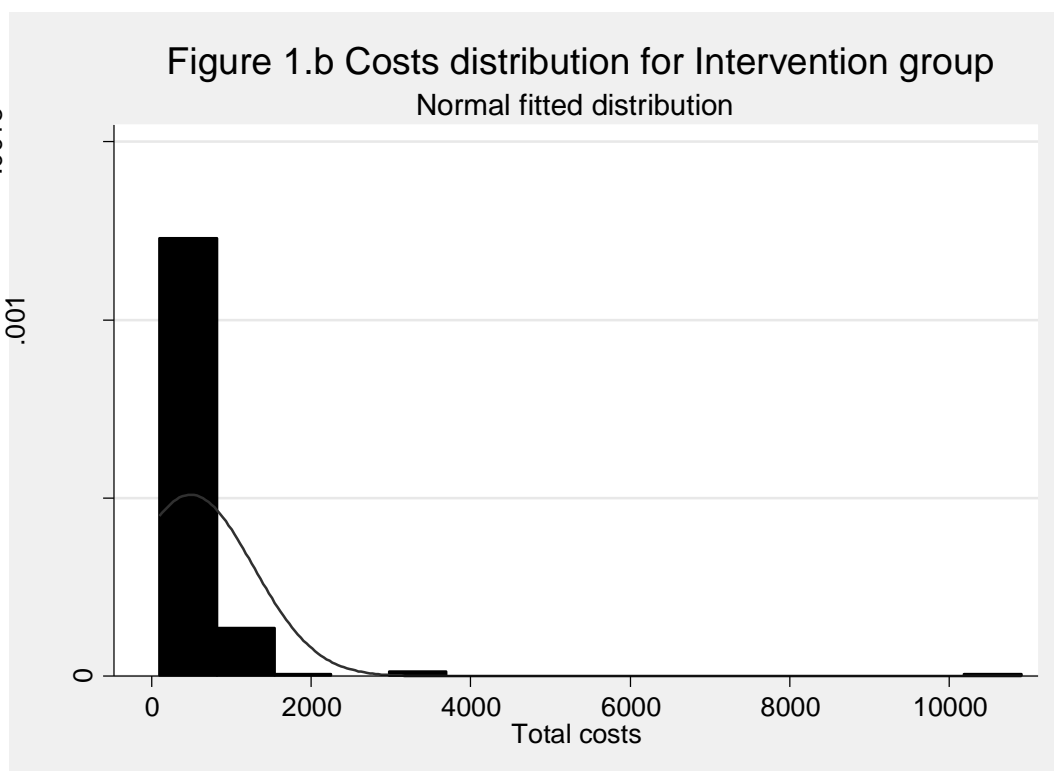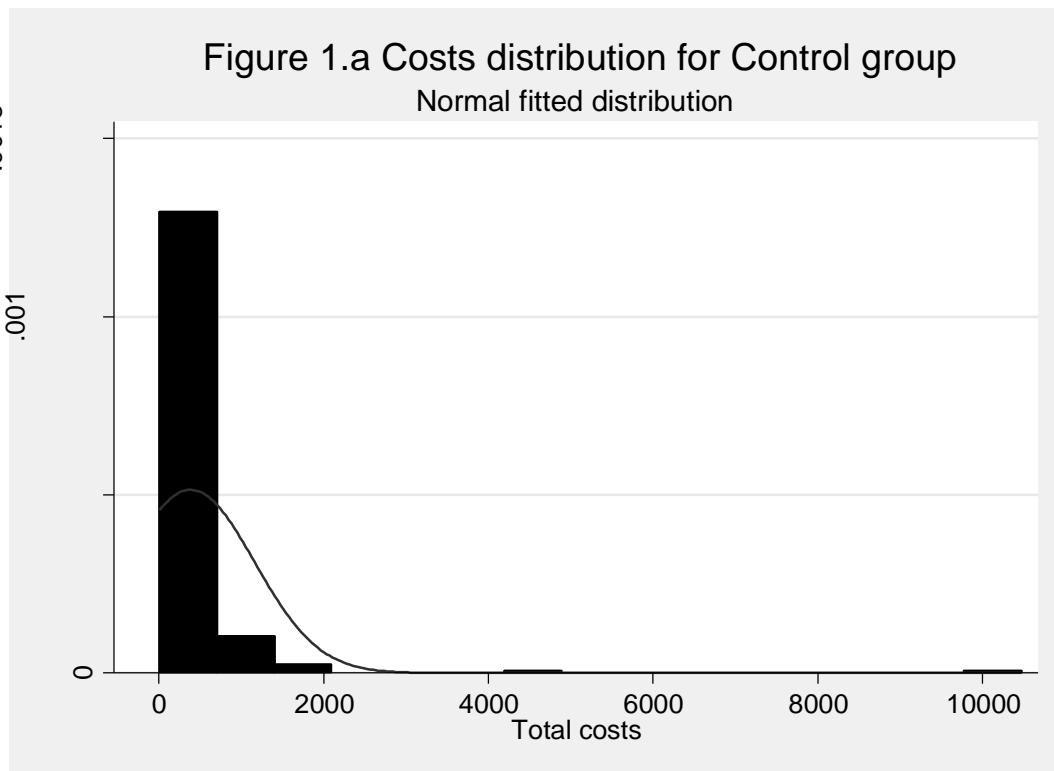Normal fitted distribution

Figure 2.a QALYs distribution for Control group
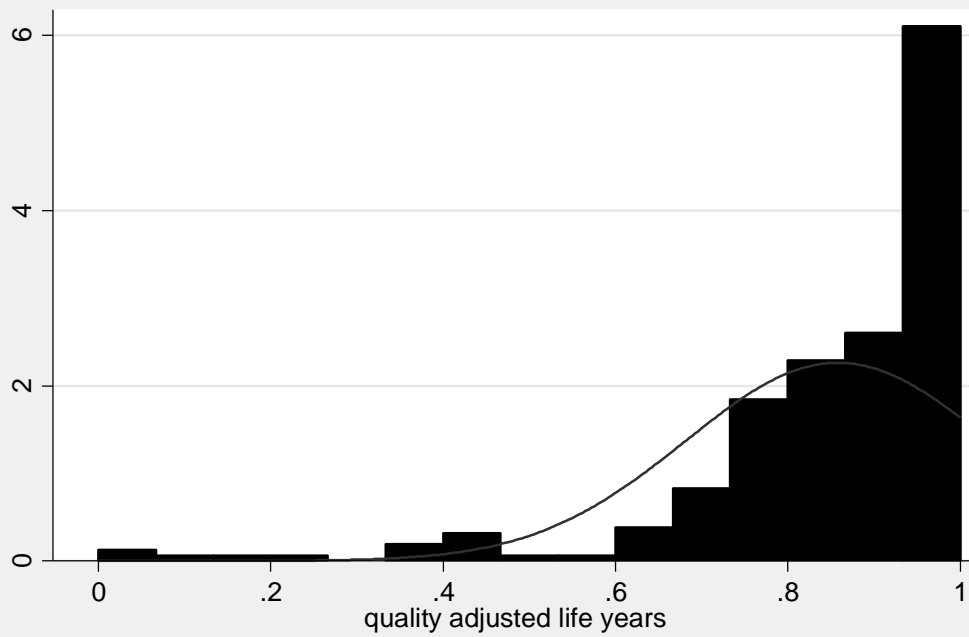Normal fitted distribution

Figure 2.b QALYs distribution for Intervention group
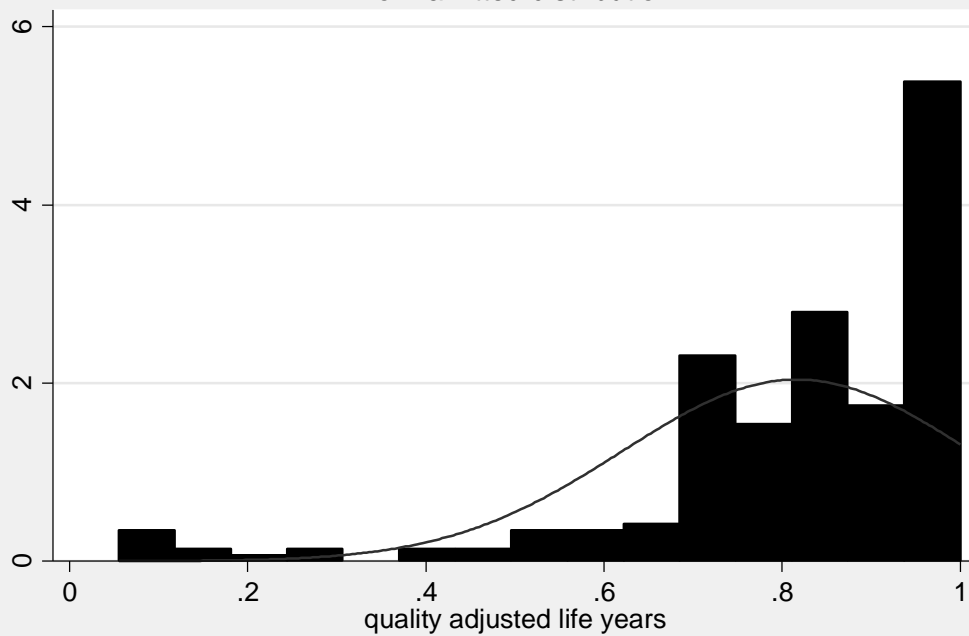Normal fitted distribution

Figure 3a Net Benefits distribution for Control group
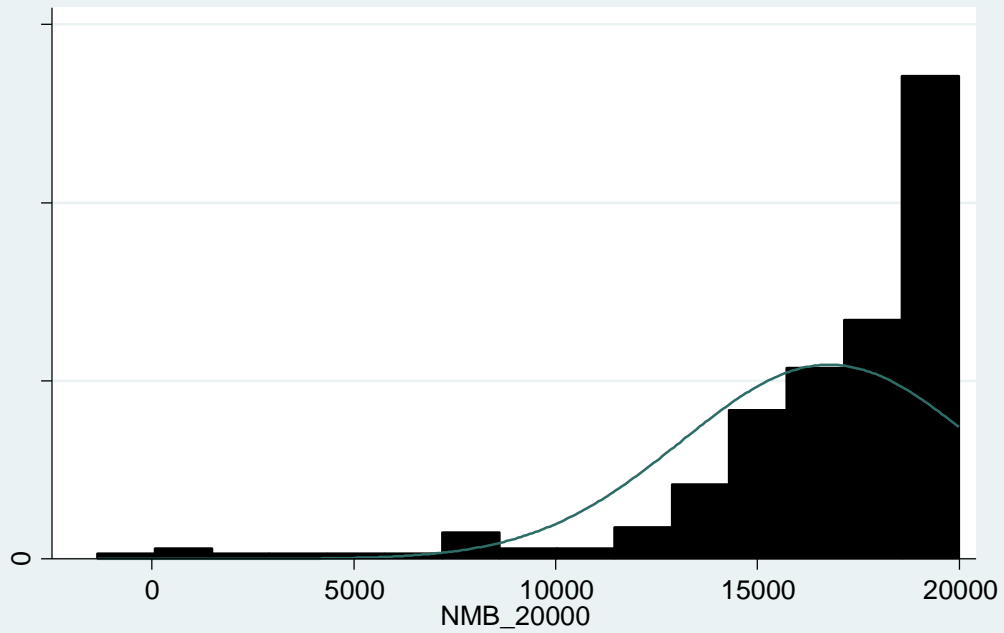Normal fitted distribution

Figure 3b Net Benefits distribution for Intervention group
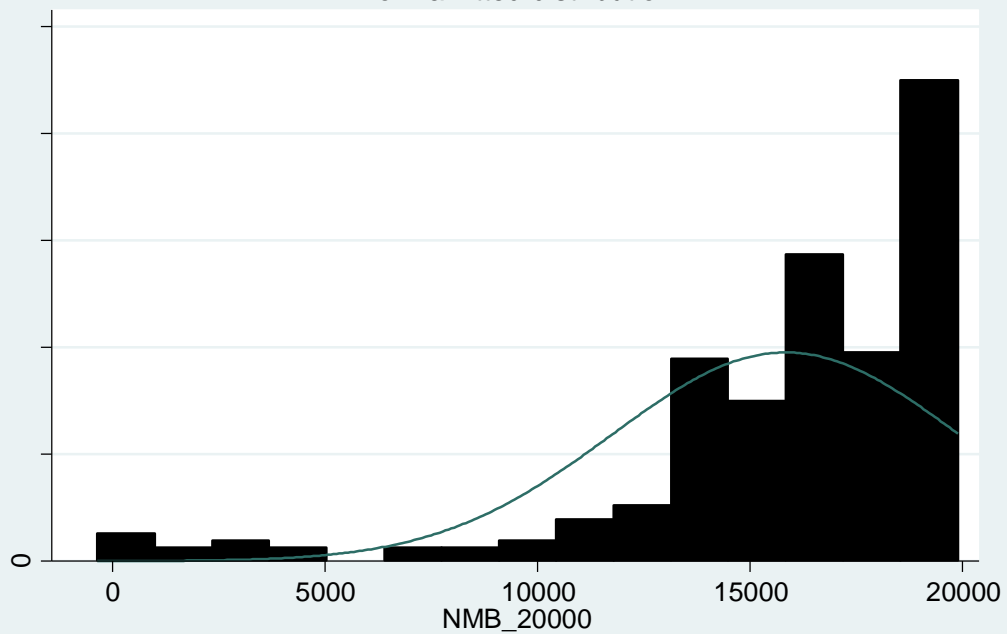Normal fitted distribution

Figure 4a 1-QALYs distribution for Control group
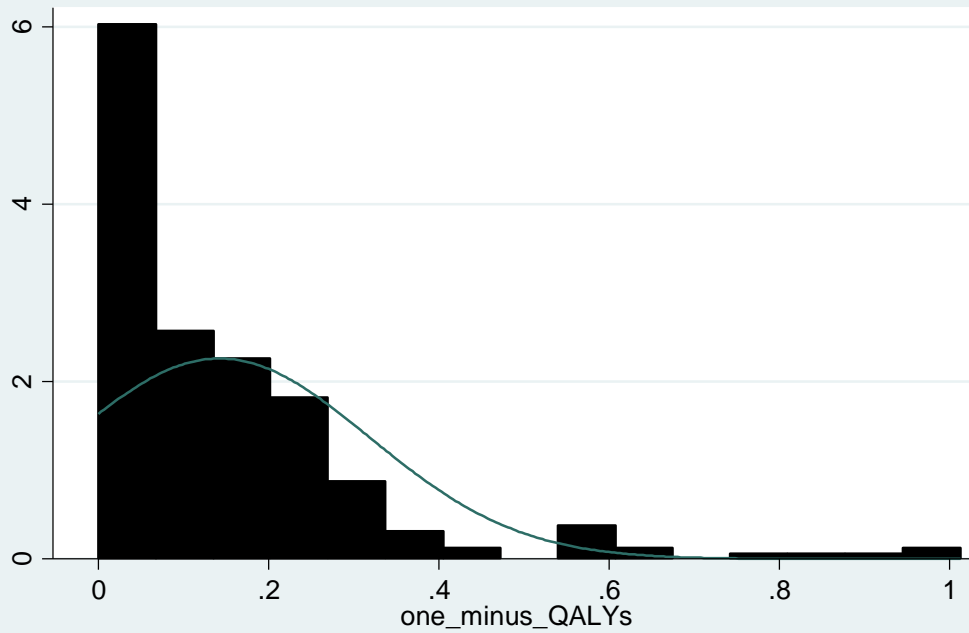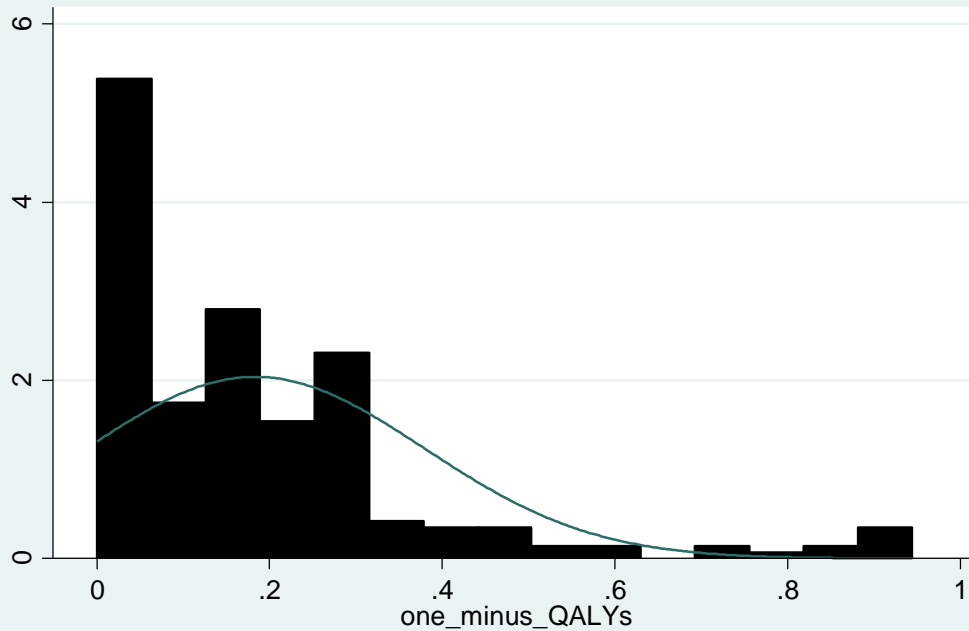Normal fitted distribution

Figure 4b 1-QALYs distribution for Intervention group
Normal fitted distribution

Model 1: gamma distribution on costs and normal on effects

```
model{
for (i in 1:N){

        efic[i] ~ dnorm(phi.e[i], tau1)

        cost[i]~dgamma(shape.c, rate.c[i])
        rate.c[i]<-shape.c/phi.c[i]

        phi.e[i] <- mu.e + beta[1]*x1[i] + beta[2]*(x2[i]-mean(x2[]))
        phi.c[i] <- mu.c + delta[1]*x1[i] + delta[2]*(efic[i]-phi.e[i])
        }


#covariate prior distributions
for (m in 1:2){
beta[m]~dnorm(0, 0.01);
delta[m]~dnorm(0, 0.01);
}

#prior distributions
tau1~dunif(0, 8)
log(shape.c)<-log.shape.c
log.shape.c~dnorm(0,0.01)


mu.e~dnorm(0,0.01)

mu.c~dnorm(0,0.01)

#Incremental QALY
effect1<-mu.e
effect2<-mu.e + beta[1]
diff.e <- effect2 - effect1

#Incremental Cost
cost1<- mu.c + delta[2]*effect1
cost2<-mu.c + delta[1] + delta[2]*effect2
diff.c<- cost2 - cost1

#ICER
ICER<-diff.c/diff.e

#INB lamda=20000
INB<-diff.e*20000-diff.c

}

# Starting values.
list(tau1=0.1, log.shape.c=0, mu.e=1, mu.c=1,
beta =c(1,1),
delta=c(1,1))

list(tau1=0.5, log.shape.c=1 mu.e=0.1, mu.c=0.1,
beta=c(0.1,0.1),
delta=c(0.1,0.1))
```

Model 2: normal distribution on cost and beta on effects

```
model{
for (i in 1:N){
        efic[i] ~ dbeta(shape.a, shape.b[i])
        shape.b[i] <- (shape.a*(1-phi.e[i]))/phi.e[i]

        cost[i]~dnorm(phi.c[i], tau)

        logit(phi.e[i]) <- mu.e + beta[1]*x1[i] + beta[2]*(x2[i]-mean(x2[]))
        phi.c[i] <- mu.c + delta[1]*x1[i]  + delta[2]*(efic[i]-phi.e[i])
                }

#covariate prior distributions
beta[1]~dnorm(0,0.1)
beta[2]~dnorm(0,0.1)
delta[1]~dnorm(0,0.1)
delta[2]~dnorm(0,0.1)


#prior distributions
tau~dunif(0, 8)

mu.e~dnorm(0, 0.01)
mu.c~dnorm(0,0.01)

log(shape.a)<-log.shape.a
log.shape.a~dnorm(0, 0.01)

#Incremental QALY
logit(effect1) <- mu.e
logit(effect2) <- mu.e + beta[1]
diff.e<- effect2 - effect1

#Incremental Cost
cost1<-mu.c + delta[2]*effect1
cost2 <- mu.c + delta[1] + delta[2]*effect2
diff.c<- cost2 - cost1

#ICER
ICER<- diff.c/diff.e

#INB lamda=20000
INB<- diff.e*20000- diff.c
}

# Starting values.
list(tau=0.1, mu.e=5, mu.c=2, log.shape.a=0.1,
beta =c(0.1,0.1),
delta=c(0.1,0.1))

list(tau=0.5, mu.e=-1, mu.c=0.2, log.shape.a=0.5,
beta =c(-1,1),
delta=c(1,-1))
```

Model 3: gamma distribution on costs and beta on effects

```
model{
for (i in 1:N){
        efic[i] ~ dbeta(shape.a, shape.b[i])
        shape.b[i] <- (shape.a*(1-phi.e[i]))/phi.e[i]

        cost[i]~dgamma(shape.c, rate.c[i])
        rate.c[i]<-shape.c/phi.c[i]

        logit(phi.e[i]) <- mu.e + beta[1]*x1[i] + beta[2]*(x2[i]-mean(x2[]))
        phi.c[i] <- mu.c + delta[1]*x1[i]  + delta[2]*(efic[i]-phi.e[i])
        }


#covariate prior distributions
for (m in 1:2){
beta[m]~dnorm(0, 0.01)
delta[m]~dnorm(0, 0.01)
}

#prior distributions
log(shape.c)<-log.shape.c
log.shape.c~dnorm(0,0.01)

mu.e~dnorm(0,0.01)
mu.c~dnorm(0,0.01)

log(shape.a)<-log.shape.a
log.shape.a~dnorm(0, 0.01)

#Incremental QALY
logit(effect1) <- mu.e
logit(effect2) <- mu.e + beta[1]
diff.e<- effect2 - effect1

#Incremental Cost
cost1<- mu.c + delta[2]*effect1
cost2<-mu.c + delta[1] + delta[2]*effect2

diff.c<- cost2 - cost1

#ICER
ICER<- diff.c/diff.e

#INB lamda=20000
INB<- diff.e*20000- diff.c
}

# Starting values
list(log.shape.c=0, mu.e=1, mu.c=2, log.shape.a=0,
beta =c(0.1,0),
delta=c(0.1,0))

list(log.shape.c=0.5, mu.e=2, mu.c=0.2, log.shape.a=0.2,
beta =c(-1,1),
delta=c(1,1))
```

Model 4: gamma distribution on costs and gamma on effects

```
model{
for (i in 1:N){
        efic.minus[i]<-1-efic[i]
        efic.minus[i] ~ dgamma(shape.e, rate.e[i])
        rate.e[i] <- shape.e/phi.e[i]

        cost[i]~dgamma(shape.c, rate.c[i])
        rate.c[i]<-shape.c/phi.c[i]

        phi.e[i] <- mu.e + beta[1]*(x1[i]) + beta[2]*(x2[i]-mean(x2[]))
        phi.c[i] <- mu.c + delta[1]*(x1[i]) + delta[2]*(efic[i]-phi.e[i])
        }


#covariate prior distributions
for (m in 1:2){
beta[m]~dnorm(0, 0.01);
delta[m]~dnorm(0, 0.01);
}

#prior distributions
log(shape.c)<-log.shape.c
log.shape.c~dnorm(0,0.01)

log(shape.e)<-log.shape.e
log.shape.e ~ dnorm(0, 0.01)

mu.e~dnorm(0,0.01)
mu.c~dnorm(0,0.01)

#Incremental QALY
effect1<-mu.e
effect2<-mu.e + beta[1]
diff.e <- -(effect2 - effect1)

#Incremental Cost
cost1 <- mu.c +delta[2]*effect1
cost2 <- mu.c +delta[1] +delta[2]*effect2
diff.c <- cost2 - cost1

#ICER
ICER<-diff.c/diff.e

#INB lamda=20000
INB <- diff.e*20000-diff.c

}

# Starting values
list(log.shape.c=0.6, log.shape.e=0.6, mu.e=0.2, mu.c=2.2,
beta =c(0.15,0.15),
delta=c(0.15,0.15))

list(log.shape.c=0.5, log.shape.e=0.5, mu.e=0.1, mu.c=2,
beta=c(0.1,0.1),
delta=c(0.1,0.1))
```

Model 5: log-normal distribution on costs and beta on effects

```
model{
for (i in 1:N){
        efic[i] ~ dbeta(shape.a, shape.b[i])
        shape.b[i] <- (shape.a*(1-phi.e[i]))/phi.e[i]

        cost[i]~dlnorm(phi.c[i], tau)

        logit(phi.e[i]) <- mu.e + beta[1]*x1[i] + beta[2]*(x2[i]-mean(x2[]))
        phi.c[i] <- mu.c + delta[1]*x1[i]  + delta[2]*(efic[i]-phi.e[i])
                }

#covariate prior distributions
beta[1]~dnorm(0,0.01)
beta[2]~dnorm(0,0.01)
delta[1]~dnorm(0,0.01)
delta[2]~dnorm(0,0.01)


#prior distributions
tau~dunif(0, 8)

mu.e~dnorm(0, 0.01)
mu.c~dnorm(0,0.01)

log(shape.a)<-log.shape.a
log.shape.a~dnorm(0, 0.01)

#Incremental QALY
logit(effect1) <- mu.e
logit(effect2) <- mu.e + beta[1]
diff.e<- effect2 - effect1

#Incremental Cost
log(cost1)<-mu.c + delta[2]*effect1
log(cost2)<- mu.c + delta[1] + delta[2]*effect2
diff.c<- cost2 - cost1

#ICER
ICER<- diff.c/diff.e

#INB lamda=20000
INB<- diff.e*20000- diff.c
}

# Starting values.
list(tau=0.1, mu.e=5, mu.c=2, log.shape.a=0.1,
beta =c(0.1,0.1),
delta=c(0.1,0.1))

list(tau=0.5, mu.e=-1, mu.c=0.2, log.shape.a=0.5,
beta =c(-1,1),
delta=c(1,-1))
```