Methods for dealing with missing outcome data in randomised controlled trials (RCTs) – a methodological scoping review

Missing outcome data is often handled poorly in RCTs, including cancer trials. To identify missing data methods, we undertook a scoping review. We found 75 reports on a variety of robust methods, but there were few descriptions of the application of these methods in RCTs. Increased uptake of robust methods will decrease bias in effect estimates and ensure fair representation of all participants.

Ellie Medcalf [1]
ellie.medcalf@sydney.edu.au
Robin Turner[2] David Espinoza[3] Katy JL Bell[1]
[1] *Sydney School of Public Health, The University of Sydney*
[2] *Biostatistics Centre, University of Otago*
[3] *NHMRC Clinical Trials Centre, The University of Sydney*

Background

Missing outcome data is often considered a greater issue in observational studies than in RCTs. However, missing outcome data can seriously compromise the validity of an intention-to-treat analysis in RCTs, including cancer trials. **Common methods** for handling missing outcome data in RCTs include complete case analysis (CCA) and multiple imputation (MI). However, depending on the **assumptions** made regarding the missing outcome data, these methods are not always appropriate to use.

Missing Data Assumptions

Missing Completely At Random is considered unrealistic in RCTs. Using **Missing At Random** is recommended when conducting the primary statistical analysis, while it is suggested that sensitivity analyses be conducted under **Missing Not At Random**. However, many RCTs do not adhere to these recommendations.

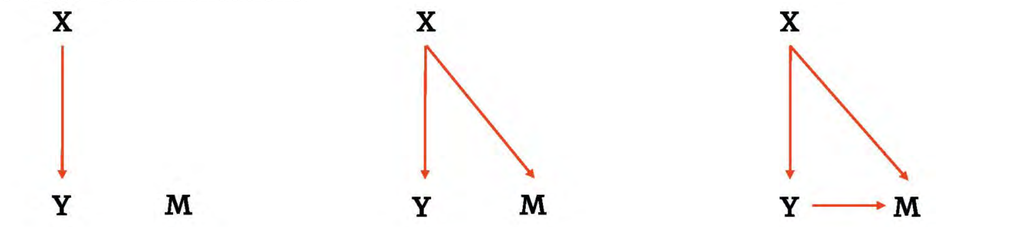

Figure 1: Missing data assumptions

Scoping Review

Given that many trials use inappropriate methods in the primary analysis and do not conduct suitable sensitivity analyses, we are conducting a **methodological scoping review** to identify methods for handling missing data in RCTs. This review will be used to inform the statistical analysis of the **MEL-SELF Trial of patient-led surveillance for the early detection of melanoma**, as well as assist other RCTs in selecting appropriate missing data methods.

Objective

Identify and summarise the current literature on analytical methods and/or frameworks that can be used to handle missing outcome data in RCTs.

Methods

Databases (MEDLINE, CINAHL, CENTRAL and EMBASE) searched from 2015 to present as previous reviews on missing outcome data methods have been conducted up to this year.

Papers will be included if:

- their primary objective was to discuss methods or theoretical approaches for handling missing outcome data in RCTs or in simulation studies motivated by an RCT design, or they are tutorial, guideline or theoretical papers discussing missing outcome data methods that are explicitly applicable to RCTs
- paper describes an application of the method or the pathway to application of the method is clear

Preliminary Findings

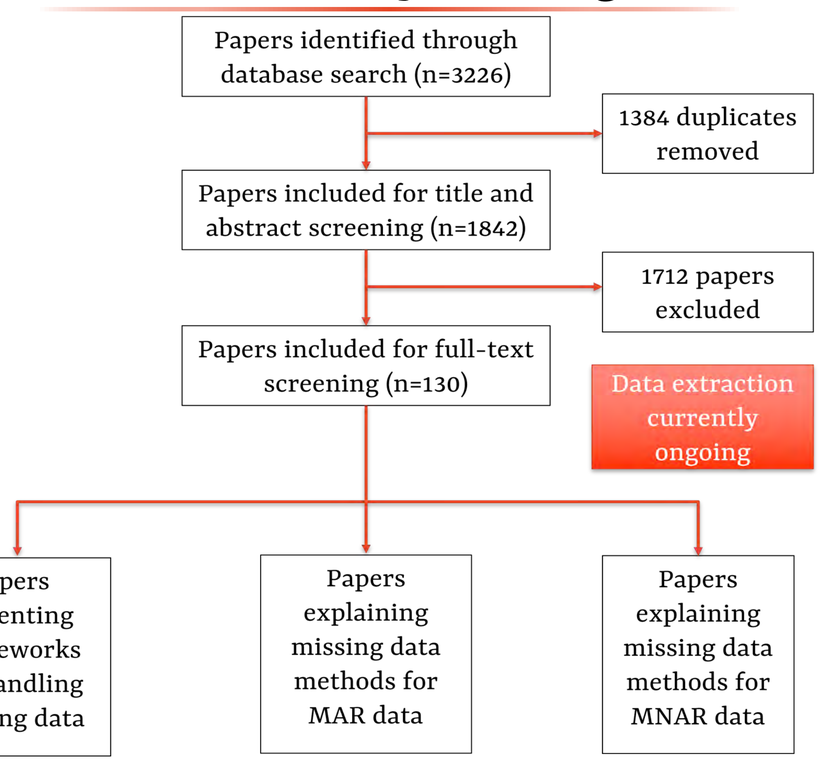

Figure 2: Flow diagram of scoping review process

Frameworks for handling missing data	• Elicitation frameworks – eliciting expert opinion • Machine learning framework
Methods for handling MAR data	• MI illustrated under different scenarios and software • Extensions to generalised estimating equations
Methods for handling MNAR data	• Controlled MI • Machine learning methods, e.g., random forest
Methods for handling MAR and/or MNAR data	• Maximum likelihood methods • Trimmed means

Table 1: Preliminary Findings

This project is supported by a National Health and Research Council (NHMRC) Investigator Grant 1174523. **Ellie Medcalf** is a recipient of a PhD Scholarship and received conference registration support from Sydney Cancer Partners via a grant from the Cancer Institute NSW.

Sydney Cancer Partners
Collaboration, Translation, Impact

NSW GOVERNMENT
Proudly supported by
Cancer Institute NSW

Please scan this QR code for further information on this research.