A Literature Review to Assess Methods for Handling Missing Data within the Health Economic Evaluation of Clinical Trials

Juan Jose Garcia Sanchez[1], Oliver Darlington[2], Rebecca Boyce[2], Mario Ouwens[3], Phil McEwan[2], Dan Jackson[1]

[1] AstraZeneca, Cambridge, UK, AstraZeneca; [2] Health Economics and Outcomes Research Ltd, Cardiff, UK, [3] AstraZeneca, KC4, Gothenburg, Sweden

POSC307

Introduction

Data missingness is a common issue in economic evaluations and can lead to substantial bias if not accounted for.[1,2]

Despite this, there is limited guidance on how to address missing data in economic evaluations, and methods are not commonly reported.

Missing data in clinical trials is often due to patient drop out,[2] which can often occur at differential rates between trial arms.

The study objective was to conduct a literature review to identify published literature describing methods of handling missing data in economic evaluations of clinical trials.

Methods

A targeted literature review designed according to the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (**PRISMA-P**) checklist was undertaken to identify publications describing methods for handling missing data in health economic evaluations.[3]

Electronic databases (Medline and Medline In-Process) were searched, according to a pre-defined search protocol, from inception to October 2020.

Abstracts were initially reviewed, followed by full texts, to determine study eligibility, based on predefined Population-Intervention-Comparators Outcomes-Study (**PICOS**) criteria, to include health economic evaluations of trial data or any systematic literature review considering methods for handling missing data.

Data from eligible studies were extracted and reported methods for handling missing data were summarised.

Results

Literature review

The review identified 44 publications meeting the eligibility criteria (**Figure 1**).

Mechanisms of missing data

The nature of the missing data can determine the most appropriate method for handling it.

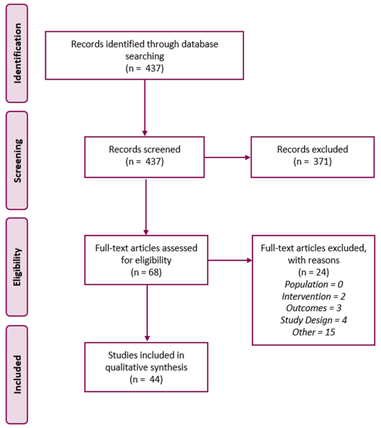

Figure 1. PRISMA flow diagram

Missing data in clinical trials can be classified into three mechanisms:

1. Missing completely at random (**MCAR**) - missing data is not linked to the analysis question.
2. Missing at random (**MAR**) - the probability of missing data is dependent upon observed data and independent of the true value.
3. Missing not at random (**MNAR**) - the probability of missing data is dependent on its true value.

Methods for handling missing data

The most common methods identified to address missing data are described in **Table 2**.

Multiple imputation (**MI**) with chained equations (**MI-MICE**) was the most common approach when using MI (**Figure 2A**), and complete-case analysis (**CCA**) was typically compared against more robust methods in sensitivity analysis.

Inverse probability weighting (**IPW**, **Figure 2B**) was used in four studies but was found to be outperformed by MI methods due both to ease of implementation and inclusion of variables which are predictive of missingness.

All approaches produced unbiased results when data are MCAR.[4]

Table 2. Methods to address missing data (in order of simplicity)

Methods (no. studies)	Summary	Use with:	Advantages	Disadvantages
Complete Case Analysis (CCA) (n=14)	Data from patients with missing data are excluded from the analysis.	MCAR	CCA is a simple solution when data is MCAR.	CCA is only valid under MCAR assumption as the results could be biased if missing data is linked to treatment.
Single Imputation (SI) (n=8)	The missing value is replaced with an imputed value. Last observation carried forward (LOCF), where the last data point is used for missing values, was most common (n=4).	MCAR MAR	More robust method than CCA and can be used under MAR assumptions.	SI methods have a limited ability to address systematic bias associated with missing data.
Multiple Imputation (MI) (n=27)	Missing values are estimated by creating several imputed datasets based on the distribution of the missing values. MI-MICE (14 studies) was most common.	MCAR MAR MNAR	MI builds on SI and produces estimates which are closer to the true values.	MI could produce inaccurate results if there is mis-specification of the model, even under MAR. Bias can occur with high levels of missing data and under the MNAR assumption.
Inverse probability weighting (IPW) (n=4)	Weights are defined as the reciprocal of the estimated probability of being observed.	MCAR MAR MNAR	Greater weight on observed values of similar patients to those with missing data.	Frequently outperformed by MI when data is MCAR. Unlikely to account for MNAR data appropriately.

MI: multiple imputation, CCA: complete-case analysis, PS: propensity scoring, IPW: inverse probability weighting, LOCF: last observation carried forward; SI: single imputation.

When data are MAR, MI and IPW is more precise than CCA.

When data are MNAR, pattern-mixture models and Bayesian approaches are the best methods to handle missingness.[5,6]

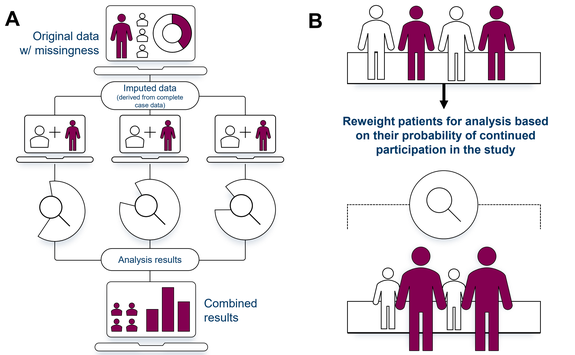

Figure 2. Missing data methods: A) multiple imputation B) inverse probability weighting

However, any method to handle missing data under MNAR will be associated with uncertainty and no single method was found to be robust under the MNAR assumption.[7,8]

Conclusions

Robust handling of missing data is essential as the choice of methods can have a direct impact on the conclusions drawn.

There is little difference in results when using MI or CCA under MCAR data.

MI or IPW should be used where data are MAR and compared with results under an assumption of missing not at random, to evaluate the effect of differing assumptions on cost-effectiveness conclusions.

Pattern-mixture models and Bayesian approaches can be used when data is MNAR, however methods rely on untestable assumptions, and often produce less precise estimates with greater bias.

References

1. Leurent B, Gomes M, Carpenter JR. Health Economics. 2018;27(6):1024-40.
2. Marvida D, White IR. Res Synth Methods. 2020;11(1):2-13.
3. Moher D, et al. Syst Rev. 2015;4:1.
4. Beiger M, Harc JM, Reed C, et al. BMC Medical Research Methodology. 2016;16(1):83
5. Leurent B, Gomes M, Faria R, et al. PharmacoEconomics. 2018;36(8):889-901.
6. Mason AJ, Gomes M, Grieve R, et al. Health Economics. 2018;27(11):1670-83
7. Gomes M, Diaz-Ordaz K, Grieve R, et al. Medical Decision Making. 2013;33(8):1051-63.
8. Michalowsky B, Hoffmann W, Kennedy K, et al. The European Journal of Health Economics. 2020:1-12

Acknowledgements

This study was funded by AstraZeneca.

Supported by AstraZeneca

Presented online at ISPOR EU, 1-3 December 2021