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

Andrea Gabrio

Abstract

Health economic evaluations are an increasingly important component in the decision making process for the approval or reimbursement of new healthcare technologies in many developed countries. Economic evaluations can be generally defined as a suite of analytic approaches for combining costs and effectiveness of intervention(s) compared with a control, to aid decision making associated with resource allocation. Individual level cost and effectiveness data, e.g. from a clinical trial, are typically characterised by a series of complexities, such as high degress of skewness or correlations between the outcomes, that need to be taken into account in the statistical analysis to avoid biased results. An additional complication in the analysis is that effectiveness and cost data for some individuals are almost inevitably missing. If missingness is not handled through appropriate statistical methods, then inferences may be biased and can lead to misleading cost effectiveness conclusions. Bayesian statistical methods are well suited to handle missingness as they allow the incorporation of external evidence (e.g. expert opinion) into the analysis to inform the assumptons about the missing values and perform sensitivity analysis to a range of plausible assumptions. In addition, under the Bayesian approach, the imputation of the missing data is jointly performed with the estimation of the parameters of interest, which avoids potential issues due to a mismatch between the imputation and analysis models. In this paper, we review the features of the R package **missingHE**, which is designed to implement different approaches to handle missingness in health economic evaluation. In particular, **missingHE** uses full Bayesian methods (via the R package **R2jags**) to implement a wide range of flexible parametric models, which can jointly deal with missingness and other typical complexities that affect individual level data .

Keywords: Bayesian methods, economic evaluations, missing data, selection models, pattern mixture models, hurdle models.

1. Introduction

A typical problem in health economic evaluations based on individual patient data, especially from a randomised controlled trial (RCT), is that either or both cost and effectiveness outcome measures are partially observed (Faria, Gomes, Epstein, and White 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 (The National Academies Press 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 possibly introduces biases in the parameter

estimates (Rubin 1987; Schafer 1997; Molenberghs and Kenward 2007). Nevertheless, recent literature reviews suggest that most applied within trial cost effectiveness analyses (CEAs) use these approaches or are unclear on the methodology used to handle the missing data (Noble, Hollingworth, and Tilling 2012; Diaz-Ordaz, Kenward, Cohen, Coleman, and Eldridge 2014; Gabrio, Mason, and Baio 2017; B., Leurent, M., Gomes, and Carpenter 2018). 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 and repeat this process for a number of times M. This leads to the creation of M complete (i.e. without missing data) replicates of the original dataset, each of which is analysed separately using standard methods. The parameter estimates from each dataset are then pooled together using meta analytic tools, known as *Rubin's rules*, to reflect the inherent uncertainty in imputing the missing values. Because of the separation between the imputation and the analysis steps, inferences under MI are subject to potential biases when the imputation model is not correctly specified or *uncongenial*, i.e. the imputation model is specified as more restrictive than the analysis model (Van Buuren 2012). In addition, 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 and it is therefore crucial to perform *sensitivity analysis* (SA) to explore how variations in the assumptions about the missing values impact the results (Carpenter, Kenward, and White 2007).

The problem associated with the handling of the missing values in economic evaluations is often coupled with the fact that the empirical distributions of resource use and health related quality of life data, e.g. quality adjusted life years (QALYs), are generally affected by some complexities (e.g. correlation, skewness and spikes) that may bias the estimates from standard regression models (Rascati, Smith, and Neilands 2001; SG. Thompson and Nixon 2005; Basu and Manca 2012). In routine analyses, appropriate methods to handle these complexities are rarely used, favouring standardised methods which however make strong simplifying assumptions. Examples include the normality and independence assumptions for the underlying cost and effectiveness data and the failure to adjust for some baseline variables that are highly predictive of the outcomes (O'Hagan and Stevens 2001; Manca, Hawkins, and Sculpher 2005; Vazquez Polo, Hernandez, and Lopez-Valcarcel 2005; Van Asselt, van Mastrigt, Dirksen, Arntz, Severens, and Kessels 2009). Different methods have been proposed to handle these complexities, such as the use of bootstrapping or alternative parametric models to address skewness in the costs (Rascati *et al.* 2001). However, when a combination of these complexities affects the data, the building of a more statistically complex model that can simultaneously account for them is desirable.

A full Bayesian modelling framework provides a framework to jointly tackle the complexities discussed above and has some advantages in comparison to a frequentist counterpart (Spiegelhalter, Abrams, and Myles 2004; Baio 2012). More specifically, the Bayesian ap-

proach 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. In addition, the modular structure used by Bayesian methods allows for a relatively easy increase of model complexity to deal with the complexities of the data while ensuring the full characterisation and quantification of the uncertainty for each variable in the model. 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, effectively surrendering their computational advantage over intensive but efficient sampling algorithms such as Markov chain Monte Carlo methods (MCMC; Brooks, Gelman, Jones, and Meng 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 the impact on the inferences of alternative missingness assumptions and jointly account for different potential sources of bias 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. In fact, **missingHE** can be considered a wrapper for some other R packages. The first package, **R2jags** (Su 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 (Lunn, Jackson, Thomas, and Spiegelhalter 2012). The second, **loo** (Vehtari, Gelman, and Gabry 2017), implements the methods described in Vehtari, Gelman, and Gabry (2016) for estimating pointwise out of sample predictive accuracy from a fitted Bayesian model using within sample fits for the purpose of model comparison. The third, **BCEA** (Baio, Berardi, and Heath 2016), produces an economic evaluation output from the posterior inference generated from a JAGS model. **missingHE** also relies on other packages such as **ggplot2** (Wickham and Chang 2016), **gridExtra** (Auguie and Antonov 2016), **ggthemes** (Arnold, Daroczi, Werth, Weitzner, Kunst, Auguie, Rudis, and Wickham 2017), **mcmcplots** (Curtis, Goldin, and Evangelou 2015) and **ggmcmc** (Marin 2016), mainly for graphics purposes.

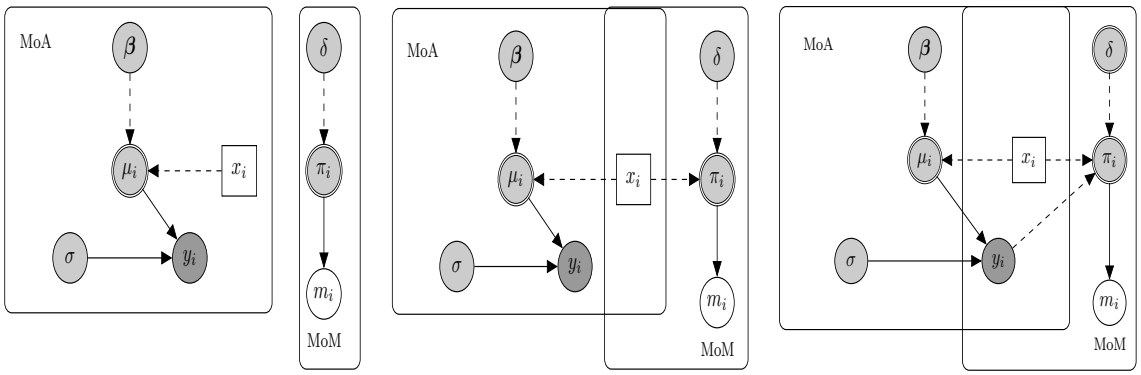
3. Missing data mechanism

When analysing partially observed data, it is essential to investigate the possible reasons behind missingness. This formally translates into an *assumed* missing data mechanism (Rubin 1987), which is 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 data mechanism.

We provide an intuitive representation of the missing data mechanism in a simplified setting with an univariate outcome variable. Consider a sample of $i = 1, \ldots, n$ individuals enrolled

in a trial 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 $\boldsymbol{x}_i = (x_{1i}, \ldots, x_{iJ})$, e.g. sex, age or comorbidities. 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 which takes value 1 if the $i-$th subject is associated with a missing outcome and 0 otherwise. We can represent this setting using two submodels, 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 (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 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. 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. Conse-

quently, 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.

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 mechanism 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.

Figure 1 (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 (Molenberghs, Fitzmaurice, Kenward, Tsiatis, and Verbeke 2015). Thus, it is crucial to explore the sensitivity of the results with respect to different missingness assumptions and quantify their impact on the inferences.

When dealing with missingness, (a) plausible benchmark(s) assumption(s) about the missing values 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** provides three alternative approaches to handle the missing data under a Bayesian framework and to assess the sensitivity of the results to alternative assumptions about the missing data mechanism.

4. Selection models

One approach is called *selection model* (Mason, Richardson, Plewis, and Best 2012; Molenberghs *et al.* 2015), which directly specifies the structure of the MoM and MoA modules. Consider a dataset 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 in a general Bayesian framework

which allows the user to choose among a set of alternative structures for both outcomes. Consider some patient level data that are available 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. To account for correlation between the outcomes, the joint distribution $p(e, c)$ can be generally specified 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** always specifies the joint distribution assuming 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 may be expressed as a function of the marginal variance σ_c^2 (Nixon and Thompson 2005; Baio 2012). 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 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.

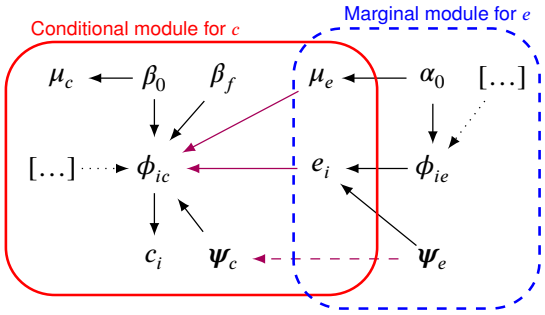


Figure 2: Joint distribution $p(e, c)$, expressed in terms of a marginal distribution for the effectiveness and a conditional distribution for the costs, respectively indicated with a solid red line and a dashed blue line. The solid black and magenta arrows show the dependence relationships between the 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 1 shows the different types of distributions for both the effectiveness and costs in the MoA that can be specified 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	$\boldsymbol{\alpha} \overset{iid}{\sim} \text{Normal}(0, 1000^2)$ $\log \sigma_e \sim \text{Uniform}(-5, 10)$
$e_i \sim \text{Beta}(\phi_{ie}\tau_{ie}, (1-\phi_{ie})\tau_{ie}))$	$\mu_e = \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}$	$\tau_{ie} = \frac{\phi_{ie}(1-\phi_{ie})}{\sigma_e^2} - 1$	$\boldsymbol{\alpha} \overset{iid}{\sim} \text{Normal}(0, 1000^2)$ $\sigma_e \sim \text{Uniform}(0, \sqrt{\mu_e(1-\mu_e)})$
Cost			
$c_i \sim \text{Normal}(\phi_{ic}, \sigma_c^2)$	$\mu_c = \beta_0$	σ_c	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000^2)$ $\log \sigma_c \sim \text{Uniform}(-5, 10)$
$c_i \sim \text{Gamma}(\phi_{ic}\tau_{ic}, \tau_{ic})$	$\mu_c = \exp(\beta_0)$	$\tau_{ic} = \frac{\phi_{ic}}{\sigma_c^2}$	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000^2)$ $\sigma_c \sim \text{Uniform}(0, 10000)$
$c_i \sim \text{LogNormal}(\phi_{ic}, \tau_c^2)$	$\mu_c = \exp(\beta_0 + \frac{\tau_c^2}{2})$	τ_c	$\boldsymbol{\beta} \overset{iid}{\sim} \text{Normal}(0, 1000^2)$ $\tau_c \sim \text{Uniform}(0, 100)$

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, the cost regression model is indexed by an addtional parameter β_f, which captures the extent of the linear association between the outcomes.

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 $[+\ldots]$.

Selection models directly specificies the missingness mechanism, which can be neatly interpreted within the Rubin's categories according to the variables included and form assumed for the MoM. For example, 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 2 shows how **missingHE** specifies the MoM for 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	$\bar{\pi}_e = \frac{\exp(\gamma_{0e})}{1+\exp(\gamma_{0e})}$	$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000^2)$
	MNAR	$\bar{\pi}_e = \frac{\exp(\gamma_{0e}+\delta_e \bar{e}_i)}{1+\exp(\gamma_{0e}+\delta_e \bar{e}_i)}$	$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000^2)$; $\delta_e \sim \text{Normal}(0, 1)$
Cost			
$m_{ic} \sim \text{Bernoullli}(\pi_{ic})$	MAR	$\bar{\pi}_c = \frac{\exp(\gamma_{0c})}{1+\exp(\gamma_{0c})}$	$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000^2)$
	MNAR	$\bar{\pi}_c = \frac{\exp(\gamma_{0c}+\delta_c \bar{c}_i)}{1+\exp(\gamma_{0c}+\delta_c \bar{c}_i)}$	$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000^2)$; $\delta_c \sim \text{Normal}(0, 1)$

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 directly specify 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 these assumptions can be translated 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). Thus, sensitivity analysis should always be performed to explore alternative assumptions and assess the impact of missingness uncertainty on the results. Two different types of sensitivity analysis that have been proposed for selection models are *assumption sensitivity* and *parameter sensitivity* (Mason *et al.* 2012).

Assumption sensitivity varies the distributional assumptions of the MoA for e_i and c_i, while parameter sensitivity varies the prior distributions for δ_e and δ_c to explore alternative MoM structures. These priors are typically informed based on the available information (e.g. expert opinion) to define a set of plausible missingness assumptions to explore.

After the required sensitivity analyses have been performed, it is important to examine the results and establish how much the quantities of interest vary across the scenarios assessed. In the event conclusions are not robust to these alternative, then more information may need to be gathered to better specify the model.

5. Pattern mixture models

Another approach implemented in **missingHE** to handle MNAR mechanisms is called *Pattern mixture model* (Little and Rubin 2002; Daniels and Hogan 2008; Molenberghs *et al.* 2015), which factors the joint distribution $p(y, m)$ into the product between the marginal distribution $p(m)$ and the conditional distribution $p(y \mid m)$, that is:

$$p(y, m \mid x, \boldsymbol{\eta}^{MoR}, \boldsymbol{\eta}^{MoP}) = p(m \mid \boldsymbol{\eta}^{MoP}) p(y \mid m, x, \boldsymbol{\eta}^{MoR})$$

where, $\boldsymbol{\eta}^{MoR}, \boldsymbol{\eta}^{MoP}$ are the sets of parameters associated with the *Model of Response* (MoR), i.e. the model of the outcome defined within each missingness pattern, and the *Model of Patterns* (MoP), respectively. The marginal distribution of the outcome is then typically obtained as

$$p(y \mid x, \boldsymbol{\eta}) = \sum p(y \mid m, x, \boldsymbol{\eta}^{MoR}) p(m \mid \boldsymbol{\eta}^{MoP})$$

that is, $p(y)$ is dervied as the weighted average of the MoR distributions over the missingness patterns, where the weights are represented by the estimated probabilities associated with each pattern.

5.1. Model of Patterns

The pattern specific probabilities in the MoP are typically estimated by directly modelling the missingness patterns indicator variable m. In health economic evaluations, the bivariate cross sectional outcome e, c allows up to a total of 4 different missingness patterns. **missingHE** assigns each individual to his/her corresponding pattern, denoted with a different number, using an indicator variable d_i. The numbers associated with each potential pattern are: 1) both outcomes observed (completers); 2) c observed and e missing; 3) e observed and c missing; 4) both outcomes missing. The patterns indicatir variable d_i is modelled using a

Multinomial distribution, defined on the support $1, \ldots, D$, where D is the total number of observed patterns.

The joint set of the probabilities associated with each pattern $\boldsymbol{\pi}_d = \left(\pi_d^{(1)}, \ldots, \pi_d^{(D)}\right)$ is then assigned a Dirichlet prior

$$\boldsymbol{\pi}_d \sim \text{Dirichlet}\left(\alpha_d^{(1)}, \ldots, \alpha_d^{(D)}\right)$$

where $\boldsymbol{\alpha}_d = \left(\alpha_d^{(1)}, \ldots, \alpha_d^{(D)}\right)$ is the joint set of the concentration parameters of the Dirichlet distribution. By default, **missingHE** assumes a noninformative prior $\boldsymbol{\alpha}_d = \mathbf{1}$, which assigns the same weight to each probability.

5.2. Model of Response

missingHE specifies the MoR using the same modelling approach of Section 4.1 but specified within each missingness patter. The distributions assigned to the effectiveness and cost variables (and their default priors) are the same to those in Table 1. However, the MoR cannot be fully identified based only on the observed data because in some patterns the observations are missing.

To identify the MoR a combination of *identifying restrictions* and *sensitivity parameters* is typically used (Daniels and Hogan 2008). Different types of identifying restrictions are available but their common purpose is to identify the distribution of the missing values in each pattern d using some constraints, usually calibrated on the observed data. For example, the parameters that index the effectiveness model in $d = 2$ can be identified by equating them to the corresponding parameters estimated from the observed data in $d = 1$. In general, different types of identifying restrictions could be used to identify the unidentified parameters in the MoR. **missingHE** identifies these parameters using the corresponding parameters estimated from $d = 1$, where both outcomes are observed. Under this type of identifying restrictions, the distribution of the missing data in each pattern is identified using the information from the observed data (more specifically from the completers), the missingness mechanism is assumed to be ignorable. However, pattern mixture models can be extended to explore some MNAR departues through some sensitivity parameters $\boldsymbol{\Delta} = (\Delta_e, \Delta_c)$. These are combined with the identifying restrictions to incorporate some external information about the distribution of the missing effectiveness and cost data into the model. For example, we can identify the mean outcome parameters in $d = 4$ ($\boldsymbol{\mu}^{(4)}$) using the corresponding parameters in $d = 1$ ($\boldsymbol{\mu}^{(1)}$) and then add $\boldsymbol{\Delta}$:

$$\mu_e^{(4)} = \mu_e^{(1)} + \Delta_e \quad \text{and} \quad \mu_c^{(4)} = \mu_c^{(1)} + \Delta_c.$$

Typically, sensitivity analysis is performed by setting a benchmark scenario (e.g. $\boldsymbol{\Delta} = \mathbf{0}$), and then explore alternative departures from it using different priors on $\boldsymbol{\Delta}$. In general, $\boldsymbol{\Delta}$ could be included in the model in different ways, e.g. as additive or multiplicative terms, and can be specified for each unidentified parameter of the model, e.g. both means and standard deviations. However, given that only a limited number of plausible MNAR scenarios are usually assessed, in practice, only variations related to the main parameters of interest are considered. Specifically, **missingHE** assesses MNAR only with respect to the mean outcome

parameters (object of interest of the economic evaluation) by additively including $\boldsymbol{\Delta}$ in the models. These parameters are then assigned Uniform priors

$$\Delta_e \sim \text{Uniform}(L_e, U_e) \quad \text{and} \quad \Delta_c \sim \text{Uniform}(L_c, U_c),$$

where (L_e, U_e) and (L_c, U_c) are the lower and upper boundaries for the mean effectiveness and cost sensitivity parameters' distributions. In **missingHE**, these hyperparameter values must be provided by the user according to the type of MNAR assumption that he/she wants to explore in sensitivity analysis.

6. Hurdle models

The last approach implemented in **missingHE** to handle missingness is called *hurdle model* (Ntzoufras 2009; Baio 2014), which does not correspond to a proper principled missing data method. However, hurdle models allow the imputation of specific values in the range of the observed data, which are often difficult to replicate using standard parametric distributions, and to assess the impact on the results of alternative assumptions about the potential proportion of these values that could be observed in the data.

In economic evaluations, hurdle models are well suited to handle both cost and effectiveness, which are typically characterised by some spikes at one of the boundaries of the range for the data. These values may not fall in the support of some parametric distributions, in which case some rescaling of the observed data is required to fit the models (Cooper, Sutton, Mugford, and Abrams 2003; Basu and Manca 2012). 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, some individuals may be associated with perfect health, i.e. unit QALY, which makes it difficult to use a Beta, defined on the open interval $(0, 1)$. 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, Briggs, O'Hagan, and Thompson 2011).

Hurdle models can efficiently handle the spikes using a mixture model approach, which is 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 weight 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, Grunwald, and Jones 2002; Harkanen, Maljanen, Lindfors, Virtala, and Knekt 2013; Baio 2014).

In **missingHE** all the models described in Table 1 can be extended to a hurdle version for either or both the outcome variables. For example, consider the situation in which some structrual values are observed in both the effectiveness (se) and the cost (sc) data. Then, **missingHE** specifies an hurdle model for both outcomes by generating for each subject $i = 1, \ldots, n$ two indicator variables d_{ie} and d_{ic}, which respectively take value 1 if the $i-$th individual is associated with a structural value in the effectiveness ($e_i = se$) and costs ($c_i = sc$) and 0 otherwise. These are then modelled as

$$d_{ie} := \mathbb{I}(e_i = se) \sim \text{Bernoulli}(\pi_{ie}) \quad \text{and} \quad d_{ic} := \mathbb{I}(c_i = sc) \sim \text{Bernoulli}(\pi_{ic})$$

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

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}. The inclusion of other centered covariates in the model of d_{ie} and d_{ic} is indicated by the terms $[+\ldots]$. 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 in the model. To make a parallel with the missing data literature, when no covariates are included in the models of d_{ie} and d_{ic}, then $\bar{\pi}_e$ and $\bar{\pi}_c$ are randomly predicted. This implies what we call a *structural completely at random* (SCAR) mechanism, in which the chance of observing an individual associated with a 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 3 shows how **missingHE** specifies the model for the structural indicators for the effectiveness and costs under either a SCAR or SAR mechanism 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	$\bar{\pi}_e = \frac{\exp(\gamma_{0e})}{1+\exp(\gamma_{0e})}$	$\gamma_{0e} \sim \text{Normal}(0, 1000^2)$
	SAR		$\boldsymbol{\gamma}_e \overset{iid}{\sim} \text{Normal}(0, 1000^2)$
Cost			
$d_{ic} \sim \text{Bernoullli}(\pi_{ic})$	SCAR	$\bar{\pi}_c = \frac{\exp(\gamma_{0c})}{1+\exp(\gamma_{0c})}$	$\boldsymbol{\gamma}_{0c} \sim \text{Normal}(0, 1000^2)$
	SAR		$\boldsymbol{\gamma}_c \overset{iid}{\sim} \text{Normal}(0, 1000^2)$

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 taken by d_{ie} and d_{ic}, the observed data can be effectively partition 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 corresponding variables $e_i^{se} = se$ and $e_i^{sc} = sc$.

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 corresponding variables $e_i^{-se} \neq se$ and $c_i^{-sc} \neq sc$.

missingHE models the non-structural components in the outcomes (e_i^{-se} and c_i^{-sc}) by fitting to these values one of the distributions among those in Table 1, parameterised in terms of the mean parameters μ_e^{-se} and μ_c^{-sc}. The parameterisation and priors used are exactly the same as those in Table 1. Then, the overall population averages effectiveness and costs in both treatment groups are obtained by the linear combination of the two mixture components

$$\mu_e = (1 - \bar{\pi}_e)\mu_e^{-se} + \bar{\pi}_e se \quad \text{and} \quad \mu_c = (1 - \bar{\pi}_c)\mu_c^{-sc} + \bar{\pi}_c sc,$$

where $\bar{\pi}_e$ and $\bar{\pi}_c$ are the estimated marginal probabilities of the structural values. For example, when $se = 1$ and $sc = 0$, then the mean effectiveness and costs are $\mu_e = (1 - \bar{\pi}_e)\mu_e^{-se} + \bar{\pi}_e$ and $\mu_c = (1 - \bar{\pi}_c)\mu_c^{-sc}$, respectively.

6.1. Hurdle models and missingness

When outcome data are partially observed, then also the structural value indicators in the hurdle model d_{ie} and d_{ic} cannot be directly computed. In general, these will be imputed based on the available information in the model. 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. However, hurdle models offer a convenient framework to explore the robustness of the results to some alternative MNAR 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. by varying the number of structural values potentially observed in each scenario. 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 the impact of missingness uncerrtainty on the results more thoroughly. In particular, if one of the scenarios explored 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 or pattern mixture models.

7. Example

We use a running example to show how selection, pattern mixture and hurdle models can be specified in **missingHE**. All types of models in **missingHE** are implemented 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 MCMC algorithms.

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 dataframe (say, `data`) that can be visualised using, for example, histogram plots

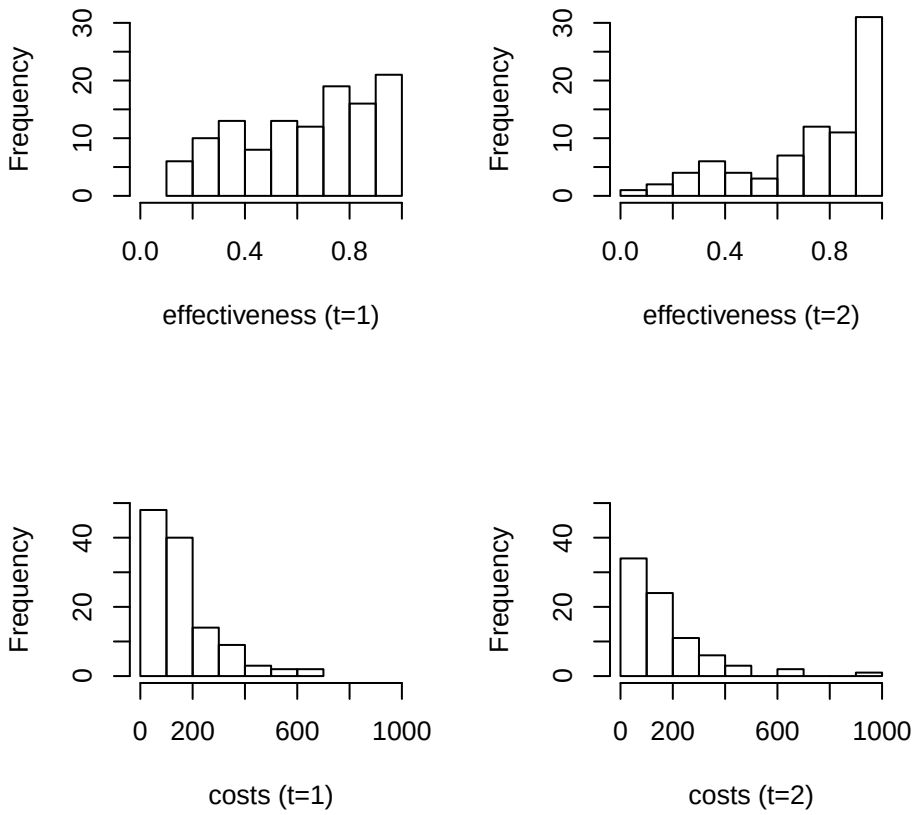


Figure 3: Histograms of the observed data distributions for the effectiveness and costs in the control (t=1) and intervention (t=2) groups

Figure 3 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 ($se = 1$) and costs ($sc = 0$).

7.1. Selection models in missingHE

To illustrate how selection models can be implemented in **missingHE**, 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. We then assume a MNAR structure for the MoM of the effectiveness while keeping a MAR assumption for the MoM of the costs in both treatment groups. In **missingHE** the model is implemented 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 = 1000, prior = "default")
```

where the arguments of the function have the following interpretations:

1. `data` must contain the data to analyse, specified in a dataframe format
2. `dist_e` and `dist_c` indicate the assumed effectiveness and cost MoA distributions, specified as character names among a set of predefined 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 contained 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 the model assumes independence between the outcomes.
4. `model.me` and `model.mc` are formulas that specify which variables should be included in the effectiveness and cost MoM as covariates (among those contained 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 covariates are included (indicated with `1`), implicitly assuming a MAR mechanism.
5. `type` specifies the type of missing data mechanism, 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 1 and Table 2. 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 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 if the model text file should be saved in the current working directory (`save_model`).

missingHE allows the user to overwrite the deafult hyperprior values in the prior distributions for all parameters in the model. For example, in the model specified above, user specific 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`. **missingHE** requires 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 parameters can be accessed by typing `help(selection)`.

The results of the model fitted using **missingHE** 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 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 elements are related to the data provided as inputs to the function `selection`, such as the effectiveness and cost data, the total number of individuals in each group, the number of observed and unobserved individuals and the covariate data (if included in the model). Other objects contained in `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 and contain the posterior samples of the parameters of interest. 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, and obtained using the function `bcea` in the package **BCEA**. This object can be analysed using tailored functions of **BCEA** to derive standard CEA outputs such as the cost effectiveness plane (CEP; Black 1990) and the cost effectiveness acceptability curve (CEAC; Van Hout, Al, Gordon, Rutten, and Kuntz 1994).

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 of the model.

```
R> print(x = model.sel, value.mis = FALSE)
##                mean     sd     2.5%    97.5% Rhat n.eff
## alpha[1]      0.631  0.028    0.578    0.685 1.02    74
## alpha[2]      0.740  0.032    0.676    0.800 1.02    69
## beta[1]     150.410 12.893  125.777  175.661 1.00  1000
## beta[2]     159.082 19.008  121.101  195.929 1.00  1000
## beta_f[1]  -133.390 46.467 -225.451  -37.810 1.00   390
## beta_f[2]   -13.827 74.278 -163.791  128.173 1.00   820
## delta_e[1]   -0.432  0.891   -2.090    1.308 1.19    12
## delta_e[2]   -0.321  0.881   -2.188    1.221 1.09    24
## deviance   3087.957  6.509 3076.069 3101.276 1.01    99
## gamma_c[1]   -1.293  0.206   -1.712   -0.896 1.00   990
## gamma_c[2]   -1.440  0.256   -1.996   -0.975 1.00  1000
## gamma_e[1]   -1.054  0.570   -2.211   -0.029 1.16    14
## gamma_e[2]   -1.223  0.672   -2.505    0.117 1.10    21
## mu_c[1]     150.410 12.893  125.777  175.661 1.00  1000
## mu_c[2]     159.082 19.008  121.101  195.929 1.00  1000
## mu_e[1]       0.631  0.028    0.578    0.685 1.02    74
## mu_e[2]       0.740  0.032    0.676    0.800 1.02    69
## p_c[1]        0.217  0.035    0.153    0.290 1.00  1000
## p_c[2]        0.195  0.039    0.120    0.274 1.00  1000
## p_e[1]        0.213  0.032    0.153    0.279 1.00   510
## p_e[2]        0.193  0.039    0.123    0.276 1.00   550
## s_c[1]      132.917  8.613  117.600  151.238 1.00  1000
## s_c[2]      172.275 14.322  145.924  202.063 1.00  1000
## s_e[1]        0.263  0.018    0.228    0.303 1.00  1000
## s_e[2]        0.263  0.022    0.226    0.311 1.01  1000
```

The optional argument `value.mis` allows to either exclude (`FALSE`) or include (`TRUE`) the posterior results associated with the imputed values; by default, these values are omitted from the results displayed.

7.2. Pattern mixture models in missingHE

In **missingHE**, pattern mixture models can be implemented using the function `pattern`. For example, we specify a joint normal distribution for both outcomes within each missingness pattern and specify a MNAR assumption for the effectiveness. The range of the variation of the sensitivity parameters Δ_e in both treatment groups is specified by generating the following object

which is then passed into the `pattern` function.

```
R> model.pat <- pattern(data = data, dist_e = "norm", dist_c = "norm",
+               model.eff = e ~ 1, model.cost = c ~ e,
+               Delta_e = Delta_e_range, Delta_c = 0,
+               type = "MNAR", n.iter = 1000, prior = "default")
```

Many arguments have similar interpretations to those of the `selection` function, with the exception of `Delta_e` and `Delta_c`, which contain the boundary values for the Uniform prior on the sensitivity parameters for the mean efectiveness and cost parameters, respectively. Under MAR both sensitivity parameters are set to 0 (default). Under MNAR, the values must be provided for either one or both outcomes (depending for which variable MNAR is assumed). In this example, we have specified the values in the object `Delta_e_range` assuming that the missing effects are associated with lower values compared with the observed effects. Specifically, we have specified the range for Δ_e in each treatment group between minus the observed standard deviation for the effectiveness (lower bound `[L]`) and 0 (upper bound `[U]`). This information is passed to the function `pattern` through the object `Delta_e_range`, which is computed as

```
R> Delta_e_range <- matrix(NA, 2, 2)
R> Delta_e_range[1, ] <- c(L[1], U[1])
R> Delta_e_range[2, ] <- c(L[2], U[2])
```

It corresponds to a 2x2 matrix, where the rows indicate the treatment group and the columns represent the lower and upper bounds for the range of Δ_e.

As for the function `selection`, user specific priors on the parameters of the model can be provded through a list object using the argument `prior`. A list of the character names to be used by type of model and parameters can be accessed by typing `help(pattern)`. Finally, a summary of the posterior results of the model can be accessed using the function `print`.

```
R> print(x = model.pat, value.mis = FALSE)
##                   mean     sd     2.5%    97.5% Rhat n.eff
## Delta_e[1]      -0.124  0.074   -0.250   -0.007 1.01   260
## Delta_e[2]      -0.132  0.076   -0.256   -0.007 1.00  1000
## alpha_p1[1]      0.639  0.025    0.590    0.687 1.01   200
## alpha_p1[2]      0.639  0.025    0.590    0.687 1.01   200
## alpha_p2[1]      0.743  0.030    0.682    0.801 1.00   720
## alpha_p2[2]      0.743  0.030    0.682    0.801 1.00   720
## beta_f_p1[1] -133.281 48.315 -228.257  -36.847 1.00  1000
## beta_f_p1[2] -133.281 48.315 -228.257  -36.847 1.00  1000
## beta_f_p2[1]  -12.180 72.925 -148.971  127.544 1.00  1000
```

```
## beta_f_p2[2]  -12.180 72.925 -148.971  127.544 1.00  1000
## beta_p1[1]    148.918 12.658  124.767  175.314 1.01   380
## beta_p1[2]    148.918 12.658  124.767  175.314 1.01   380
## beta_p2[1]    159.616 18.717  124.038  196.083 1.00  1000
## beta_p2[2]    159.616 18.717  124.038  196.083 1.00  1000
## deviance     2834.942  5.037 2827.112 2846.975 1.01   170
## mu_c[1]       148.918 12.658  124.767  175.314 1.01   380
## mu_c[2]       159.616 18.717  124.038  196.083 1.00  1000
## mu_c_p1[1]    148.918 12.658  124.767  175.314 1.01   380
## mu_c_p1[2]    148.918 12.658  124.767  175.314 1.01   380
## mu_c_p2[1]    159.616 18.717  124.038  196.083 1.00  1000
## mu_c_p2[2]    159.616 18.717  124.038  196.083 1.00  1000
## mu_e[1]         0.612  0.030    0.554    0.667 1.00  1000
## mu_e[2]         0.717  0.034    0.648    0.781 1.00   910
## mu_e_p1[1]      0.639  0.025    0.590    0.687 1.01   200
## mu_e_p1[2]      0.514  0.077    0.380    0.645 1.00   730
## mu_e_p2[1]      0.743  0.030    0.682    0.801 1.00   720
## mu_e_p2[2]      0.610  0.082    0.469    0.754 1.00   820
## p_prob1[1]      0.784  0.034    0.716    0.849 1.00   610
## p_prob1[2]      0.216  0.034    0.151    0.284 1.00   720
## p_prob2[1]      0.805  0.039    0.727    0.876 1.00  1000
## p_prob2[2]      0.195  0.039    0.124    0.273 1.00  1000
## s_c_p1[1]     133.085  9.001  117.246  151.793 1.00   610
## s_c_p1[2]     133.085  9.001  117.246  151.793 1.00   610
## s_c_p2[1]     171.923 13.416  148.472  200.624 1.00  1000
## s_c_p2[2]     171.923 13.416  148.472  200.624 1.00  1000
## s_e_p1[1]       0.259  0.018    0.228    0.297 1.02   660
## s_e_p1[2]       0.259  0.018    0.228    0.297 1.02   660
## s_e_p2[1]       0.263  0.021    0.224    0.308 1.00   690
## s_e_p2[2]       0.263  0.021    0.224    0.308 1.00   690
```

The results are displayed for each parameter of interest conditional on each missingness pattern contained in the data. The aggregate mean effectiveness `mu_e` and costs `mu_c` are derived by averaging the estimates across the patterns using the estimated pattern probabilities as weights.

7.3. Hurdle models in missingHE

We illustrate how to specify a Hurdle model in **missingHE** assuming a joint Beta Gamma distribution for the effectiveness and cost variables. Individuals associated with a unit effectiveness and zero costs are considered structural values and are explicitly handled. We assume a SCAR mechanism for the structural values in both outcomes in the hurdle model. 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 = 1000, prior = "default", d_e = my.d_e)
```

Some of the arguments have the same interpretation of those in the function `selection` and `pattern`, 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 covariates are 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` indicate 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 1 and Table 3. 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)`.

Similarly to the `selection` and `pattern` function, it is possible to access the posterior results of the model using the `print` function.

```
R> print(model.hur, value.mis = FALSE)
##                mean     sd     2.5%    97.5% Rhat n.eff
## alpha[1]      0.364  0.090    0.189    0.540 1.01  1000
## alpha[2]      0.829  0.135    0.551    1.078 1.01   320
## beta[1]       5.093  0.089    4.931    5.275 1.01   250
## beta[2]       5.241  0.106    5.046    5.448 1.00  1000
## beta_f[1]    -0.313  0.237   -0.750    0.175 1.00  1000
## beta_f[2]     0.216  0.240   -0.225    0.697 1.00  1000
## deviance   1088.878  5.440 1080.779 1101.478 1.00  1000
## gamma_c[1]   -1.966  0.261   -2.480   -1.466 1.00   610
## gamma_c[2]   -1.922  0.336   -2.643   -1.288 1.00  1000
## gamma_e[1]   -0.801  0.185   -1.163   -0.445 1.00  1000
## gamma_e[2]   -0.888  0.212   -1.320   -0.476 1.00   860
## mu_c[1]     143.026 13.807  119.852  172.187 1.00   450
## mu_c[2]     164.920 19.531  132.627  206.370 1.00  1000
## mu_e[1]       0.718  0.022    0.675    0.760 1.00  1000
## mu_e[2]       0.785  0.024    0.737    0.828 1.01   240
## p_c[1]        0.126  0.028    0.077    0.188 1.00   590
## p_c[2]        0.132  0.038    0.066    0.216 1.00  1000
```

```
## p_e[1]        0.311  0.039    0.238    0.390 1.00  1000
## p_e[2]        0.293  0.043    0.211    0.383 1.00   910
## s_c[1]      139.035 14.571  113.228  168.837 1.02    99
## s_c[2]      169.667 23.720  133.384  219.426 1.00   290
## s_e[1]        0.227  0.011    0.206    0.252 1.00  1000
## s_e[2]        0.253  0.017    0.223    0.288 1.00  1000
```

8. 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**. These are: 1) *Deviance information criterion* (DIC; Spiegelhalter, Best, Carlin, and van der Linde 2002); 2) *Widely applicable information criterion* (WAIC; Watanable 2010); 3) *Leave-one-out information criterion* (LOOIC; Vehtari *et al.* 2016).

The three methods estimate the predictive accuracy of the model 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 the details about how these measures are constructed or their theoretical justifications as these are outside the scope of this article. For a comprehensive analysis of the different types of predictive information criteria we refer to Vehtari, Gelman, and Gabry (2014).

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 compute the three types of information criteria using 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 of the model. 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 outcomes (`"both"`).

The output obtained from the function `pic` is a list which elements vary according to the information criterion specified. In this example, we choose the 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.

3. `dic` is the DIC, computed as `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 criteria, type `help(pic)`. The actual DIC value for the model stored in `model.hur` can be accessed by typing:

```
R> pic.dic$dic
## [1] 1036.273
```

9. 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 via the functions `selection`, `pattern` and `hurdle`. The function `diagnostic` allows the user to choose among a set of predefined diagnostic plots through the argument `type`. For example, we can plot the histograms of 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 4.

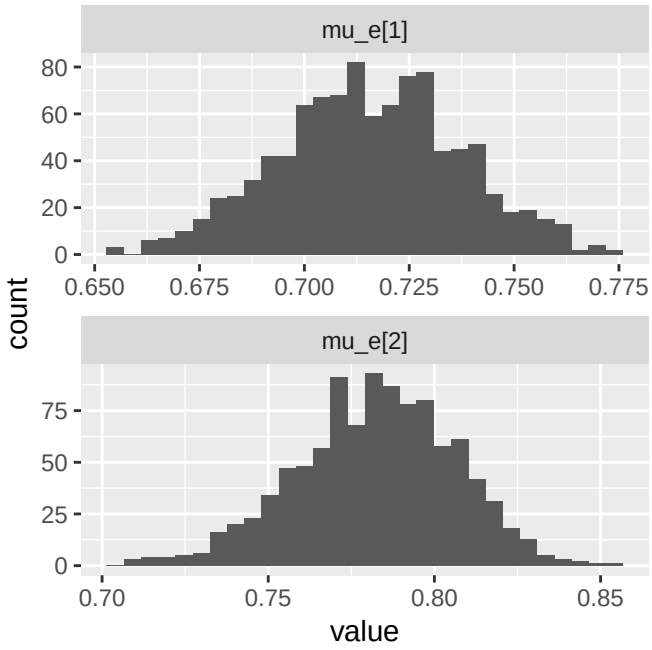


Figure 4: Histograms of the posterior distributions for the mean effectiveness parameters in the two groups

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, pattern mixture 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 predefined 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 9 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, trace_eff, ncol = 2)
```

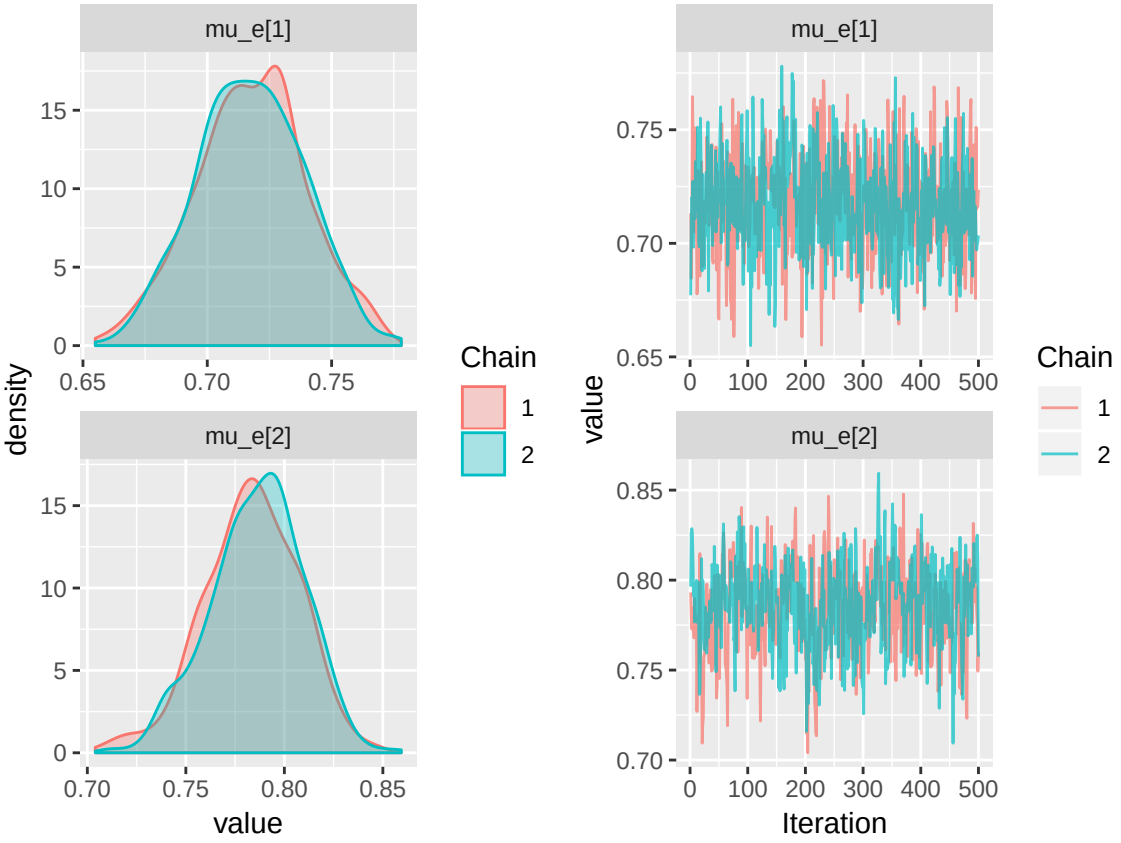


Figure 5: Density and trace plots of the posterior distribution of the mean effectiveness in the two groups

10. 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 6

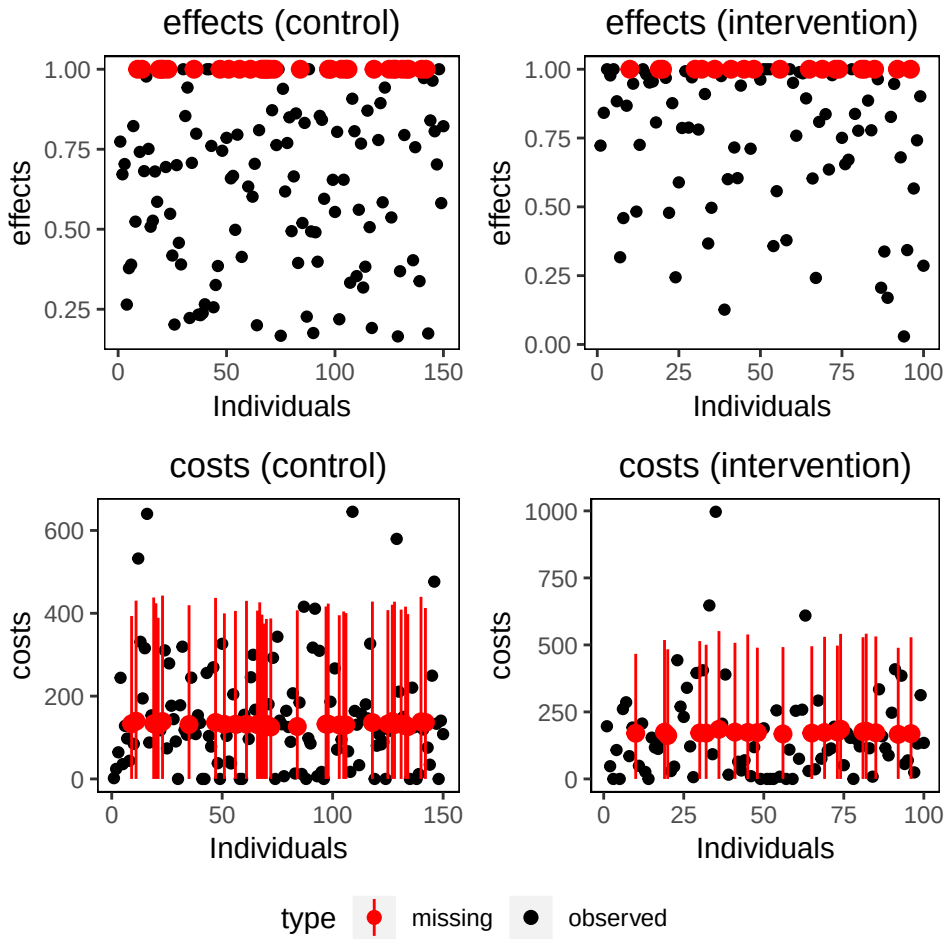


Figure 6: Scatter plots of the observed (black dots) and imputed (red dots and lines) values for both outcomes in the control and intervention groups. Imputation uncertainty is represented in terms of posterior means and 95% credible intervals.

The only compulsory argument to provide to the function 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 associated with lines representing their posterior credible intervals. By default these are the 95% CI but they

can be modified by changing the values for the upper and lower bounds of the credible intervals using the `prob` argument in the function `selection`, `pattern` 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 prespecified themes similarly to what shown for the function `diagnostic`.

11. Economic evaluation

Results from the economic evaluation from a model implemented in **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   0.718  0.022   0.681   0.754
## mean costs   143.026 13.807 122.607 166.079
##
##  Reference intervention
##                 mean     sd      LB     UB
## mean effects  0.785  0.024   0.743  0.821
## mean costs   164.92 19.531 135.475 198.22
##
##  Incremental results
##                   mean     sd      LB     UB
## delta effects   0.067  0.032   0.016  0.121
## delta costs    21.894 23.909 -16.931 61.601
## ICER          325.679
```

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 and are named `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** to summarise 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 7

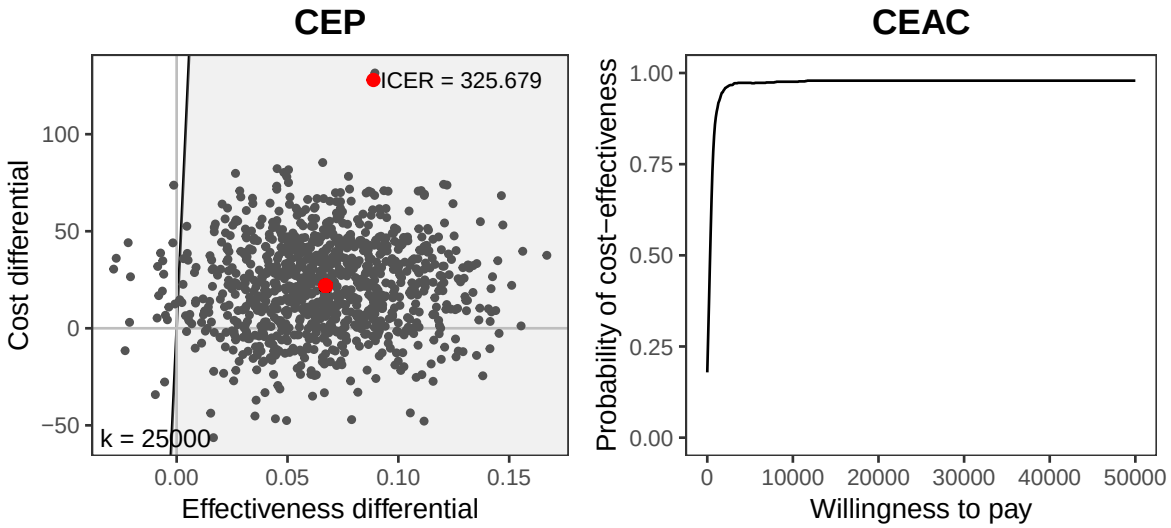


Figure 7: Cost effectiveness plane (CEP) and cost effectiveness acceptability curve (CEAC) obtained using the functions `ceplane.plot` and `ceac.plot` in the package **BCEA** and applied to the model results contained in the object `model.hur$cea`

12. Conclusions

The **missingHE** package can be used to handle missing outcome data in economic evaluations for two arms trial data. The package implements different approaches to address missingness under alternative missing data mechanism assumptions and provides a set of parametric

models to jointly handle the typical complexities which affect the effectiveness and costs (e.g. correlation, skewness and structural values).

Missing data represent a serious threat for economic evaluations as, when confronted with a partially observed dataset, each analysis makes assumptions about the missing values that cannot be verified from the data at hand. Therefore, 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 alternative 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/.

B, Leurent, M, Gomes, Carpenter J (2018). "Missing data in trial-based cost-effectiveness analysis: An incomplete journey." *Health Economics*, pp. 1024–1040.

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.

Diaz-Ordaz K, Kenward M, Cohen A, Coleman C, Eldridge S (2014). "Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines." *Clinical Trials*, **11**, 590–600.

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.

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**.

Little R, Rubin D (2002). *Statistical Analysis with Missing Data, Second Edition*. John Wiley & Sons, New York.

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/.

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