

Dealing with missing data in randomized and cluster randomized trials

Citation for published version (APA):

Kayembe, M. T. (2021). *Dealing with missing data in randomized and cluster randomized trials*. [Doctoral Thesis, Maastricht University]. ProefschriftMaken. <https://doi.org/10.26481/dis.20211216mk>

Document status and date:

Published: 01/01/2021

DOI:

[10.26481/dis.20211216mk](https://doi.org/10.26481/dis.20211216mk)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Propositions of the thesis

“Dealing with missing data in randomized and cluster randomized trials”

Mutamba Tonton Kayembe

1. Statistical analysis methods used to handle missing data in randomized controlled trials (RCTs) or cluster randomized trials (CRTs), as in any study, should account for the underlying causes of the missing data, as failure to do so may lead to bias and/or to inefficient results.
2. For RCTs with a continuous outcome and missingness in one or more baseline covariate (s), all methods for handling covariate missingness considered in this thesis produce unbiased treatment effect estimates regardless of the missingness mechanism, provided all covariates are measured before randomization. In this case, simple methods for covariate missingness, such as overall mean imputation (ME) and the missing-indicator method (M) perform similarly to advanced methods such as multiple imputation (MI) and linear mixed effects models (LMM).
3. For RCTs and CRTs with a continuous outcome measured at baseline and after treatment and missingness occurring at both time points, bias in treatment effect estimation depends on outcome (and not on covariate) missingness for all missing data methods in this thesis, as long as all covariates are measured before randomization.
4. Multiple imputation (MI) is not necessarily the method of choice for handling missing data in RCTs, because there is always a simple alternative method that can perform at least as well as MI for treatment effect estimation when the covariates are measured before randomization.
5. Considering the practical importance of RCTs and CRTs in health research, appropriate handling of missing data in statistical analysis of RCTs and CRTs has a direct impact on science because of interest in unbiased and efficient estimation of treatment effects from such trials, and hence an indirect beneficial impact on society (in terms of individual's or public health safety) because of informed policy decision making based on uncompromised inferences from clinical trials.
6. “Randomization is too important to be left to chance.” (J. D. Petruccelli).
7. Statistical techniques for data analysis are best used, and the results thereof best interpreted, if the context in which they are being used is well understood.
8. Statistical inferences are very important for decision making in many scientific disciplines with societal impact. They are, however, not inherently superior to sound judgments and should, therefore, always be used with common sense for decision making.
9. In a society drowned in information but still starved for knowledge, statistics is particularly important because of its ability to fill the crucial gap between information and knowledge.
10. “Statisticians, like artists, have the bad habit of falling in love with their models.” (George Box).