missingHE: Health Economic Evaluations with Missing Data

Andrea Gabrio
University College London
Department of Statistical Science
Gower Street, London, WC1E 6BT (UK)
E-mail: ucakgab@ucl.ac.uk

Abstract. This is my abstract.

1 Introduction

A well-known issue in health economic evaluations based on individual-patient data, especially within a randomised controlled trial (RCT) setting, is the presence of missing data in the outcome variables (Manca and Palmer, 2006, Faria et al. (2014)). More specifically, patients may be lost to follow-up, questionnaires may be lost or unreturned and responses to individual questionnaire items may be illegible, nonsensical or nonexistent (National Research Council, 2010).

Removing the unobserved cases (a method usually referred to as Complete Case Analysis, CCA) or replacing a missing observation with a single predicted value (Single Imputation, SI) generally leads to a loss in efficiency and possible serious biases in the parameter estimates (Rubin, 1987, Schafer (1997),Little et al. (2010),Molenberghs and Kenward (2007)). Nevertheless, many reviews concluded that most applied within-trial Cost Effectiveness Analyses (CEAs) widely use these approaches or are unclear on the methodology used (Wood et al., 2004, Noble et al. (2012),Gabrio et al. (2017)). As a result, it can be difficult to understand the assumptions underpinning these analyses as well as to use their findings in subsequent research or in resource allocation decisions.

A more flexible missing data method is Multiple Imputation (MI, (Rubin, 1987)), which has become extremely popular in clinical studies (Carpenter and Kenward, 2013). In a nutshell, MI proceeds by replacing each missing data point with a value simulated from a suitable model M complete (i.e. without missing data) replicates of the original dataset are thus created, each of which is then analysed separately using standard methods. The individual estimates are pooled together using meta-analytic tools such as *Rubin's rules*, to reflect the inherent uncertainty in imputing the missing values. As a consequence of the separation between the imputation and the analysis steps, MI requires the correspondence between the model used to impute the missing values and the one used to obtain the estimates of interest, a feature often called *congeniality*. Non-congeniality can occur if the imputation model is specified as more restrictive than the analysis model, possibly causing biased and inefficient estimates (Van Buuren, 2012).

Additionally, in many applications, MI is based upon a *Missing At Random* (MAR) mechanism, which implies the assumption that the observed data can explain fully the reason for why some observations are missing. However, this may not be reasonable in practice (e.g. for self-reported questionnaire data) and it is important to explore whether the resulting inferences are robust to a range of plausible *Missing Not At Random* (MNAR) mechanisms, which cannot be explained fully by observed data only. The validity of neither of these mechanisms can be tested using the available data alone and thus it is crucial to perform *Sensitivity Analysis* (SA) to explore how variations in assumptions about the missing values impact the results (Carpenter et al., 2007).

The problem associated with the missing values in CEA is often coupled with the fact that resource use and health related quality of life data, e.g. Quality-Adjusted Life Years (QALYs), are generally affected by different types of idiosyncrasies (e.g. skewness, spikes at the boundary) that may bias estimates from standard regresison models (Rascati et al., 2001, Thompson and Nixon (2005),Basu and Manca (2012)). Typical simplifying assumptions often (implicitly) made in routine analyses are: normality for the underlying cost and effectiveness data, independence between the outcomes and failure to adjust for some potentially relevant baseline variables (Manca et al., 2005, van Asselt et al. (2009),Vazquez Polo et al. (2005),O'Hagan and Stevens (2001)). Different methods have been proposed to handle each of these issues, such as the use of bootstrapping or alternative parametric models (Rascati et al., 2001). However, when a combination of these issues affects the data, the building of a more complex model that accounts for all of them simultaneously is desirable.

A full Bayesian modelling framework provides a unified framework that allows to jointly tackle the different types of issues discussed above, which in turns produces several advantages in comparison to a frequentist counterpart, specifically in health care technology assessments (Spiegelhalter et al., 2004, Baio (2013)). Among these, the Bayesian approach naturally allows for the principled incorporation of external evidence (e.g. expert opinions) through the use of prior distributions. This is often crucial for conducting sensitivity analysis to a plausible range of missingness assumptions including MNAR (Daniels and Hogan, 2008), particularly when the evidence produced by the current study is limited, as is the case of small pilot trials, whose objective is to aid decision making about larger investments.

Moreover, we note that MI can be considered as an approximation to a full Bayesian analysis in the sense that it separates the imputation and analysis steps in two estimation procedures. Conversely, within a full Bayesian approach, the parameters of interest are estimated simultaneously with the imputation of the missing values and no additional analysis or *ad hoc* pooling is necessary. Even though it has been shown that MI performs well in most standard situations, when the complexity of the analysis grows, a full Bayesian approach is likely to be a preferable option that naturally allows to propagate uncertainty to the wider economic model and perform sensitivity analysis.

Interestingly, the often-quoted objection to Bayesian modelling, i.e. that it is too computationally intensive in comparison to simpler frequentist counterparts, is likely to dissolve in the presence of extremely complex models, which would require tailor-made routines for the optimisation of non-standard multivariate likelihood functions, thus effectively surrendering their computational advantage over intensive but efficient sampling methods such as Markov Chain Monte Carlo (MCMC, (Brooks et al., 2011)).

The objective of this work is to develop a suite of functions and tools for the freely available statistical software `R`, specifically designed to provide a unified full Bayesian framework that allows to explore alternative plausible missingness assumption scenarios, while simultaneously account for different potential bias sources in CEAs.

2 The R package of missingHE

missingHE is a package designed to aid in the process of economic evaluations and cost-effectiveness analysis in Health Economics in the presence of missing data in the outcome variables. The modelling perspective used is that of the Bayesian approach, exploiting its natural suitability to assess the intrinsic uncertainty of the missing data and the uncertainty underpinning decision-making problems. In fact, **missingHE** can be considered a wrapper for some other `R` packages. The first package, **R2jags** (YS. and Yajima, 2015), allows to interface `R` with `JAGS` (Plummer, 2010), a program for simulation from Bayesian hierarchical models using MCMC methods that is based on the `BUGS` modelling language (Gelman et al., 2017). The second, **BCEA** (Baio et al., 2016), is used to produce an economic evaluation output from the posterior inference generated via `JAGS`. The package also relies on other packages such as **ggplot2** (Wickham and Chang, 2016), **gridExtra** (Auguie and Antonov, 2016), **ggthemes** (Arnold et al., 2017), **mcmcplots** (Curtis et al., 2015) and **ggmcmc** (Marin, 2016), mainly for graphics purposes.

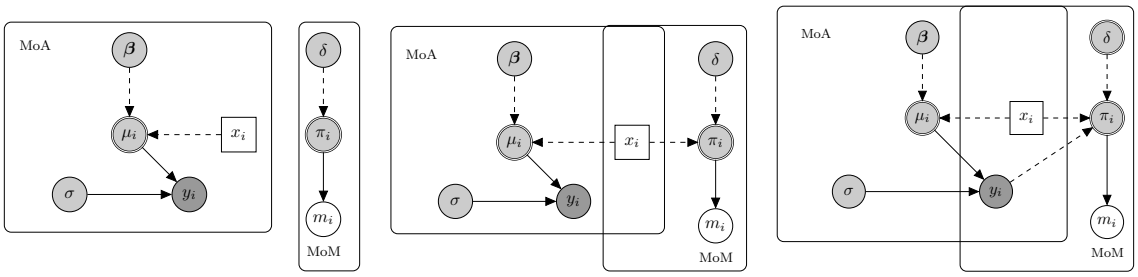
3 Missing Data Mechanism

When analysing partially observed data, it is essential to investigate the possible reasons behind the missingness. This formally translates into an *assumed* missing data mechanism (Rubin, 1987) that is linked to the data generating process, as a key concept to address missingness in a "principled" way. We specifically refer to "principled" methods for missing data as those based on a well-defined statistical model for the complete data, and explicit assumptions about the missing value mechanism.

We consider a sample of $i = 1, \ldots, n$ individuals and for each the relevant outcome is indicated as y_i, which is unobserved for some individuals. Typically, trial data also include a set of J covariates $\boldsymbol{x}_i = (x_{1i}, \ldots, x_{Ji})$, e.g. sex, age or co-morbidities. While in general these may be partially or fully observed, in this section we consider only the latter case. In addition, we define a missingness indicator m_i taking value 1 if the $i-$th subject is associated with missing outcome and 0 otherwise.

This setting can be modelled using two sub-models, or "modules". The first module is the missing data mechanism, denoted as *Model of Missingness* (MoM). It describes a probability distribution for m_i, as a function of some unobserved parameters π_i and δ, defining the probability of missingness in the outcome variable y_i. The second module is the data generating process of the outcome variable, denoted as *Model of Analysis* (MoA). This contains the main parameters of interest (e.g. the population average costs and effectiveness) and describes a probability model for the outcome y_i. As a general example, we can think of a simple regression model where $y_i \sim \text{Normal}(\mu_i, \sigma)$, and $\mu_i = \beta_0 + \beta_1 x_i$. In this case, the parameters of the MoA are the regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1)$ representing respectively the intercept and the slope, and the individual standard deviation σ.

The most accepted classification of missingness mechanisms is based on three classes (Rubin, 1987), which are distinguished according to the way the probability of missingness in the MoM is modelled. A simple graphical representation of the three classes is provided in Figure 1.



(a) Missing Completely At Random (MCAR)

(b) Missing At Random (MAR)

(c) Missing Not At Random (MNAR)

Figure 1: Graphical representation of Rubin's missing data mechanism classes, namely (a) MCAR, (b) MAR and (c) MNAR. Variables and parameters are represented through nodes of different shapes and colours. Parameters are indicated by grey circles with logical parameters defined by double circles, while predictor variables are assumed fixed and drawn as white squares. Fully observed variables are denoted by white circles, partially observed variables by darker grey circles. Nodes are related to each other through dashed and solid arrows which respectively represent logical functions and stochastic dependence. MoA=Model of Analysis, MoM=Model of Missingness.

Variables and parameters are denoted by nodes of different shapes and colours according to their nature. Parameters (β_0, β_1, σ, δ) are represented through grey circles. "Logical" quantities such as μ_i and π_i, which are defined as a function of other parameters, are denoted by a double circle notation. Fully observed variables (m_i) are represented with a white circle while partially observed variables (y_i) are denoted by a darker grey circle. Finally, we show covariates (x_i) as white squares to indicate that they are fully observed and not modelled. Rounded rectangles are used to show the extent of the two modules in terms of variables/parameters included. Arrows show the relationships between the nodes, with dashed and solid lines indicating logical functions and stochastic dependence, respectively.

The missing data mechanism specifies a probability model for the distribution of m_i conditional on all other variables, broadly distinguished, according to Rubin's classification, into three classes.

1. Figure 1 (a) illustrates the class of *Missing Completely At Random* (MCAR), in which the probability of missingness is fully independent of any other partially or fully observed variable. Consequently, the MoA and MoM are not connected and π_i does not depend on any quantity in the MoA. This amounts to assuming that there is no systematic difference between partially and fully observed individuals in terms of the outcome y_i. In other words, in this case we would be assuming that observed cases are a representative sample of the full sample.
2. Figure 1 (b) shows a case of *Missing At Random* (MAR), in which the missingness probability may depend on a fully observed variable. As a result, MoA and MoM are connected by means of the predictor variable affecting both the mechanisms generating y_i and m_i. Because of this relationship, the partially observed cases are systematically different from the fully observed cases; crucially, however, the difference is fully captured by x_i.
3. Figure 1 (c) provides an example of *Missing Not At Random* (MNAR). This is characterised by dependence of the probability of missingness on both the partially and fully observed variables. Thus, π_i depends on both the fully observed predictor x_i and the partially observed outcome y_i. This means that the difference between fully and partially observed cases still depends on the missing values, even after taking x_i into account. Therefore it is necessary to make more structured assumptions about this relationship that go beyond the information contained in the data.

Regardless of the setting, it is important to notice that it is never possible to definitively distinguish between MAR and MNAR models. The data alone do not provide all the information necessary to make this choice and, at the same time, different MNAR models can provide identical fits to the observed data. However, they may have quite different implications for the unobserved data, leading to different conclusions (Molenberghs et al., 2015). Therefore, it becomes crucial to explore the sensitivity of the results with respect to different missing data assumptions and quantify results' uncertainty. What is generally recommended is to set MAR as the reference assumption and then explore different MNAR departures. However, the base-case analysis should be primarily defined accorging to the available state of knowledge in the given setting.

4 Selection Models

When informative missingness it thought to be the most realistic scenario, then setting-specific MNAR assumptions should be set as the reference case, with suitably-defined departures being explored in *Sensitivity Analysis* (SA). Usually, SA for nonignorable/informative models is implemented through advanced statistical methods, which can explicitly model a MNAR mechanism. In this package we focus on a specific class of these, named *Selection Models* (Molenberghs et al., 2015; Daniels and Hogan, 2008; Mason et al., 2012).

To represent the application of selection models we consider a simple example. We assume a data set comprising a partially observed response variable y, the corresponding missing data indicator vector m, and a fully-observed covariate x. Under the SM approach, the joint distribution $p(y, m)$ is factored as the product of the marginal distribution $p(y)$ and the conditional distribution $p(m \mid y)$.

$$p(y, m \mid x, \boldsymbol{\theta}^{MoA}, \boldsymbol{\theta}^{MoM}) = p(y \mid x, \boldsymbol{\theta}^{MoA}) p(m \mid y, x, \boldsymbol{\theta}^{MoM})$$

where, $\boldsymbol{\theta}^{MoA}, \boldsymbol{\theta}^{MoM}$ are the set of parameters associated with the MoA and the MoM, respectively. We need to specify the complete data model for the response, so that the probability of nonresponse is modelled conditionally on the possibly unobserved outcomes. Model identifiability comes from some parametric assumption and the assumed form of $p(y)$, which wil implicitly set up the relationships between the parameters indexing the distribution of the observed and unobserved cases.

4.1 Model of Analysis

In health economic evaluations, the MoA corresponds to a joint model assumed for the effectiveness and cost data. **missingHE** specifies this model within a general Bayesian framework which allows the user to choose among a set of alternative structures for both outcomes.

Assume that some patient-level data are collected from a trial on $i = 1, \ldots, n$ individuals who are randomly allocated to either a control ($t = 1$) or intervention ($t = 2$) group, with sample sizes n_1 and n_2, respectively. We denote by e_{it} and c_{it} the effectiveness and cost outcome variables for the $i-$th person in group t of the trial. To simplify the notation unless necessary we suppress the treatment subscript t. In order to account for correlation between the outcomes, in general we can specify the joint distribution $p(e, c)$ using the following factorisation (Nixon and Thompson, 2005):

$$p(e, c) = p(c)p(e \mid c) = p(e)p(c \mid e)$$

where, for example, $p(e)$ is the *marginal* distribution of the effectiveness and $p(c \mid e)$ is the *conditional* distribution of the costs given the effectiveness. Note that while it is possible to use interchangeably either factorisation, **missingHE** always assumes a marginal for the effectiveness and a conditional for the costs.

Without loss of generality, we can consider for each individual a marginal distribution $p(e_i \mid \boldsymbol{\theta}_e^{MoA})$, indexed by a set of parameters $\boldsymbol{\theta}_e^{MoA}$, composed by a *location* ϕ_{ie} and a set of *ancillary* parameters $\boldsymbol{\psi}_e$ typically including some measure of *marginal* variance, σ_e^2. We can model the location parameter using a generalised linear structure, e.g.

$$g_e(\phi_{ie}) = \beta_{0e} \; [+ \ldots],$$

where β_{0e} is the intercept and the notation $[+\ldots]$ indicates that other terms (e.g. quantifying the effect of relevant covariates) may or may not be included in the model. In the absence of covariates or assuming that a centered version $x_i^* = (x_i - \bar{x})$ is used, the parameter $\mu_e = g_e^{-1}(\beta_{0e})$ represents the population average effectiveness.

As for the costs, we can consider a model $p(c_i \mid e_i, \boldsymbol{\theta}_c^{MoA})$, which explicitly depends on the effectiveness variable, as well as on a set of quantities $\boldsymbol{\theta}_c^{MoA}$, again comprising of the location and ancillary parameters. Note that in this case $\boldsymbol{\psi}_c$ includes a *conditional* variance τ_c^2, which can be typically expressed as a function of the marginal variance σ_c^2 (Nixon and Thompson, 2005, Baio (2013)). The location can be modelled as a function of the effectiveness variable as a *conditional regression*:

$$g_c(\phi_{ic}) = \beta_{0c} \; [+ \ldots] + \rho(e_i - \mu_e),$$

Here, $(e_i - \mu_e)$ is the centered version of the effectiveness variable, while ρ quantifies the correlation between costs and effectiveness. Assuming other covariates are either also centered or not present at all, $\mu_c = g_c^{-1}(\beta_{0c})$ is the population average cost.

Note that **missingHE** expands any categorical covariates to a set of dummy variables: so if a covariate has four categories, in line with R notation, **missingHE** considers three binary indicators. Thus the profile (0,0,0) indicates the first (reference) category, while the profiles (1,0,0), (0,1,0) and (0,0,1) indicate the second, third and fourth category, respectively. In **missingHE**, the total number of covariates depends on this full expansion of the design matrix.

Figure 2 shows a graphical representation of the general modelling framework described above. The effectiveness and cost distributions are represented in terms of combined "modules" — the blue and the red boxes — in which the random quantities are linked trough logical relationships. This ensures the full characterisation of the uncertainty for each variable in the model. Notably, this is general enough as to be extended to any suitable distributional assumption, as well as to handle covariates in either or both the modules.

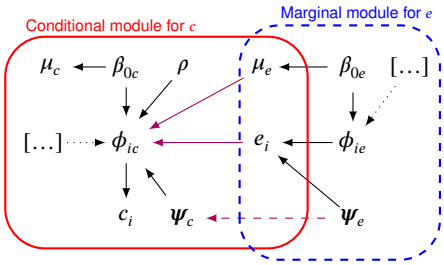


Figure 2: Joint distribution $p(e, c)$, expressed in terms of a marginal distribution for the effectiveness and a conditional distribution for the costs, respectively indicated with a solid red line and a dashed blue line. The parameters indexing the corresponding distributions or "modules" are indicated with different Greek letters, while i denotes the individual index. The solid black and magenta arrows show the dependence relationships between the parameters within and between the two models, respectively. The dashed magenta arrow indicates that the ancillary parameters of the cost model may be expressed as a function of the corresponding effectiveness parameters. The dots enclosed in the square brackets indicate the potential inclusion of other covariates at the mean level for both modules.

Table 1 shows the different types of models for both the effectiveness and costs that are implemented in **missingHE**. More distribution choices will be available in the next versions. In each, by default, minimally informative priors are specified on all the relevant parameters.

MoA	Marginal Mean	Ancillary	Default Priors
Effectiveness			
$e_i \sim \text{Normal}(\phi_{ie}, \sigma_e^2)$	$\mu_e = \beta_{0e}$	σ_e	$\boldsymbol{\beta}_e \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\log \sigma_e \sim \text{Uniform}(-5, 10)$
$e_i \sim \text{Beta}(\phi_{ie}\tau_{ie}, (1-\phi_{ie})\tau_{ie}))$	$\mu_e = \frac{\exp(\beta_{0e})}{1+\exp(\beta_{0e})}$	$\tau_{ie} = \frac{\phi_{ie}(1-\phi_{ie})}{\sigma_e^2} - 1$	$\boldsymbol{\beta}_e \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\sigma_e \sim \text{Uniform}(0, \sqrt{\mu_e(1-\mu_e)})$
Cost			
$c_i \sim \text{Normal}(\phi_{ic}, \sigma_c^2)$	$\mu_c = \beta_{0c}$	σ_c	$\boldsymbol{\beta}_c \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\log \sigma_c \sim \text{Uniform}(-5, 10)$
$c_i \sim \text{Gamma}(\phi_{ic}\tau_{ic}, \tau_{ic})$	$\mu_c = \log \beta_{0c}$	$\tau_{ic} = \frac{\phi_{ic}}{\sigma_c^2}$	$\boldsymbol{\beta}_c \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\sigma_c \sim \text{Uniform}(0, 1000)$

Table 1: A list of the distributions supported by **missingHE** for the effectiveness (e_i) and cost (c_i) variables. The set of logistic regression parameters for ϕ_{ie} and ϕ_{ic} are indicated with $\boldsymbol{\beta}_e = (\beta_{0e}, \beta_{1e}, \ldots)$ and $\boldsymbol{\beta}_c = (\beta_{0c}, \beta_{1c}, \ldots)$. Notice that `JAGS` specifies Normal distributions in terms of the precision rather than the variance (precision=1/variance).

For all these distributions, when a joint between e_i and c_i is assumed, the additional parameter ρ indexes the costs model and captures the correlation with the effectiveness. The default prior for this parameter is $\rho \sim \text{Normal}(0, 1000)$.

4.2 Model of Missingness

The existence of a bivariate outcome in economic evaluations requires the specification of two different MoM when missingness affects both the effectiveness and cost variables. Under a selection model, the MoMs for the missingness indicators $p(m_{ie} \mid \boldsymbol{\theta}_e^{MoM})$ and $p(m_{ic} \mid \boldsymbol{\theta}_c^{MoM})$ are indexed by the two different sets of parameters $\boldsymbol{\theta}_e^{MoM}$ and $\boldsymbol{\theta}_c^{MoM}$. These include the missingness probabilities in the effectiveness (π_{ie}) and costs (π_{ic}), respectively, which in turn are typically expressed as functions of some other parameters.

More specifically, **missingHE** models the two missing value probabilities as using a general logistic regression form:

$$\text{logit}(\pi_{ie}) = \gamma_{0e} + \delta_e e_i \ [+\ldots] \quad \text{and} \quad \text{logit}(\pi_{ic}) = \gamma_{0c} + \delta_c c_i \ [+\ldots]$$

where γ_{0e} and γ_{0c} are the baseline parameters, while δ_e and δ_c are the parameters that capture the impact of the unobserved values on the missingness probabilities for the effectiveness and costs on the logit scale, respectively. The possible inclusion of other centered covariates in either or both the MoMs is indicated by the terms $[+\ldots]$.

An advantage of using selection models is that the MoM can be neatly fitted inside the Rubin's categories according to the types of variables included in the model for the missingness probabilities. In particular, the inclusion of the parameters δ_e and δ_c denotes the existence of a MNAR mechanism, as they are not well-identified from the data; when instead these parameters are set to zero, the mechanism is MAR (which becomes MCAR if no covariate is included in the models).

Table 2 shows how **missingHE** specifies the MoM in the effectiveness and costs under either a MAR or MNAR assumption. By default, minimally informative priors are used on all parameters.

MoM	Mechanism	Marginal Missingness Probability	Default Priors
Effectiveness			
$m_{ie} \sim \text{Bernoullli}(\pi_{ie})$	MAR	$\bar{\pi_e} = \frac{\exp(\gamma_{0e})}{1+\exp(\gamma_{0e})}$	$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000)$
	MNAR	$\bar{\pi_e} = \frac{\exp(\gamma_{0e}+\delta_e \bar{e}_i)}{1+\exp(\gamma_{0e}+\delta_e \bar{e}_i)}$	$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000)$; $\delta_e \sim \text{Normal}(0, 1)$
Cost			
$m_{ic} \sim \text{Bernoullli}(\pi_{ic})$	MAR	$\bar{\pi_c} = \frac{\exp(\gamma_{0c})}{1+\exp(\gamma_{0c})}$	$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000)$
	MNAR	$\bar{\pi_c} = \frac{\exp(\gamma_{0c}+\delta_c \bar{c}_i)}{1+\exp(\gamma_{0c}+\delta_c \bar{c}_i)}$	$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000)$; $\delta_c \sim \text{Normal}(0, 1)$

Table 2: The distributions used by **missingHE** for the missing indicators for the effectiveness (m_{ie}) and cost (m_{ic}) variables. The set of logistic regression parameters for π_{ie} and π_{ic} are indicated with $\boldsymbol{\gamma}_e = (\delta_e, \gamma_{0e}, \gamma_{1e}, \ldots)$ and $\boldsymbol{\gamma}_c = (\delta_c, \gamma_{0c}, \gamma_{1c}, \ldots)$. Notice that `JAGS` specifies Normal distributions in terms of the precision rather than the variance (precision=1/variance).

Selection models allow to directly model the target distribution of the full data (observed and missing) under MNAR. This has the advantage to straightforwardly formulate assumptions about the nonresponse mechanism. The drawback is how we can translate these assumptions into assumptions on the distribution of the missing data.

Indeed, model identification depends on assumptions on the distribution of the MoA (often difficult to check) and on the form of the MoM (on which unverifiable assumptions have to be made). SA is an important tool that should always be used in order to handle missingness uncertainty. Specifically, two different types of sensitivity analysis can be used (Mason et al., 2012):

1. **Assumption Sensitivity** (AS) varies the distributional assumptions in the MoA for e_i and c_i.

2. **Parameter Sensitivity** (PS) varies the prior distributions for δ_e and δ_c in the MoM. Priors must be chosen according to the available information (e.g. expert opinions) and define a set of plausible missingness assumptions to explore.

After all required SAs have been performed, it is important to examine the results to establish how much the quantities of interest vary. In the event conclusions are not robust we may need to gather more information to better specify the model. One of the reasons is that, while a model fit to observed data can always be assessed, its fit to unobserved data can never be assessed and therefore compared.

Finally, it is important to stress that any conclusion derived from a nonignorable missing data model must be treated very cautiously. Each model specification makes assumptions about the behaviour of missing data that can only be formulated using external information, should such information exist. The key fact is that we cannot ultimately distinguish between MAR and MNAR. Hence, SA is required to the test the robustness of results to different plausible assumptions.

5 Hurdle Models

A common type of idiosyncrasy in economic evaluations relates to data showing spikes at one or both of the boundaries of the range for the underlying distribution. For example, some patients in a trial may not accrue any cost at all (i.e. $c_i = 0$), thus invalidating the assumptions for the Gamma distribution, which is defined on the range $(0, +\infty)$. Similarly, we may observe individuals who are associated with perfect health, i.e. unit QALY (Basu and Manca, 2012), which makes it difficult to use a Beta distribution, defined on the open interval $(0, 1)$.

A simple solution is to add/subtract a small constant ϵ to the entire set of observed values for the cost/effectiveness variable, thus artificially re-scaling it in the desired interval (Cooper et al., 2003). Despite being very easy to implement, this strategy is potentially problematic as the results are likely to be strongly affected by the actual choice of the scaling parameter ϵ and no clear guideline exists about the value to use (e.g. $0.1, 0.01, \ldots$). In addition, when the proportion of these values is substantial, they may induce high skewness in the data and the application of simple methods may lead to biased inferences (Mihaylova et al., 2011). A more efficient solution suggested to handle this issue is the application of *hurdle models* (Ntzoufras, 2009, Baio (2014)).

These are mixture models defined by two components: the first one is a mass distribution at the spike, while the second is a parametric model applied to the natural range of the relevant variable. Usually, a logistic regression is used to estimate the probability of incurring a *structural* value (e.g. 0 for the costs, or 1 for the QALYs); this is then used to weigh the mean of the *non-structural* values estimated in the second component. Hurdle models have been discussed and applied in CEA mainly for handling structural zero costs (Tooze et al., 2002, Harkanen et al. (2013),Baio (2014)).

Within the **missingHE** general framework for the MoA shown in Figure 2, we can extend all the models described in Table 1 to a hurdle version for both outcome variables. Specifically, for each subject in the trial $i = 1, \ldots, n$ we define two indicator variables d_{ie} and d_{ic} taking value 1 if the $i-$th individual is associated with a structural value in the effectiveness ($e_i = se$) and costs ($c_i = sc$), respectively, and 0 otherwise. This is then modelled as

$$d_{ie} := \mathbb{I}(e_i = 1) \sim \text{Bernoulli}(\pi_{ie}) \quad \text{and} \quad d_{ic} := \mathbb{I}(c_i = 1) \sim \text{Bernoulli}(\pi_{ic})$$
$$\text{logit}(\pi_{ie}) = \gamma_{0e} \, [+ \ldots] \quad \text{and} \quad \text{logit}(\pi_{ic}) = \gamma_{0c} \, [+ \ldots],$$

where π_{ie} and π_{ic} are the individual probabilities of a structural value in the effectiveness and costs, which are estimated on the logit scale as a function of some baseline parameters γ_{0e} and γ_{0c}. Other centered covariates ($[+ \ldots]$) can be additively included in the model of d_{ie} and d_{ic}.Within this framework, the quantities

$$\bar{\pi}_e = \frac{\exp(\gamma_{0e})}{1+\exp(\gamma_{0e})} \quad \text{and} \quad \bar{\pi}_c = \frac{\exp(\gamma_{0c})}{1+\exp(\gamma_{0c})}$$

represents the estimated marginal probability of structural e_i and c_i. The parameters $\bar{\pi}_e$ and $\bar{\pi}_c$ in effect represent the weights used to mix the two components.

To make a parallel with the missing data literature, when no covariates are included in the models of π_{ie} and π_{ic}, these probabilities are randomly predicted and describe what we call a *Structural Completely At Random* (SCAR) mechanism, in which the chance of observing an individual associated with structural value does not depend on any other variable. When, instead, some relevant covariates are inlcuded and they contribute to estimate the probabilities, we have a *Structural At Random* (SAR) mechanism.

Table 3 shows how **missingHE** specifies the structural value mechanism in the effectiveness and costs under either a SCAR or SAR assumption. By default, minimally informative priors are used on all parameters.

Model	Mechanism	Marginal Structural Value Probability	Default Priors
Effectiveness			
$d_{ie} \sim \text{Bernoullli}(\pi_{ie})$	SCAR	$\bar{\pi}_e = \frac{\exp(\gamma_{0e})}{1+\exp(\gamma_{0e})}$	$\gamma_{0e} \sim \text{Normal}(0, 1000)$
	SAR		$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000)$
Cost			
$d_{ic} \sim \text{Bernoullli}(\pi_{ic})$	SCAR	$\bar{\pi}_c = \frac{\exp(\gamma_{0c})}{1+\exp(\gamma_{0c})}$	$\boldsymbol{\gamma}_{0c} \sim \text{Normal}(0, 1000)$
	SAR		$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000)$

Table 3: The distributions used by **missingHE** for the structural value indicators for the effectiveness (d_{ie}) and cost (d_{ic}) variables. The set of logistic regression parameters for π_{ie} and π_{ic} are indicated with $\boldsymbol{\gamma}_e = (\gamma_{0e}, \gamma_{1e}, \ldots)$ and $\boldsymbol{\gamma}_c = (\gamma_{0c}, \gamma_{1c}, \ldots)$. Notice that `JAGS` specifies Normal distributions in terms of the precision rather than the variance (precision=1/variance).

To notice that the correspondence between missingness and structural value mechanisms is not perfect. In the missing data, the assumption of MCAR implies that the MoM and MoA modules are completely separated (and thus the chance of observing a missing value is assumed to be independent on any other variable, including the outcome). This implies that under MCAR there is no need to include the MoM in the analysis. On the other hand, in the structural values the two modules are always linked, because the distribution of the outcome depends on the structural indicator. Nevertheless, under SCAR, we are assuming the absence of other (observed or unobserved) factors that can influence the chance of observing a structural value.

Finally, depending on the value of d_{ie} and d_{ic}, we can partition the observed data on the outcome in two subsets.

1. In the first one (*structural*), defined in correspondence of the n^{se} subjects for whom $d_{ie} = se$ the n^{sc} subjects for whom $d_{ic} = sc$, we define two variables $e_i^{se} = se$ and $e_i^{sc} = sc$.
2. The second one (*non-structural*) is made by the $n^{-se} = (n - n^{se})$ subjects for whom $d_{ie} = 0$ and the $n^{-sc} = (n - n^{sc})$ subjects for whom $d_{ic} = 0$. For these individuals we define two variables e_i^{-se} and c_i^{-sc}.

We can model the non-structural component using a distribution characterised by overall means μ_e^{-se} and μ_c^{-sc}. For example, among the distributions available in **missingHE**, when we observe the structural values $se = 1$ and $sc = 0$, we can directly apply the Beta for the effectiveness and the Gamma for the costs. The

parameterisation and priors assigned are exactly the same of those in Table 1, except the fact that they now apply to the non-structural components only, i.e. e_i^{-se} and c_i^{-sc}.

Finally, using the estimated value for $\bar{\pi}_e$ and $\bar{\pi}_c$, we can compute the overall population average effectiveness and cost measures in both treatment groups μ_{et} and μ_{ct} as the linear combinations

$$\mu_{et} = (1 - \bar{\pi}_{et})\mu_{et}^{-se} + \bar{\pi}_{et} se \quad \text{and} \quad \mu_{ct} = (1 - \bar{\pi}_{ct})\mu_{ct}^{-sc} + \bar{\pi}_{ct} sc$$

5.1 Hurdle models and missingness

Hurdle models represent an interesting approach to handle missingness. More specifically, if outcome data are unobserved then also the structural value indicators d_{ie} and d_{ic} cannot be computed. However, within a Bayesian framework where the uncertainty is propagated through each variable, the indicator variables and the probabilities π_{ie} and π_{ic} are imputed based on the available information in the model. This information may come either from the observed data or from informative priors and may be associated with different types of MoM structures.

Under MAR, no information other than that contained in the observed data is used to impute the missing values, both in the structural and non-structural components. Under MNAR, Hurdle models offer a convenient framework to explore the robustness of the results to some pluasible scenarios and therefore allow to perform a simple type of sensitivity analysis to the missingness assumptions. More specifically, it is possible to arbitrarily set the unobserved values in d_{ie} and d_{ic} to either 1 or 0, using different configurations, i.e. vary the number of structural values potentially observed in a given scenario, and assess how they affect the inferences in both treatment groups.

This corresponds to assessing the impact on the inferences of alternative missingness assumptions. Since these configurations are based on assumptions that cannot be verified from the data at hand (but are in fact arbitrarily set by the experimenter), they effectively represent a way to assess the robustness of the results to some MNAR departures.

Even though this approach associates the missing data with specific MNAR values (structural values), it has the advantage of being easy to implement and offers a starting point to investigate MNAR more thoroughly. Specifically, if one of these scenarios is thought to be more realistic, then it can be explored by means of methods that explicitly account for variability in the MNAR values, e.g. Selection Models.

6 Example

We use a running example to show how selection models and hurdle models can be specified in **missingHE**. Suppose that the user has a suitable dataset, perhaps obtained from a trial, in which data for each individual are recorded for the effectiveness and cost variables as well as for an arm indicator specifying whether the individual to whom the data refer belongs in the control or the active treatment arm of the trial. Of course, other variables may be observed, e.g. relevant covariates, such as sex, age or co-morbidity. Both outcome variables can have missing values while no unobserved values should be observed for the covariates as **missingHE** can only deal with missingness in the outcomes.

Assume that the data are available in the `R` workspace as a data-frame (say, `data`) that can be visualised using, for example, histogram plots

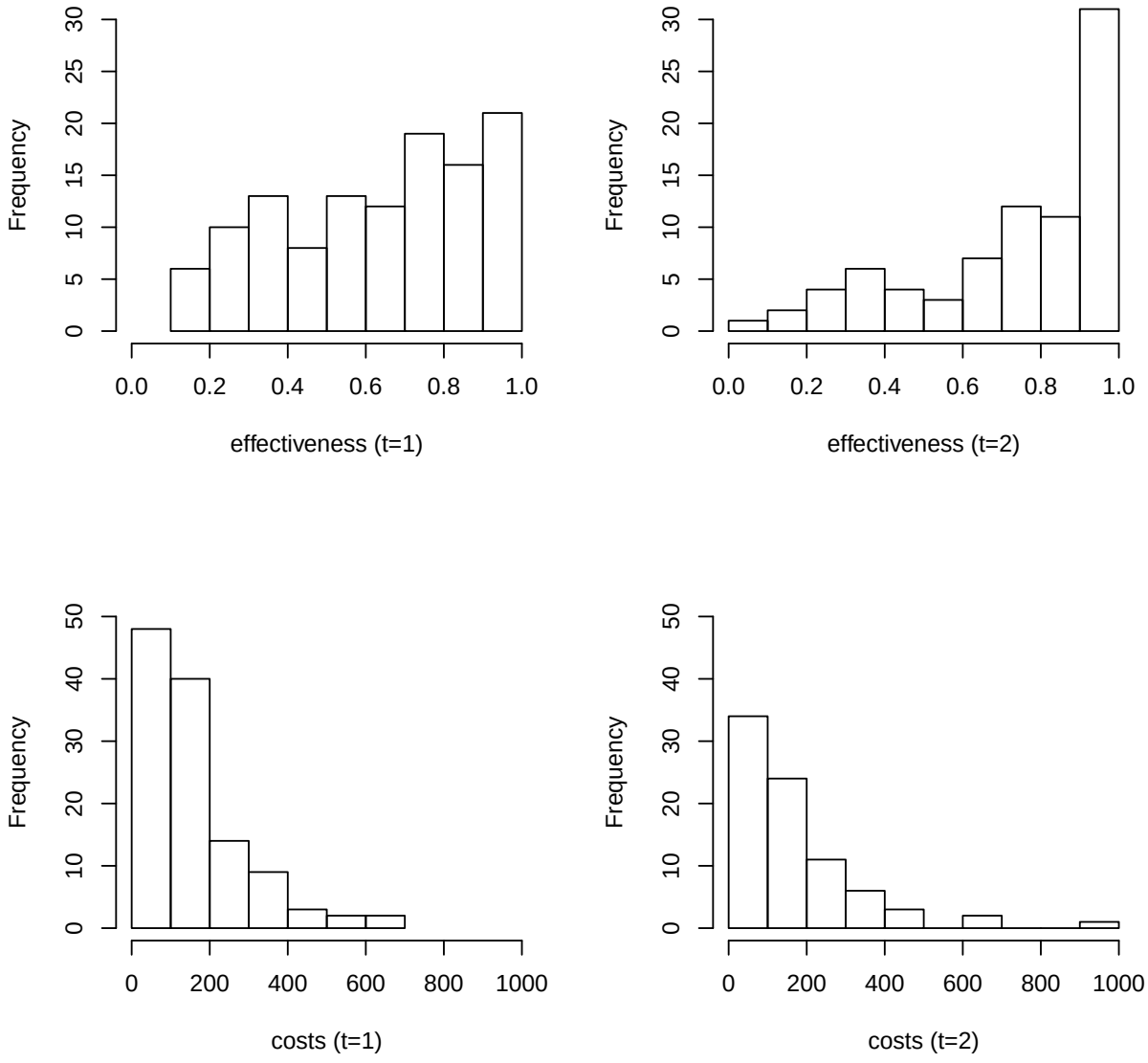


Figure 3: Histograms of the observed data distributions for the effectiveness and costs in the control (t=1) and intervention (t=2) arms

Figure 3 shows the histograms for the observed data distributions. The dataset consists of 250 individuals in total, grouped in two arms (here arm = 1 indicates the control and arm = 2 indicates the active treatment). 51 individuals have unobserved outcome data (32 in the control and 19 in the intervention), and 24 individuals are associated with structural values in both the effectiveness (1) and costs (0).

6.1 Selection models in missingHE

Selection models are implemented in **missingHE** using the Bayesian software program `JAGS` (Plummer, 2010) (Just Another Gibbs Sampler), which is called from the `R` package **R2jags**. The program is based on

the BUGS (Bayesian inference Using Gibbs Sampling) (Lunn et al., 2012) language and performs Bayesian inference using the Gibbs Sampling as a specific type of MCMC algorithm.

To illustrate how **missingHE** interfaces with these programs we assume a joint bivariate Normal MoA distribution to the effectiveness and cost variables in the dataset `data`. With respect to the MoM, we assume a MNAR structure for the effectiveness while we keep a MAR assumption for the costs in both treatment arms. In **missingHE** we can implement the model using the `selection` function:

```
model.sel<-selection(data = data,dist_e = "norm",dist_c = "norm",
             model.eff = e~1,model.cost = c~e,
             model.me = me~e,model.mc = mc~1,
             type = "MNAR",n.iter = 10000,prior="default")
```

where the different arguments of the function have the following interpretations:

1. `data` must contain the a data to analyse, specified in a data frame format
2. `dist_e` and `dist_c` indicate the assumed effectiveness and cost MoA distributions, specified as character names among a set of pre-defined choices. Available choices are: Normal for both outcomes, Beta for the effectiveness and Gamma for the costs.
3. `model.eff` and `model.cost` are formulas that specify which variables should be included in the effectiveness and cost models as covariates (among those included in `data`). A joint bivariate distribution can be assumed by placing `e` in the formula for the costs. By default both formulas do not contain any covariate (indicated with `1`) and assume independence between the outcomes.
4. `model.me` and `model.mc` are formulas that specify which variables should be included in the effectiveness and cost missingness models as covariates (among those included in `data`). It is possible to specify a MNAR mechanism by placing `e` and `c` in the formulas for the missing effectiveness and cost models, respectively. By default no covariate is included (indicated with `1`), implicitly assuming a MAR mechanism.
5. `type` specifies the type of mechanism to be assumed, either MAR or MNAR, respectively indicated by the character values `MAR` and `MNAR`.
6. `n.iter` specifies the number of iterations in each chain of the MCMC algorithm. Other additional arguments that may be provided are: the burnin period to be discarded (`n.burnin`), the number of the chains (`n.chains`), the thinning interval (`n.thin`), the initialised values for the parameters in each chain (`inits`), the upper and lower bounds of the credible intervals for describing the uncertainty around the imputed values (`prob`) and whether the model text file should be saved in the current working directory (`save_model`).
7. `prior` specifies the prior distribtions to be used for the parameters of the model, by default those shown in Table 1 and Table 2. These priors can be overwritten by the user who may desire to change them. In this case, the new hyperpriors for each parameter in the model can be provided by creating a list object that contains the new values. For example, for the model above, the priors for the parameters β_{0e} and δ_e can be defined as follows:

```
myprior<-list("beta0.prior.e"=c(0,0.00001),"delta.prior.e"=c(1,1))
```

The list object `myprior` can then be supplied to the argument `prior` in the function `selection`. It is necessary that specific character names are used to indicate for which parameter the prior should be overwritten. A list of the character names to be used by type of model and parameter can be accessed using the `help` function on `selection`.

Executing the command above creates an object `model.sel` in the class `missingHE`, in which the results of the economic analysis are stored for the given MoA-MoM specification considered. The usual `R` command

```
names(model.sel)
```

returns the names of the elements in the list

```
## [1] "data_set"     "model_output" "cea"          "type"
```

The objects `data_set`, `model_output` and `cea` are themselves lists that contain different elements related to the data provided, the model results and the economic analysis, respectively. For example, the elements in the first object can be accessed using the standard `R` notation `model.sel$data_set[]` (i.e. using double square brackets) and can be inspected typing the command

```
names(model.sel$data_set[])
```

```
##  [1] "effects"                       "costs"
##  [3] "N in reference arm"            "N in comparator arm"
##  [5] "N observed in reference arm"   "N observed in comparator arm"
##  [7] "N missing in reference arm"    "N missing in comparator arm"
##  [9] "covariates_effects"            "covariates_costs"
## [11] "covariates_missing_effects"    "missing_effects"
## [13] "covariates_missing_costs"      "missing_costs"
```

These are merely the data related to the inputs given to the function `selection`, such as effect and cost data, total number of individuals in each arm, number of observed and unobserved individuals in each arm and covariate data (if included in the model). The other elements of the object `model.sel` are

1. `model_output` is a list storing the output of the `JAGS` model. Depending on the type of model, the results shown in this list can vary as they contain the posterior samples of the parameters of interest based on the MoA-MoM structure assumed. In the list a summary of the posterior estimates of the `JAGS` model is also available, taken directly from the output of the function `jags` in the package **R2jags**.
2. `cea` is another list that stores the output of the economic evaluation based on the mean posterior samples of the marginal mean effectiveness and cost parameters and which is implemented using the function `bcea` in the package **BCEA**. This object can be analysed using tailored functions of **BCEA** to visually represent standard CEA outputs such as the Cost-Effectiveness Plane (CEP) (Black, 1990) and the Cost-Effectiveness Acceptability Curve (CEAC) (Van Hout et al., 1994).
3. `type` is a string variable that specifies the type of missingness mechanism assumed.

Model results can be shown using the `print` function which returns the table related to the posterior estimates of the parameters of the model.

```
print(x=model.sel,value.mis=FALSE)
```

```
##                mean     sd     2.5%    97.5% Rhat n.eff
## beta_c[1]   147.342 12.499  122.444  171.298 1.00 10000
## beta_c[2]   153.967 18.583  116.494  190.454 1.00 10000
## beta_e[1]     0.633  0.028    0.578    0.687 1.00   580
## beta_e[2]     0.737  0.032    0.674    0.800 1.00   700
## delta_e[1]   -0.320  0.937   -2.112    1.510 1.01   320
## delta_e[2]   -0.354  0.950   -2.104    1.602 1.01   150
## deviance   3089.898  6.435 3077.783 3103.635 1.00 10000
## gamma_c[1]   -1.294  0.198   -1.694   -0.913 1.00  7600
## gamma_c[2]   -1.428  0.256   -1.945   -0.953 1.00  4500
## gamma_e[1]   -1.124  0.607   -2.406   -0.037 1.01   330
## gamma_e[2]   -1.209  0.728   -2.783    0.064 1.01   150
## mu_c[1]     147.342 12.499  122.444  171.298 1.00 10000
## mu_c[2]     153.967 18.583  116.494  190.454 1.00 10000
## mu_e[1]       0.633  0.028    0.578    0.687 1.00   580
## mu_e[2]       0.737  0.032    0.674    0.800 1.00   700
## p_c[1]        0.217  0.033    0.155    0.286 1.00  7400
## p_c[2]        0.196  0.040    0.125    0.278 1.00  5100
## p_e[1]        0.214  0.033    0.152    0.282 1.00 10000
```

```
## p_e[2]        0.192  0.040    0.121    0.275 1.00 10000
## rho[1]      -40.108 26.645  -91.386   13.154 1.00 10000
## rho[2]       -1.493 28.633  -57.622   55.009 1.00  1700
## s_c[1]      135.451  8.803  119.558  154.325 1.00 10000
## s_c[2]      171.521 13.556  147.569  201.178 1.00 10000
## s_e[1]        0.260  0.017    0.229    0.297 1.00 10000
## s_e[2]        0.264  0.021    0.226    0.309 1.00  8700
```

The argument `x` must contain an object of class `missingHE`, such as `model.sel`. The optional argument `value.mis` allows to exclude (`FALSE`) or include (`TRUE`) the results associated with the imputed values; by default, these values are omitted from the results displayed.

6.2 Hurdle models in missingHE

To illustrate how it is possible to specify a Hurdle model in **missingHE**, we consider the same dataset `data` and now assume a joint bivariate Beta-Gamma distribution to the effectiveness and cost variables. Individuals associated with a unit effectiveness and zero costs are considered structural values and are handled through a Hurdle appraoch.

With respect to the structural value mechanism, we assume a SCAR structure for both outcomes. In **missingHE** we can implement the model using the `hurdle` function:

```
model.hur<-hurdle(data = data,dist_e = "beta",dist_c = "gamma",
             model.eff = e~1,model.cost = c~e,
             model.se = se~1,model.sc = sc~1,se=1,sc=0,
             type = "SCAR",n.iter = 10000,prior="default",d_e=my.d_e)
```

Some of the arguments have the same interpretation of those in the function `selection`, but there are some exceptions:

1. `model.se` and `model.sc` are formulas that specify which variables should be included in the effectiveness and cost structural value models as covariates (among those included in `data`). By default no covariate is included (indicated with `1`), implicitly assuming a SCAR mechanism.
2. `type` specifies the type of mechanism to be assumed, either SCAR or SAR, respectively indicated by the character values `SCAR` and `SAR`.
3. `se` and `sc` define which values in the effectiveness and cost data should be treated as structural by the model. If structural values are observed only for one outcome it is possible to set either `se=NULL` or `sc=NULL`. In this case, no hurdle model is assumed for that outcome, which is modelled using the distribution specified in `dist_e` or `dist_c`.
4. `prior` specifies the prior distribtions to be used for the parameters of the model, by default those shown in Table 1 and Table 3. These priors can be overwritten in a similar way to that of the `selection` function by the user who may desire to change them. A list of the character names to be used by type of model and parameter can be accessed using the `help` function on `hurdle`.
5. `d_e` and `d_c` (optional) specify the vectors of structural value indicators to be used in the model for the effectiveness and costs. If not provided, **missingHE** internally computes these vectors based on the observed cases (`NA` if the cases are missing). When provided, the argument of `d_e` and `d_c` must be vectors of length equal to the number of individuals in `data`, which take value 1 or 0 to respectively associate each case with the structural or non-structural component in the hurdle model.

For example, we can define a new variable `my.d_e` to specify a scenario where all the 51 missing individuals in `data` are associated with a structural one in e_i. This variable is obtained by first computing the usual indicator based on the observed data and then by setting all its missing values to 1:

```
my.d_e<-ifelse(data$e==1,1,0)
my.d_e[is.na(data$e)==TRUE]<-1
```

Once the new variable `my.d_e` is passed to the argument `d_e` in the function `hurdle`, then **missingHE** automatically uses it as the new indicator variable in the model.

Similarly to the `selection` function, it is possible to access the names of the elements in the list `model.hur` by typing `names(model.hur)`. Each of these elements is itself a list which contains objects with similar interpretations to those of the `selection` function.

Model results can be shown again using the `print` function.

```
print(x=model.hur,value.mis=FALSE)
```

```
##                mean     sd     2.5%    97.5% Rhat n.eff
## beta_c[1]     5.099  0.089    4.920    5.274    1  1400
## beta_c[2]     5.237  0.109    5.025    5.453    1 10000
## beta_e[1]     0.364  0.090    0.189    0.542    1  2300
## beta_e[2]     0.840  0.137    0.574    1.111    1 10000
## deviance   7703.241  5.329 7694.860 7715.430    1 10000
## gamma_c[1]   -1.978  0.279   -2.566   -1.463    1  3000
## gamma_c[2]   -1.915  0.331   -2.614   -1.306    1 10000
## gamma_e[1]   -0.809  0.175   -1.154   -0.472    1 10000
## gamma_e[2]   -0.884  0.218   -1.321   -0.458    1 10000
## mu_c[1]     143.933 13.707  118.398  172.801    1  2600
## mu_c[2]     164.056 19.409  129.157  205.609    1 10000
## mu_e[1]       0.717  0.021    0.674    0.757    1  5500
## mu_e[2]       0.787  0.024    0.736    0.832    1 10000
## p_c[1]        0.125  0.030    0.071    0.188    1  2800
## p_c[2]        0.133  0.037    0.068    0.213    1 10000
## p_e[1]        0.309  0.037    0.240    0.384    1 10000
## p_e[2]        0.294  0.045    0.211    0.387    1 10000
## rho[1]       -0.310  0.230   -0.750    0.145    1  4100
## rho[2]        0.214  0.255   -0.268    0.740    1 10000
## s_c[1]      140.015 14.661  114.680  172.237    1  2000
## s_c[2]      168.567 22.634  130.046  219.506    1 10000
## s_e[1]        0.227  0.012    0.206    0.251    1 10000
## s_e[2]        0.252  0.017    0.221    0.286    1 10000
```

7 Diagnostic Checks

As with any MCMC estimation, it is important to thoroughly assess convergence. The function `diagnostic` in **missingHE** allows to visualise the model output and assess convergence. Different diagnostic tools and plots for the model parameters are taken from the package **ggmcmc** and **mcmcplots** and are displayed using functions from **ggplot2** according to the inputs provided by the user. For simplicity, we consider only the model output generated in `model.hur` for the current example, but similar considerations apply any model generated through the functions `hurdle` or `selection`.

We can visually represent via histograms the posterior samples for the mean effect parameters in the two arms in the following way.

```
check.hur<-diagnostic(x=model.hur,type="histogram",param="mu.e",theme=NULL)
```

which displays the graph in Figure 4

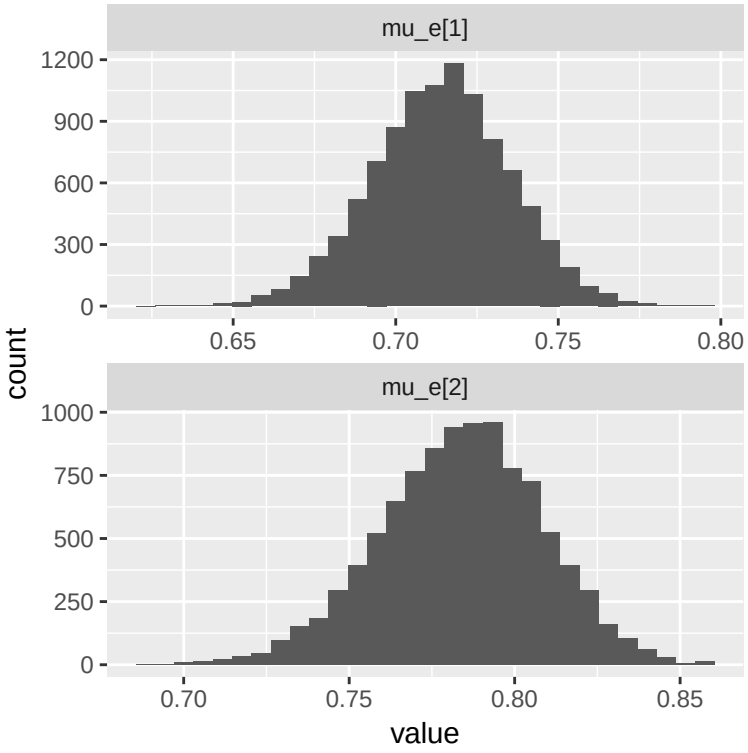


Figure 4: Histograms of the posterior distributions for the mean effectiveness parameters in the two arms

The arguments of `diagnostic` are the followings:

1. `x` is an object of class `missingHE`.
2. `type` specifies which type of diagnostic tools to use for assessing convergence. If `type=summary`, a summary of some of the most important diagnostic plots for the family of parameters indicated in `param` is displayed. A variety of plots are available using specific character names, such as histograms (`histogram`), density plots (`denplot`), traceplots (`traceplot`), autocorrelation plots (`acf`), etc. The full list of all available types of diagnostics can be found using the function `help` on `diagnostic`.
3. `param` specifies for which family of parameters the diagnostic output should be displayed. It must correspond to a character name among a set of pre-defined choices. For example, the mean effectiveness and cost parameters can be accessed via the expressions `mu.e` and `mu.c`, respectively. The type of parameters that are available vary according to the type of model implemented (either Selection or Hurdle models) and the assumptions of the model (MAR/MNAR or SCAR/SAR). The list of all parameters that can be specified with the corresponding character names can be found using the function `help` on `diagnostic`. By default, all model parameters are selected using the character name `all`.
4. `theme` modifies the pre-defnied background theme of the plots generated. Pre-defined themes are taken from the package **ggthemes** and must be indicated with corresponding character names. For a full list of available themes use `help` on `diagnostic`.

It is also possible to combine multiple graphs by running `diagnistic`, setting different parameters to monitor, and saving the plots in corresponding `R` objects. We can then combine these plots into a single one using the function `grid.arrange` from the **gridExtra** package (that should be loaded). For example, we can combine the density and trace plots for the mean effectiveness parameters in `model.hur` in the following way.

```
require(gridExtra)
dens_eff<-diagnostic(x=model.hur,type = "denplot",param = "mu.e")
trace_eff<-diagnostic(x=model.hur,type = "traceplot",param = "mu.e")
grid.arrange(dens_eff$plot, trace_eff$plot, ncol=2)
```

which returns the graphs in Figure 5

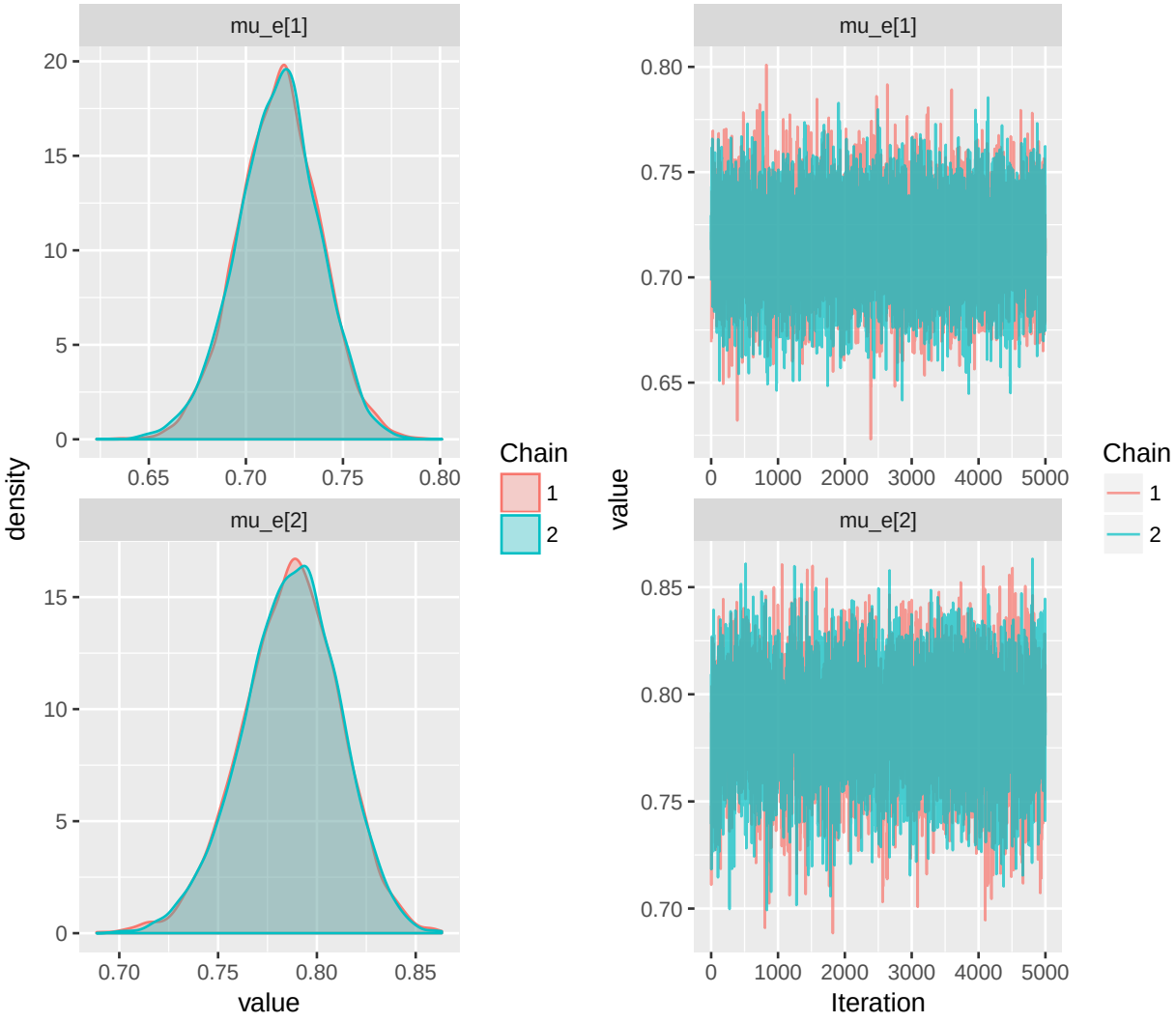


Figure 5: Posterior density and trace plots for the posterior distributions of the mean effectiveness parameters in the two arms

8 Missing Data Plots

Once the model has been estimated, we can visually inspect how missing data in the outcome variables are imputed and compare them to the observed data. **misisngHE** has a specialised function `plot` that can do this, by typing:

```
plot(x=model.hur,class="scatter",outcome="all",theme="base")
```

which displays the graphs in Figure 6

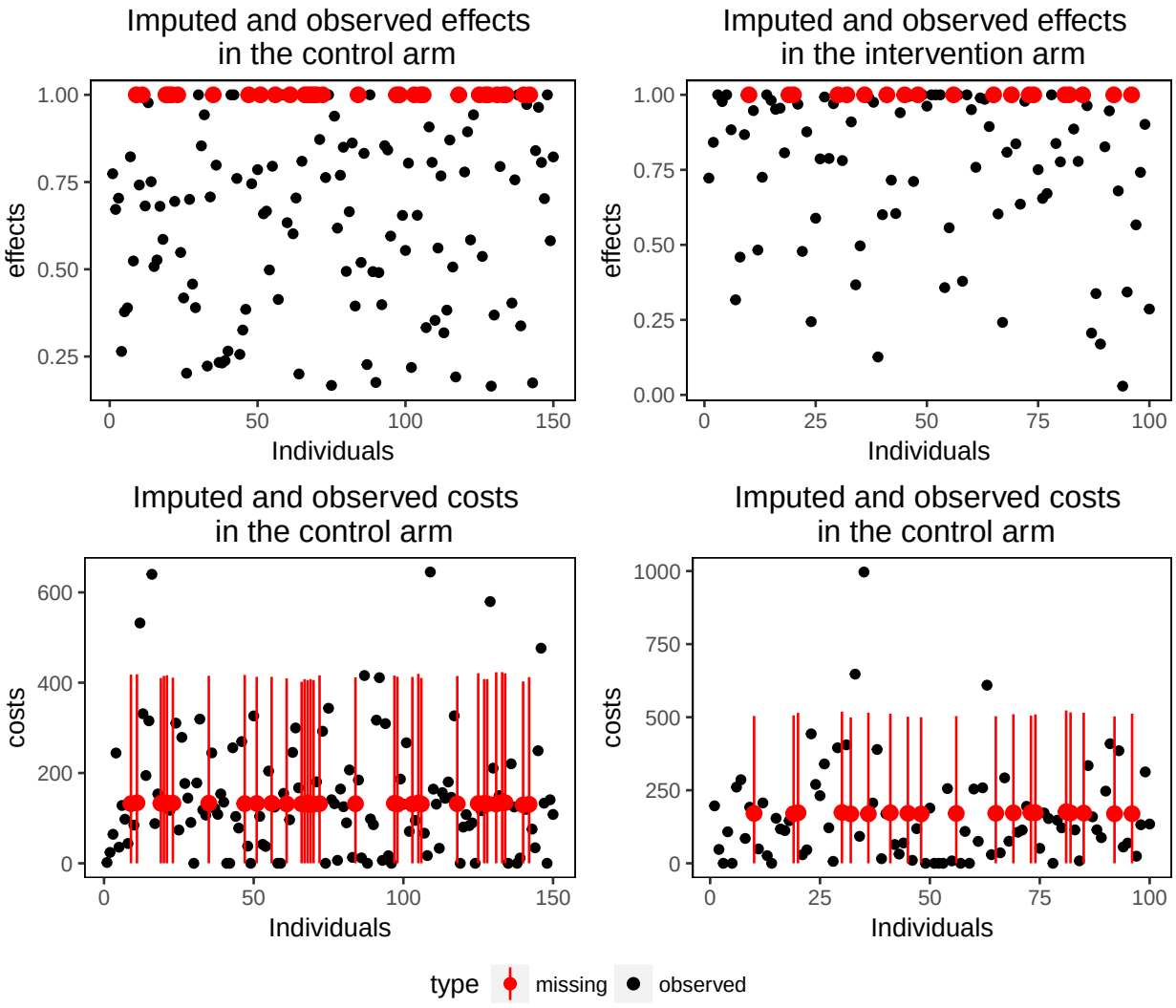


Figure 6: Scatter plots of the observed (black dots) and imputed (red dots and lines) values for both outcomes in the control and intervention arms. Imputations uncertainty is represented in terms of posterior means and 95% credible intervals.

The only compulsory argument to be provided is `x`, which must be a `missingHE` object. All the other optional arguments are mainly related to the type of plot to be shown, which outcome and treatment arm to consider, and other graphics parameters. These are:

1. `class` specifies the type of plot to be displayed. Two alternative character names are available: `scatter` and `histogram`. In the former the observed and imputed values (evaluated at the posterior means) are shown in a scatter plot, with unobserved data also associated with lines representing their posterior credible intervals. By default these are the 95% CI but they can also be modified by changing the values for the upper and lower bounds using the `prob` argument in the function `selection` or `hurdle`. The latter compares the observed and missing value distributions with a histogram plot and associate them with different colours.

2. `outcome` specifies for which variable, either effectiveness, costs or both, and for which treatment arm, either control, intervention or both, results should be visualised. For example, the plots only for the effectiveness in both arms can be shown using the character name `effects`, while the plots by arm can be accessed using the name `arm1` (control) or `arm2` (intervention). Plots for each combination of outcome and treatment group can be specified using the specifc names, such as `effects_arm1`, `costs_arm1`, etc. By default all plots are displayed using the name `all`.
3. `theme` modifies the graphical output according to some pre-specified themes similarly to what shown for the `diagnostic` function.

9 Economic Evaluation

Results from the economic evaluation performed using **missingHE** can be summarised in a tabular form using the function `summary` by typing:

```
summary(object=model.hur)
```

which returns the following table:

```
##
##  Cost-effectiveness analysis summary
##
##  Comparator intervention: intervention 1
##  Reference intervention: intervention 2
##
##  Parameter estimates under SCAR assumption
##
##  Comparator intervention
##                  mean     sd      LB      UB
## mean effects    0.717  0.021   0.681   0.751
## mean costs    143.933 13.707  122.24 167.531
## sd effects      0.227  0.012   0.209   0.247
## sd costs      140.015 14.661 118.118 165.854
##
##  Reference intervention
##                  mean     sd      LB      UB
## mean effects    0.787  0.024   0.745   0.825
## mean costs    164.056 19.409 134.153   197.7
## sd effects      0.252  0.017   0.225    0.28
## sd costs      168.567 22.634 135.442 209.724
##
##  Incremental results
##                   mean     sd      LB     UB
## delta effects     0.07  0.032   0.016  0.123
## delta costs     20.122 23.895 -18.096 60.524
## ICER           287.967
```

The only argument provided to `object` must be a **missingHE** object. Information is reported only for the main parameters of interest in the model for the economic evaluation for both outcomes and treatment groups. In addition, the incremental mean results are provided at the bottom of the table, denoted with `delta effects` and `delta costs` respectively, with also the value of the ICER. Results are summarised in terms of posterior mean, standard deviation and 95% credible intervals for each parameter.

A series of useful functions are included in the package **BCEA** that summarise the results of the economic evaluation coputed by **missingHE**. As an example, the CEP and CEAC plots from the respective functions

`ceac.plot` and `ceplane.plot` in the **BCEA** can be obtained by applying these functions to the `BCEA` object contained in `model.hur` and that can be accessed via `model.hur$cea`. The `R` commands used to generate and combine these plots are the following

```
require(ggplot2)
require(BCEA)
ceac.plot(model.hur$cea,graph = "ggplot2")+ggtitle("CEAC")
ceplane.plot(model.hur$cea,graph = "ggplot2")+ggtitle("CEP")
```

and the resulting output is given in Figure 7

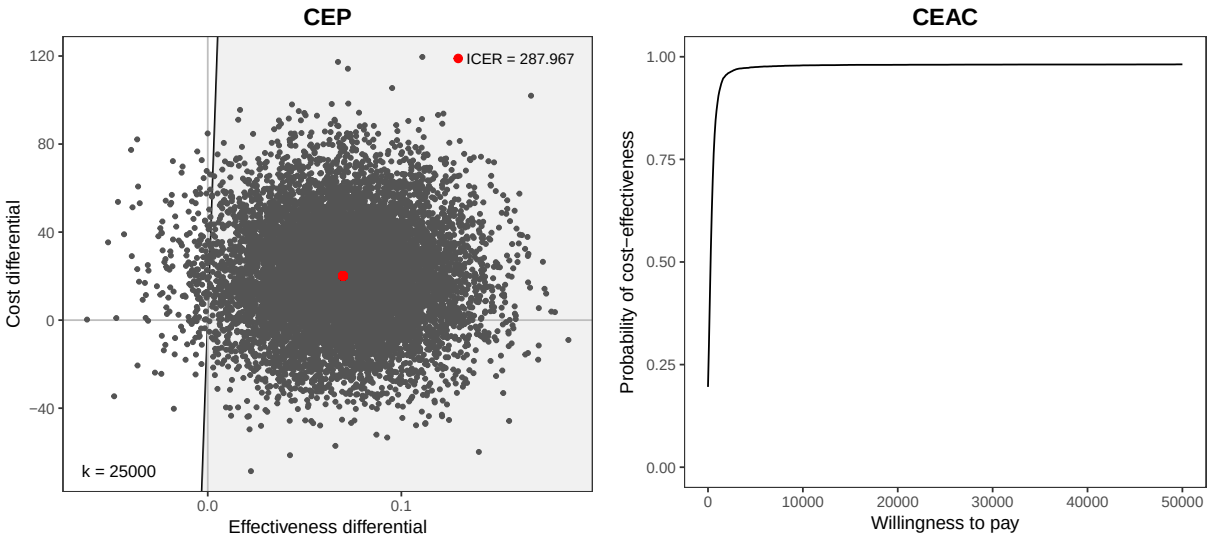


Figure 7: Cost Effectiveness Plane (CEP) and Cost Effectiveness Acceptability Curve (CEAC) obtained using respectively the functions `ceplane.plot` and `ceac.plot` in the package **BCEA** and applied to the model results contained in the object `model.hur$cea`

10 Conclusions

The **missingHE** package presented in this paper performs economic evaluations with missing outcome values for two-arms individual level trial datasets under a Bayesian framework using `JAGS`. In addition, the package provides different types of models that can be specified to jointly handle a series of issues typically affecting effectiveness and cost data which may bias parameter estimates (e.g. skewness and structural values).

Missing data represent a serious threat for the economic evaluation as, when confronted with a partially-observed dataset, each analysis makes assumptions about the missing values that cannot be ultimately verified from the data at hand. This means that any measure of fit or predictive accuracy, such as the Deviance Information Criterion or Posterior Predictive Checks (Spiegelhalter et al., 2004, Gelman et al. (2004)), cannot be easily interpreted. This in turns makes the validation of the model more complex (Mason et al., 2012). Thus, the use of sensitivity analysis to explore the impact on the results of a set of plausible missingness assumptions, included MNAR, becomes unavoidable. The Bayesian approach naturally allows to perform these assessments through the incorporation in the model of external evidence (e.g. expert opinions) using prior distributions while ensuring consistency and the correct propagation of uncertainty throughout the model.

missingHE handles missingness under alternative mechanism scenarios through the functions `selection` and `hurdle` which respectively implement selection and hurdle models for either or both outcome variables. Selection models assume a specifc MoA-MoM modelling structure, where either MAR or MNAR mechanisms

can be specified directly, while hurdle models are mostly based on MAR but allow the exploration of MNAR through specific assumptions about the structural values.

In conclusion, **missingHE** allows the analyst who wants to perform an economic evaluation to: a) jointly model costs and effectiveness; b) account for skewness and structural values; and c) assess the robustness of the results under a set of plausible missingness assumptions. These are typical issues affecting individual-level data that should be simultaneously addressed to avoid biased results, which may in turn lead to misleading cost-effectiveness conclusions. The availability of methodological and practical tools such as the ones used in this package have the potential to improve the workflow of modellers and regulators alike, thus advancing the fields of economic evaluation of health care interventions.

References

Arnold, J., Daroczi, G., Werth, B., B., W., Kunst, J., Auguie, B., Rudis, B., and Wickham, H. (2017). Package 'ggthemes'. https://cran.r-project.org/web/packages/ggthemes/.

Auguie, B. and Antonov, A. (2016). Package 'gridExtra'. https://cran.r-project.org/web/packages/gridExtra/.

Baio, G. (2013). *Bayesian Methods in Health Economics*. Chapman and Hall/CRC, University College London London, UK.

Baio, G. (2014). Bayesian models for cost-effectiveness analysis in the presence of structural zero costs. *Statistics in Medicine*, 33:1900–1913.

Baio, G., Berardi, A., and Heath, A. (2016). Package 'BCEA'. https://cran.r-project.org/web/packages/BCEA/.

Basu, A. and Manca, A. (2012). Regression estimators for generic health-related quality of life and quality-adjusted life years. *Mediacal Decision Making*, 1:56–69.

Black, W. (1990). A graphic representation of cost-effectiveness. *Medical Decision Making*, 10:212–214.

Brooks, S., Gelman, A., Jones, G., and Meng, X. (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

Carpenter, J. and Kenward, M. (2013). *Multiple Imputation and its Application*. John Wiley and Sons, Chichester, UK.

Carpenter, J., Kenward, M., and IR., W. (2007). Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Statistical Methods in Medical Research*, 16:259–275.

Cooper, N., Sutton, A., Mugford, M., and Abrams, K. (2003). Use of bayesian markov chain monte carlo methods to model cost-of-illness data. *Medical Decision Making*, 23:38–53.

Curtis, S., Goldin, I., and Evangelou, E. (2015). Package 'mcmcplots'. https://cran.r-project.org/web/packages/mcmcplots/.

Daniels, M. and Hogan, J. (2008). *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman and Hall, New York.

Faria, R., Gomes, M., Epstein, D., and White, I. (2014). A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *PharmacoEconomics*, 32:1157–1170.

Gabrio, A., Mason, A., and Baio, G. (2017). Handling missing data in within-trial cost-effectiveness analysis: A review with future recommendations. *PharmacoEconomics-Open*, 1:79–97.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis - 2nd edition*. Chapman and Hall, New York, NY.

Gelman, A., Sturtz, S., and Ligges, U. (2017). Package 'R2OpenBUGS'. https://cran.r-project.org/web/packages/R2OpenBUGS/.

Harkanen, T., Maljanen, T., Lindfors, O., Virtala, E., and Knekt, P. (2013). Confounding and missing data in cost-effectiveness analysis: comparing different methods. *Health Economics Review*, 3.

Little, R., D'Agostino, R., Dickersin, K., Emerson, S., Farrar, J., Frangakis, C., Hogan, J., Molenberghs, G., Murphy, S., Neaton, J., Rotnitzky, A., Scharfstein, D., Shih, W., Siegel, J., , and Stern, H. (2010). The prevention and treatment of missing data in clinical trials. panel on handling missing data in clinical trials. *Committee on National Statistics, Division of Behavioral and Social Sciences and Education*.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.

Manca, A., Hawkins, N., and Sculpher, M. (2005). Estimatingmean qalys in trial-based cost-e¡ectiveness analysis: the importance of controlling for baseline utility. *Health Economics*, 14:487–496.

Manca, P. and Palmer, S. (2006). Handling missing values in cost effectiveness analyses that use data from cluster randomized trials. *Appl Health Econ Health Policy*, 4:65–75.

Marin, X. (2016). Package 'ggmcmc'. https://cran.r-project.org/web/packages/ggmcmc/.

Mason, A., Richardson, S., Plewis, I., and Best, N. (2012). Strategy for modelling nonrandom missing data mechanisms in observational studies using bayesian methods. *Journal of Official Statistics*, 28:279–302.

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

Molenberghs, G., Fitzmaurice, G., Kenward, M., Tsiatis, A., and Verbeke, G. (2015). *Handbook of Missing Data Methodology*. Chapman and Hall, Boca Raton, FL.

Molenberghs, G. and Kenward, M. (2007). *Missing Data in Clinical Studies*. John Wiley and Sons, Chichester, UK.

National Research Council (2010). The prevention and treatment of missing data in clinical trials. The National Academies Press.

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

Noble, S., Hollingworth, W., and Tilling, K. (2012). Missing data in trial-based cost-effectiveness analysis: the current state of play. *Health Economics*, 21:187–200.

Ntzoufras, I. (2009). *Bayesian Modelling Using WinBUGS*. John Wiley and Sons, New York, USA.

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

Plummer, M. (2010). JAGS: Just Another Gibbs Sampler. http://www-fis.iarc.fr/~martyn/software/jags/.

Rascati, K., Smith, L., and Neilands, T. (2001). Dealing with skewed data: An example using asthma-related costs of medicaid clients. *Health Economics*, 23:481–498.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons, New York,USA.

Schafer, J. (1997). *Analysis of Incomplete Multivariate Data*. Chapman and Hall, New York,USA.

Spiegelhalter, D., Abrams, K., and Myles, J. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley and Sons, Chichester, UK.

Thompson, S. and Nixon, R. (2005). How sensitive are cost-effectiveness analyses to choice of parametric distributions? *Medical Decision Making*, 4:416–423.

Tooze, J., Grunwald, G., and Jones, K. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research*, 11:341–355.

van Asselt, A., van Mastrigt, G., Dirksen, C., Arntz, A., Severens, J., and Kessels, A. (2009). How to deal with cost differences at baseline. *PharmacoEconomics*, 27:519–528.

Van Buuren, S. (2012). *Flexible Imputation of Missing Data*. CRC press.

Van Hout, B., Al, M., Gordon, G., Rutten, F., and Kuntz, K. (1994). Costs, effects and c/e-ratios alongside a clinical trial. *Health Economics*, 3:309–319.

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

Wickham, H. and Chang, W. (2016). Package 'ggplot2'. https://cran.r-project.org/web/packages/ggplot2/.

Wood, A., White, I., and Thompson, S. (2004). Are missing outcome data adequately handled?a review of published randomized controlled trials in major medical journals. *Clinical Trials*, 1:368–376.

YS., S. and Yajima, M. (2015). Package 'R2jags'. https://cran.r-project.org/web/packages/R2jags/.