missingHE: A Package to Handle Missing Data in Health Economic Evaluations

Andrea Gabrio

Abstract

Keywords: Bayesian Statistics, Economic Evaluations, Missing Data.

1. Introduction

A well-known issue in health economic evaluations from individual-patient data, especially within a randomised controlled trial (RCT) setting, is the presence of partially-observed outcome variables (**?**). 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 (**?**).

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 possibly bias the parameter estimates (**???**). Nevertheless, many reviews concluded that most applied within-trial cost effectiveness analyses (CEAs) widely use these approaches or are unclear on the methodology used (**???**). 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; **?**), which has become extremely popular in clinical studies (**?**). In a nutshell, MI proceeds by replacing each missing data point with a value simulated from a suitable model and repeat this process for a number of times M. Thus, M complete (i.e. without missing data) replicates of the original dataset are created, each of which is analysed separately using standard methods. The individual estimates are then pooled together using meta-analytic tools, known 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 property of *congeniality*, i.e. the imputation model needs to be specified as equally or less restrictive than the analysis model (**?**). In many applications, MI is based upon a *missing at random* (MAR) assumption, which implies 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) assumptions, i.e. missingness cannot be explained fully by the observed data only. Neither MAR nor MNAR assumptions 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 (**?**).

The problem associated with the missing values in economic evaluations 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 and spikes) that may bias estimates from standard regression models (**???**). In routine analyses, appropriate methods to handle these complexities are rarely used, favouring standardised methods which however make strong simplifying assumptions such as: 1) normality for the underlying cost and effectiveness data, 2) independence between the outcomes and 3) failure to adjust for some potentially relevant baseline variables (**????**). Different methods have been proposed to handle these complexities, such as the use of bootstrapping or alternative parametric models (**?**) for addressing non-normality. However, when a combination of these affects the data, the building of a 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 all the complexities discussed above and produces several advantages in comparison to a frequentist counterpart (**??**). 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 (**?**), particularly when the evidence produced by the current study is limited, as is the case of small pilot trials. 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; **?**).

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

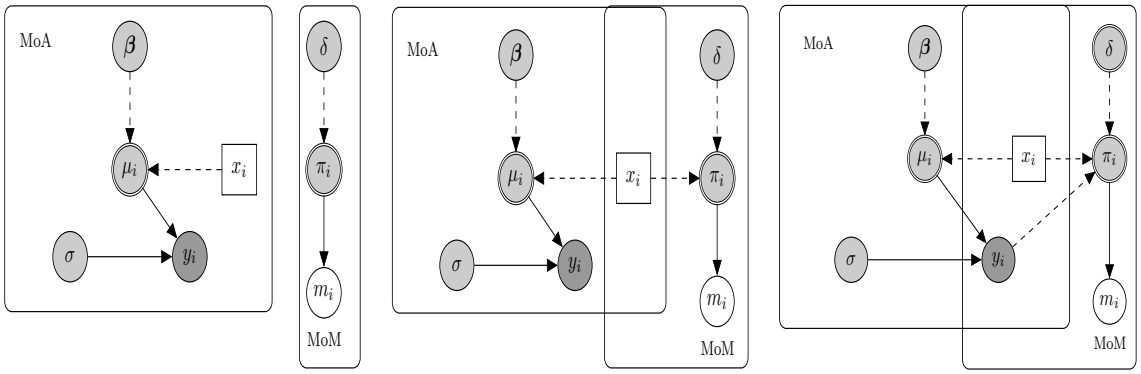
missingHE is a package designed to aid in the process of economic evaluations with missing outcome data. The modelling perspective is that of the Bayesian approach, exploiting its natural ability to assess the uncertainty associated with the missing values and its impact on decision-making problems. In fact, **missingHE** can be considered a wrapper for some other R packages. The first package, **R2jags** (**?**), allows to interface R with JAGS (**?**), a program for simulation from Bayesian hierarchical models using MCMC methods that is based on the BUGS modelling language (**?**). The second, **loo** (**?**), implements the methods described in **?** for estimating pointwise out-of-sample prediction accuracy from a fitted Bayesian model using within-sample fits for the purpose of model comparison. The third, **BCEA** (**?**), produces an economic evaluation output from the posterior inference generated from a JAGS model. **missingHE** also relies on other packages such as **ggplot2** (**?**), **gridExtra** (**?**), **ggthemes** (**?**), **mcmcplots** (**?**) and **ggmcmc** (**?**), mainly for graphics purposes.

3. Missing data mechanism

When analysing partially observed data, it is essential to investigate the possible reasons behind missingness. This formally translates into an *assumed* missing data mechanism (**?**) 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 $i = 1, \dots, n$ 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 $\boldsymbol{x}_i = (x_{1i}, \dots, x_{iJ})$, 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, 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 (**?**), 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 **??**



(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 mechanisms, 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.

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

1. Figure **??** (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 y_i.

2. Figure **??** (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 **??** (c) provides an example of *missing not at random* (MNAR). This is characterised by the dependence of the probability of missingness on both the partially and fully observed variables. Thus, π_i depends on both the predictor x_i and the 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, more structured assumptions that go beyond the information contained in the data are required.

Regardless of the setting, it is important to notice that it is never possible to distinguish between MAR and MNAR. 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 but have quite different implications for the unobserved data, leading to different conclusions (**?**). Therefore, it becomes crucial to explore the sensitivity of the results with respect to different missingness assumptions and quantify results' uncertainty.

4. Selection models

When informative missingness it thought to be the most realistic scenario, context-specific MNAR assumptions should be set as the reference case, with suitably-defined departures being explored in sensitivity analysis. This is typically implemented through advanced statistical methods, which can explicitly model a MNAR mechanism. **missingHE** focuses on a specific class of these methods, called *selection models* (**???**). To show how selection models work we consider a simple example. We assume a data set comprising a partially observed response y, the corresponding missing data indicator m, and a fully-observed covariate x. Under the selection model approach, the joint distribution $p(y, m)$ is factored as the product between the marginal distribution $p(y)$ and the conditional distribution $p(m \mid y)$, that is:

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

where, $\boldsymbol{\theta}^{MoA}, \boldsymbol{\theta}^{MoM}$ are the sets of parameters associated with the MoA and the MoM, respectively.

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 the 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, we can specify the joint distribution $p(e, c)$ using the following factorisation:

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

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** 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.

$$g_e(\phi_{ie}) = \alpha_0 \; [+ \ldots],$$

where α_0 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 $\mu_e = g_e^{-1}(\alpha_0)$ 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 (**??**). The location can be modelled as a function of the effectiveness variable as:

$$g_c(\phi_{ic}) = \beta_0 + \beta_f(e_i - \mu_e)\ [+\ldots],$$

Here, $(e_i - \mu_e)$ is the centered version of the effectiveness variable, while β_f quantifies the correlation between costs and effectiveness. Assuming other covariates are either centered or not present at all, $\mu_c = g_c^{-1}(\beta_0)$ 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 **??** 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.

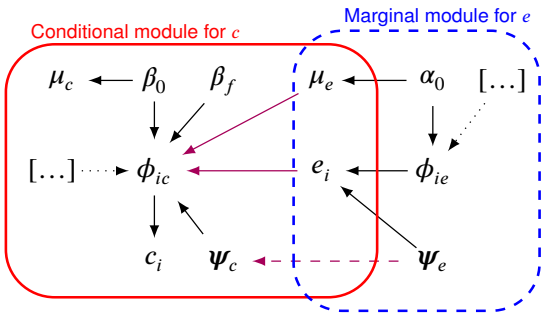


Figure 2: Joint distribution $p(e, c)$, expressed in terms of a marginal distribution for the effectiveness and a conditional distribution for the costs, respectively indicated with a solid red line and a dashed blue line. The solid black and magenta arrows show the dependence relationships between the parameters within and between the two modules, 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 **??** shows the different types of models for both the effectiveness and costs that are available in **missingHE**. In each, by default, minimally informative priors are specified on all

the relevant parameters.

MoA	Marginal Mean	Ancillary	Default Priors
Effectiveness			
$e_i \sim \text{Normal}(\phi_{ie}, \sigma_e^2)$	$\mu_e = \alpha_0$	σ_e	$\boldsymbol{\alpha} \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\log \sigma_e \sim \text{Uniform}(-5, 10)$
$e_i \sim \text{Beta}(\phi_{ie}\tau_{ie}, (1-\phi_{ie})\tau_{ie}))$	$\mu_e = \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}$	$\tau_{ie} = \frac{\phi_{ie}(1-\phi_{ie})}{\sigma_e^2} - 1$	$\boldsymbol{\alpha} \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\sigma_e \sim \text{Uniform}(0, \sqrt{\mu_e(1-\mu_e)})$
Cost			
$c_i \sim \text{Normal}(\phi_{ic}, \sigma_c^2)$	$\mu_c = \beta_0$	σ_c	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\log \sigma_c \sim \text{Uniform}(-5, 10)$
$c_i \sim \text{Gamma}(\phi_{ic}\tau_{ic}, \tau_{ic})$	$\mu_c = \exp(\beta_0)$	$\tau_{ic} = \frac{\phi_{ic}}{\sigma_c^2}$	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\sigma_c \sim \text{Uniform}(0, 1000)$
$c_i \sim \text{LogNormal}(\phi_{ic}, \sigma_c^2)$	$\mu_c = \exp(\beta_0 + \frac{\log \sigma_c^2}{2})$	σ_c	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000)$ $\log \sigma_c \sim \text{Uniform}(-5, 10)$

Table 1: A list of the distributions supported by **missingHE** for the effectiveness (e_i) and cost (c_i) variables. The set of logistic/log regression parameters for ϕ_{ie} and ϕ_{ic} are indicated with $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots)$. Notice that JAGS specifies Normal and LogNormal distributions in terms of the precision rather than the variance (precision=1/variance).

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

4.2. Model of missingness

The existence of a bivariate outcome in economic evaluations requires the specification of two different mechanisms when missingness affects both outcomes. Under a selection model, the MoM for the effectiveness $p(m_{ie} \mid e_i, \boldsymbol{\theta}_e^{MoM})$ and costs $p(m_{ic} \mid c_i, \boldsymbol{\theta}_c^{MoM})$ are indexed by the two different sets of parameters $\boldsymbol{\theta}_e^{MoM}$ and $\boldsymbol{\theta}_c^{MoM}$, respectively. These include the two missingness probabilities π_{ie} and π_{ic}, which in turn are expressed as functions of some other parameters. More specifically, **missingHE** models the two missingness probabilities using logistic regressions:

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

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

A feature of selection models is that missingness assumptions can be neatly fitted inside the Rubin's categories according to the variables included and form assumed for the MoM. The inclusion of the quantities δ_e and δ_c respectively denotes a MNAR mechanism for e and c. When, instead, these parameters are set to zero, the mechanisms become MAR (MCAR if no covariate is included in the models). Table **??** shows how **missingHE** specifies the MoM in the effectiveness and costs under either a MAR or MNAR assumption. By default, minimally informative priors are assumed for 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 distributional assumptions in the MoA (often difficult to check) and on the form of the MoM (on which unverifiable assumptions have to be made). Sensitivity analysis should always be used in order to explore alternative assumptions and assess missingness uncertainty. Two different types of sensitivity analysis for selection models are (**?**):

1. **Assumption sensitivity** varies the distribution assumptions in the MoA for e_i and c_i.
2. **Parameter sensitivity** 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 sensitivity analyses 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 more information may need to be gathered to better specify the model. Finally, it is important to stress that any conclusion derived from a MNAR model must be treated very cautiously. Each model specification makes assumptions about the missing data that can only be formulated using external information, should such information exist. Hence, sensitivity analysis is always required to the test the robustness of results to a range of plausible assumptions.

5. Hurdle models

In economic evaluations, outcome data are often affected by a substantial degree of skewness which standard methods typically fail to address. Especially under a Bayesian approach, alternative parametric modelling has been shown to be a valid approach to handle skewness, e.g. using a Gamma distribution for the costs and a Beta distribution for the effectiveness. However, when additional complexities affect the data, these methods may not be enough to obtain correct estimates. More specifically, a common type of idiosyncrasy is the presence of spikes at one of the boundaries of the range for the underlying outcome distributions. 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, which is defined on the range $(0, +\infty)$. Similarly, we may observe individuals who are associated with perfect health, i.e. unit QALY (**?**), which makes it difficult to use a Beta, 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 (**?**). 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 (**?**). A more efficient solution is the use of *hurdle models* (**??**). 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 (**???**).

Within the **missingHE** framework for the MoA shown in Figure **??**, we can extend all the models described in Table **??** 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. These are then modelled as

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

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 the structural values in 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 included and they contribute to estimate the probabilities, we have a *structural at random* (SAR) mechanism.

Table **??** 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 assumed 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).

Depending on the value of d_{ie} and d_{ic}, we can effectively partition the observed data into two sets: *structural* and *non-structural*:

1. The first one is made by the n^{se} subjects for whom $d_{ie} = 1$ and the n^{sc} subjects for whom $d_{ic} = 1$. For these individuals we define the variables $e_i^{se} = se$ and $e_i^{sc} = sc$, where se and sc denote the structural values.
2. The second one 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 the variables e_i^{-se} and c_i^{-sc}.

We can model the non-structural component using a distribution characterised by mean parameters 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 **??**, except the fact that they now apply to the non-structural components only. Then, 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 can be easily framed to handle missingness under different assumptions. If outcome data are unobserved then also the structural value indicators d_{ie} and d_{ic} cannot be directly computed. In general, these will be 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 alternative 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 each scenario, and assess the results across them. 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. 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 intervention group. 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 is allowed 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

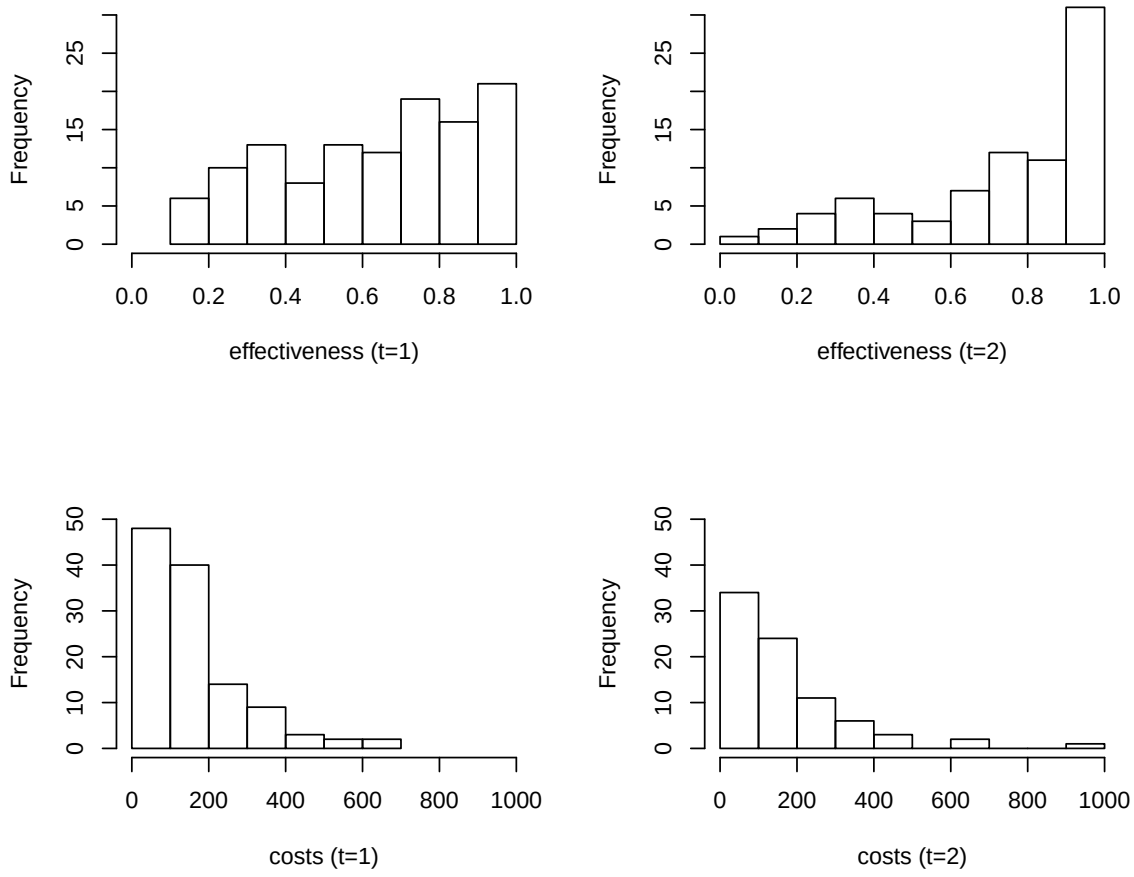


Figure 3: Histograms of the observed data distributions for the effectiveness and costs in the control (t=1) and intervention (t=2) groups

Figure **??** shows the histograms for the observed data distributions. The dataset consists of 250 individuals in total, grouped in two groups (here group 1 indicates the control and group 2 indicates the intervention). 51 individuals have unobserved outcome data (32 in the control and 19 in the intervention), and 24 individuals show 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, which is called from the R package **R2jags**. The program is based on the BUGS 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 first load the package using the command

```
R> require(missingHE)
```

Then we specify a joint bivariate Normal MoA distribution for the effectiveness and cost variables in the dataset `data`. With respect to the MoM, we assume a MNAR structure for the effectiveness while keeping a MAR assumption for the costs in both treatment arms. In **missingHE** we can implement the model using the `selection` function:

```
R> 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 arguments of the function have the following interpretations:

1. `data` must contain the 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 (`"norm"`) for both outcomes, Beta (`"beta"`) for the effectiveness and Gamma (`"gamma"`) or LogNormal (`"lnorm"`) for the costs.
3. `model.eff` and `model.cost` are formulas that specify which variables should be included in the effectiveness and cost MoA as covariates (among those available in `data`). A joint bivariate distribution can be assumed by placing `e` on the right-hand side of 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 MoMs as covariates (among those available 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.
7. `prior` specifies the prior distribtions to be used for the parameters of the model, by default those shown in Table **??** and Table **??**. These priors can be overwritten by the user. 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 (more on this later).
8. 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`).

missingHE allows user-specific priors for all parameters in the model. For example, for the model specified above, the priors for the parameters α_0 and δ_e can be provided as follows:

```
R> myprior <- list("alpha0.prior" = 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 by typing `help(selection)`.

The results of the model specified are stored in the object `model.sel`. The usual R command

```
R> 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 by typing the command

```
R> 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 related to the data provided as inputs to the function `selection`, such as effectiveness and cost data, total number of individuals in each group, number of observed and unobserved individuals and covariate data (if included in the model). The other elements in 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 elements 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 posterior samples of the marginal mean effectiveness and cost parameters, obtained 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; **?**) and the cost effectiveness acceptability curve (CEAC; **?**).

3. `"type"` is a character name that reports the type of missingness mechanism assumed by the model.

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

```
R> print(x = model.sel, value.mis = FALSE)
##                mean     sd     2.5%    97.5% Rhat n.eff
## alpha[1]      0.633  0.028    0.577    0.688    1  3800
## alpha[2]      0.737  0.032    0.673    0.800    1 10000
## beta[1]     147.593 12.473  122.784  171.954    1 10000
## beta[2]     154.011 18.669  117.305  190.638    1 10000
## beta_f[1]   -39.905 26.700  -91.975   13.313    1  3300
## beta_f[2]    -1.777 28.884  -59.506   54.533    1 10000
## delta_e[1]   -0.338  0.964   -2.201    1.619    1  1800
## delta_e[2]   -0.381  0.885   -2.078    1.438    1  2800
## deviance   3089.837  6.413 3077.766 3103.375    1 10000
## gamma_c[1]   -1.291  0.197   -1.683   -0.912    1  1200
## gamma_c[2]   -1.434  0.250   -1.950   -0.974    1  1400
## gamma_e[1]   -1.114  0.624   -2.489   -0.003    1  2300
## gamma_e[2]   -1.189  0.671   -2.663    0.005    1  1600
## mu_c[1]     147.593 12.473  122.784  171.954    1 10000
## mu_c[2]     154.011 18.669  117.305  190.638    1 10000
## mu_e[1]       0.633  0.028    0.577    0.688    1  3800
## mu_e[2]       0.737  0.032    0.673    0.800    1 10000
## p_c[1]        0.218  0.033    0.157    0.287    1  1200
## p_c[2]        0.195  0.039    0.125    0.274    1  1400
## p_e[1]        0.214  0.033    0.152    0.281    1 10000
## p_e[2]        0.192  0.039    0.121    0.275    1  4800
## s_c[1]      134.834  8.832  118.963  153.326    1  8600
## s_c[2]      171.385 13.895  146.689  200.962    1   480
## s_e[1]        0.259  0.017    0.228    0.296    1 10000
## s_e[2]        0.265  0.021    0.227    0.311    1  1600
```

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 for 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. We assume SCAR for the structural value mechanism in both outcomes. In **missingHE** we can implement the model using the `hurdle` function:

```
R> 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 available 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 directly modelled using the distribution specified in `dist_e` or `dist_c`.

4. `prior` specifies the prior distributions to be assumed for the parameters of the model, by default those shown in Table **??** and Table **??**. These priors can be overwritten in a similar way to what shown for the `selection` function. A list of the character names to be used to change each parameter's prior can be accessed by typing `help(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 rows in `data`, and should take value 1 or 0 to respectively associate each case with the structural or non-structural component in the hurdle model (more on this later).

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:

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

Once the new variable `my.d_e` is passed to the argument `d_e` in the function `hurdle`, **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.

```
R> print(model.hur, value.mis = FALSE)
##                mean     sd     2.5%    97.5% Rhat n.eff
## alpha[1]      0.363  0.093    0.176    0.543    1 10000
## alpha[2]      0.842  0.138    0.570    1.111    1 10000
## beta[1]       5.094  0.088    4.925    5.267    1  1000
## beta[2]       5.233  0.106    5.026    5.445    1  2100
## beta_f[1]    -0.312  0.232   -0.760    0.154    1 10000
## beta_f[2]     0.209  0.255   -0.264    0.739    1 10000
## deviance   7703.194  5.270 7694.839 7715.110    1 10000
## gamma_c[1]   -1.978  0.283   -2.564   -1.444    1  5000
## gamma_c[2]   -1.919  0.327   -2.600   -1.309    1 10000
## gamma_e[1]   -0.814  0.175   -1.165   -0.481    1 10000
## gamma_e[2]   -0.886  0.220   -1.332   -0.461    1  1200
## mu_c[1]     143.241 13.643  118.645  172.289    1  1700
## mu_c[2]     163.483 18.714  129.735  203.254    1  4200
## mu_e[1]       0.716  0.022    0.672    0.757    1 10000
## mu_e[2]       0.787  0.025    0.736    0.832    1  2700
## p_c[1]        0.125  0.030    0.071    0.191    1  4900
## p_c[2]        0.132  0.037    0.069    0.213    1 10000
## p_e[1]        0.308  0.037    0.238    0.382    1 10000
## p_e[2]        0.294  0.045    0.209    0.387    1  1100
## s_c[1]      139.018 14.261  114.614  171.085    1  2000
## s_c[2]      167.732 22.138  130.709  217.767    1  4200
## s_e[1]        0.227  0.011    0.206    0.250    1  1600
## s_e[2]        0.252  0.017    0.220    0.286    1 10000
```

7. Model comparison

After fitting a Bayesian model we may want to measure its predictive accuracy, for its own sake or for purposes of model comparison. Cross-validation and information criteria are two approaches to estimating out-of-sample predictive accuracy using within-sample fits. Three different measures of predictive accuracy for an object of class `"missingHE"` are available in **missingHE**:

1. Deviance information criterion (DIC; **?**).
2. Widely applicable information criterion (WAIC; **?**).
3. Leave-one-out information criterion (LOOIC; **?**).

All these methods estimate predictive accuracy using the log-likelihood evaluated at the posterior simulations of the parameters in the model. DIC is based on a point estimate of the parameter values (typically the posterior mean), while WAIC and LOOIC are fully Bayesian measures in that they use the whole posterior distribution of the parameters. We do not go into details about how these measures are constructed or their theoretical justifications as

these are outside the scope of this work. For a comprehensive analysis of the different types of predictive information criteria see **?**.

Information criteria can be used as relative measures of fit to compare nested models, with lower values typically indicating better fit. In **missingHE** it is possible to calculate the three types of predictive information criteria above usign the function `pic`. For example, considering the results of the hurdle model stored in the object `model.hur`, we can type:

```
R> pic.dic <- pic(x = model.hur, criterion = "dic", module = "total")
```

The arguments of the function `pic` are the following:

1. `x` must be an object of class `"missingHE"` that contains the model results.
2. `criterion` specifies which type of information criterion should be used to assess the predictive accuracy. Available choices are DIC (`"dic"`), WAIC (`"waic"`) and LOOIC (`"looic"`).
3. `module` specifies for which observed variables in the model the information criterion should be calculated. Available choices are all the observed data (`"total"`), only the effectiveness (`"e"`), only the cost (`"c"`) or both outcome data (`"both"`). For the purpose of model comparison, it is important that the information criteria are calculated on the basis of the same observed data to avoid misleading results. For example, it is not correct to compare the fit based on all the observed data of an hurdle model such as that stored in `model.hur` with the fit of another model which does not assume any structural values in the outcomes. The reason is that, among the observed data, the hurdle model includes the structural indicator variables which are absent from other models. In this case, we need to calculate the information criteria using only the observed data coming from either or both the effectiveness and cost modules to ensure a correct comparison between the model fits.

The output obtained from the function `pic` is a list which elements vary according to the specific information criterion specified. In this example, we choose DIC and the elements contained in the object `pic.dic` are the following:

1. `d_bar` is the posterior mean deviance.
2. `pd` is the effective number of parameters, computed as half the variance of the posterior deviance.
3. `dic` is the DIC, computed as: `DIC = d_bar + pD`.
4. `d_hat` is the deviance evaluated at the posterior mean of the parameters.

To see the complete list of the elements that can be obtained by selecting the WAIC and LOOIC choices, type `help(pic)`.

The actual DIC value for the model stored in `model.hur` can be accessed by typing:

```
R> pic.dic$dic
## [1] 7717.083
```

8. Diagnostic checks

As with any MCMC estimation, it is important to assess convergence. The function `diagnostic` in **missingHE** allows to visualise the model output and assess convergence for each parameter in the model. Different diagnostic tools and plots 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 to any model generated through the functions `hurdle` and `selection`. We can visually represent via histograms the posterior samples for the mean effectiveness parameters in the two groups using the command:

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

which displays the histograms in Figure **??**.

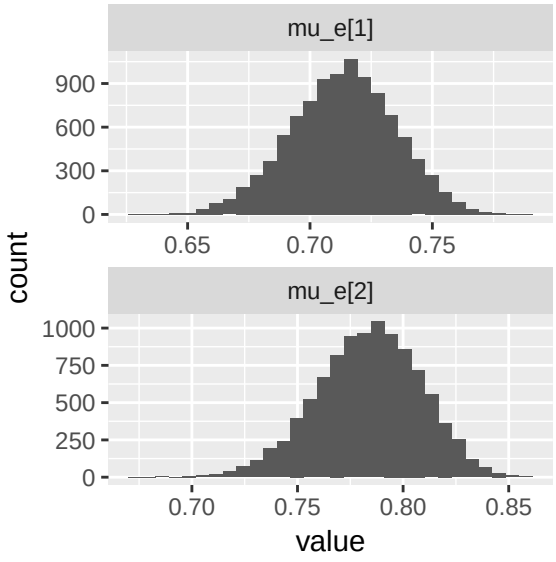


Figure 4: Histograms of the posterior distributions for the mean effectiveness parameters in the two groups

The arguments of the function `diagnostic` are the followings:

1. `x` must be an object of class `"missingHE"` which contains the model results.
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 accessed by typing `help(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 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 their corresponding character names can be accessed by typing `help(diagnostic)`. By default, all model parameters are selected using the character name `"all"`.

4. `theme` modifies the pre-defined background theme of the generated plots. Pre-defined themes are taken from the package **ggthemes** and must be indicated with corresponding character names. For a full list of available themes type `help(diagnostic)`.

It is possible to combine multiple graphs generated through the function `diagnostic` and saved 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, Figure **??** combines the density and trace plots for the mean effectiveness parameters from `model.hur`.

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

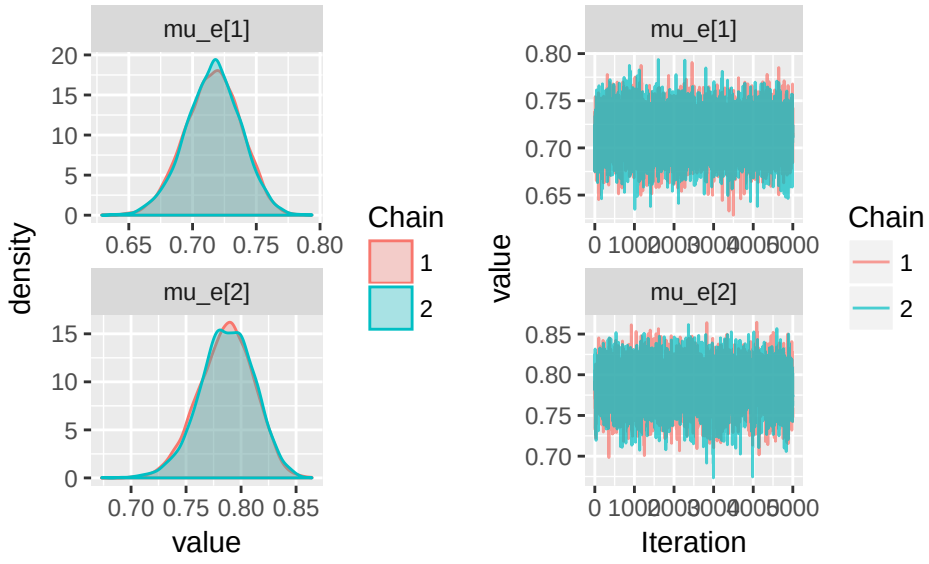


Figure 5: Density and trace plots of the posterior distribution of the mean effectiveness in the two groups

9. 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:

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

which displays the graphs in Figure **??**

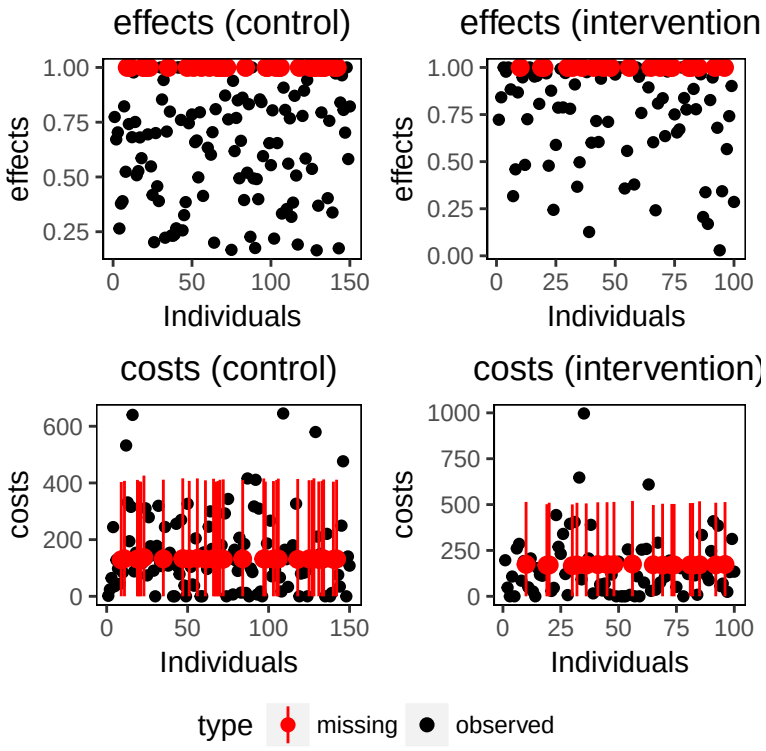


Figure 6: Scatter plots of the observed (black dots) and imputed (red dots and lines) values for both outcomes in the control and intervention groups. Imputation uncertainty is represented in terms of posterior means and 95% credible intervals.

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

1. `class` specifies the type of plot to be displayed. Two alternative character names are available: `"histogram"` and `"scatter"`. In the former, the observed and imputed data distributions are compared with an histogram plot and are distinguished by different colours. In the latter, the observed and imputed values (evaluated at the posterior means) are shown in a scatter plot, with imputed data that are 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`.

2. `outcome` specifies for which variable, either effectiveness, costs or both, and for which treatment group, either control, intervention or both, results should be visualised. For example, the plots for the effectiveness in both groups can be shown using the character name `"effects"`, while the plots by group can be accessed using the name `"arm1"` (control) or `"arm2"` (intervention). By default all plots are displayed using the character name `"all"`.

3. `theme` modifies the graphical background according to some pre-specified themes similarly to what shown for the function `diagnostic`.

10. Economic evaluation

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

```
R> summary <- summary(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    0.716  0.022   0.679   0.751
## mean costs    143.241 13.643 122.002 166.284
## sd effects      0.227  0.011   0.209   0.246
## sd costs      139.018 14.261 118.123 164.741
##
##  Reference intervention
##                  mean     sd      LB      UB
## mean effects    0.787  0.025   0.745   0.826
## mean costs    163.483 18.714 134.929 196.108
## sd effects      0.252  0.017   0.224   0.281
## sd costs      167.732 22.138 135.394 208.368
##
##  Incremental results
##                   mean     sd      LB    UB
```

```
## delta effects   0.071  0.033   0.017 0.124
## delta costs    20.242 23.424 -17.026 59.46
## ICER          286.192
```

Information is reported only for the main parameters of interest in the model for both outcomes and treatment groups. The incremental mean results for the effectiveness and costs are provided at the bottom of the table, denoted with `delta effects` and `delta costs` respectively. Results are summarised in terms of posterior mean, standard deviation and 95% credible intervals for each quantity. In addtion, the value of the incremental cost effectiveness ratio (ICER), which quantifies the cost per incremental unit of effectiveness, is also reported.

A series of useful functions are included in the package **BCEA** that summarises the results of the economic evaluation. For example, the CEP and CEAC plots can be obtained by applying their respective functions `ceac.plot` and `ceplane.plot` in **BCEA** to the `"BCEA"` object contained in `model.hur`, which can be accessed via `model.hur$cea`. The R commands used to generate and combine these plots are the following:

```
R> require(ggplot2)
R> require(BCEA)
R> cep <- ceplane.plot(model.hur$cea, graph = "ggplot2") + ggtitle("CEP")
R> ceac <- ceac.plot(model.hur$cea, graph = "ggplot2") + ggtitle("CEAC")
R> grid.arrange(cep, ceac, ncol = 2)
```

and the resulting output is given in Figure **??**

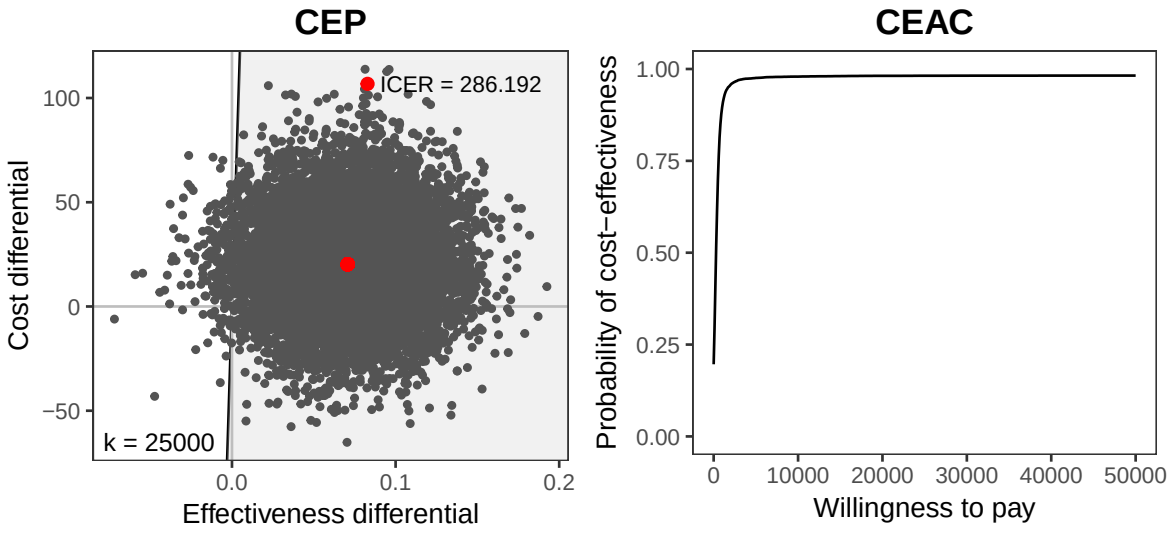


Figure 7: Cost effectiveness plane (CEP) and cost effectiveness acceptability curve (CEAC) obtained using the functions `ceplane.plot` and `ceac.plot` in the package **BCEA** and applied to the model results contained in the object `model.hur$cea`

11. Conclusions

The **missingHE** package performs economic evaluations with missing outcome values for two-arms trial data under a Bayesian framework. In addition, the package provides a set of models that can be flexibly specified to jointly handle a series of complexities typically affecting effectiveness and cost data, e.g. correlation, 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 information criteria or posterior predictive checks (**??**), cannot be easily interpreted. This, in turn, makes the validation of the model more complex. 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.

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 work of modellers and regulators alike, thus advancing the fields of economic evaluation of health care interventions.

References

Arnold J, Daroczi G, Werth B, Weitzner B, Kunst J, Auguie B, Rudis B, Wickham H (2017). ***ggthemes***. URL https://cran.r-project.org/web/packages/ggthemes/.

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

Baio G (2012). *Bayesian Methods in Health Economics*. Chapman & Hall, London.

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, Heath A (2016). ***BCEA***. URL https://cran.r-project.org/web/packages/BCEA/.

Basu A, 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, Meng X (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

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

Carpenter J, Kenward M, White I (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, 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, Evangelou E (2015). ***mcmcplots***. URL https://cran.r-project.org/web/packages/mcmcplots/.

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

Faria R, Gomes M, Epstein D, 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, 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, Rubin D (2004). *Bayesian Data Analysis*. 2nd edition. Chapman & Hall, New York.

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

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

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

Marin X (2016). ***ggmcmc***. URL https://cran.r-project.org/web/packages/ggmcmc/.

Mason A, Richardson S, Plewis I, 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, 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, Verbeke G (2015). *Handbook of Missing Data Methodology*. Chapman & Hall, Boca Raton.

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

Nixon R, 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, 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 & Sons, New York.

O'Hagan A, 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*. URL http://www-fis.iarc.fr/~martyn/software/jags/.

Rascati K, Smith L, 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 & Sons, New York.

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

SG Thompson R, Nixon (2005). "How Sensitive Are Cost-Effectiveness Analyses to Choice of Parametric Distributions?" *Medical Decision Making*, **4**, 416–423.

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

Spiegelhalter D, Best N, Carlin B, van der Linde A (2002). "Bayesian measures of model complexity and fit." *Journal of the Royal Statistical Society*, **64**, 583–639.

Su Y, Yajima M (2015). ***R2jags***. URL https://cran.r-project.org/web/packages/R2jags/.

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

Tooze J, Grunwald G, 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, 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, Kuntz K (1994). "Costs, Effects and C/E-Ratios Alongside a Clinical Trial." *Health Economics*, **3**, 309–319.

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

Vehtari A, Gelman A, Gabry J (2014). "Understanding predictive information criteria for Bayesian models." *Statistics and Computing*, **24**, 997–1016.

Vehtari A, Gelman A, Gabry J (2016). "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC." *Statistics and Computing*.

Vehtari A, Gelman A, Gabry J (2017). *loo*. URL https://cran.r-project.org/web/packages/loo/.

Watanable S (2010). "Asymptotic equivalence of Bayes cross validation and widely application information criterion in singular learning theory." *Journal of Machine Learning Research*, **11**, 3571–3594.

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

Wood A, White I, 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.

Affiliation:

Andrea Gabrio
University College London
Department of Statistical Science
Gower Street, London, WC1E 6BT (UK)
E-mail: ucakgab@ucl.ac.uk
URL: http://www.ucl.ac.uk/statistics/research/statistics-health-economics/current-projects/ag