# IMPERIAL

# Introducing retrieved dropout reference-based centred multiple imputation for estimation of treatment policy strategies with missing data


Suzie Cro
James Carpenter
James Roger


## Introduction

Missing data from participants who withdraw from treatment early complicates trial analysis for a treatment policy estimand.

Retrieved dropout (RD) imputation fills in data based on a model for outcome on- and off-treatment, but this performs poorly with limited observed off-treatment data.

Reference-based imputation (RBI) replaces data based on a model from a specified reference group, but this disregards off-treatment data and makes strong assumptions.

We introduce a novel imputation method that combines these two approaches to form an extended model for multiple imputation.

## Novel Method

1. Choose a core RBI model (e.g. J2R) and a RD model

2. Parameterise an extended model as core RBI model plus additional parameters representing difference between RBI and RD model

3. Fit (2) in a Bayesian framework using uninformative priors for core model and mildly informative zero-centred priors for the additional parameters

4. Draw parameters and predict missing data based on patients' conditional distribution of post-deviation data given their pre-deviation data

5. Repeat (4) K times → K data sets

6. Fit model of interest to each K data set

7. Use Rubin’s rules for final inference

## Statistical Methods

We analytically explored expected bias and root mean square error (RMSE).

The novel method was applied to data sets based on an anti-depression trial varying priors for the additional parameters: 1) ‘covered’ lower missingness and standard RD implementable 2) ‘perforated’ with greater missingness and standard RD poor.

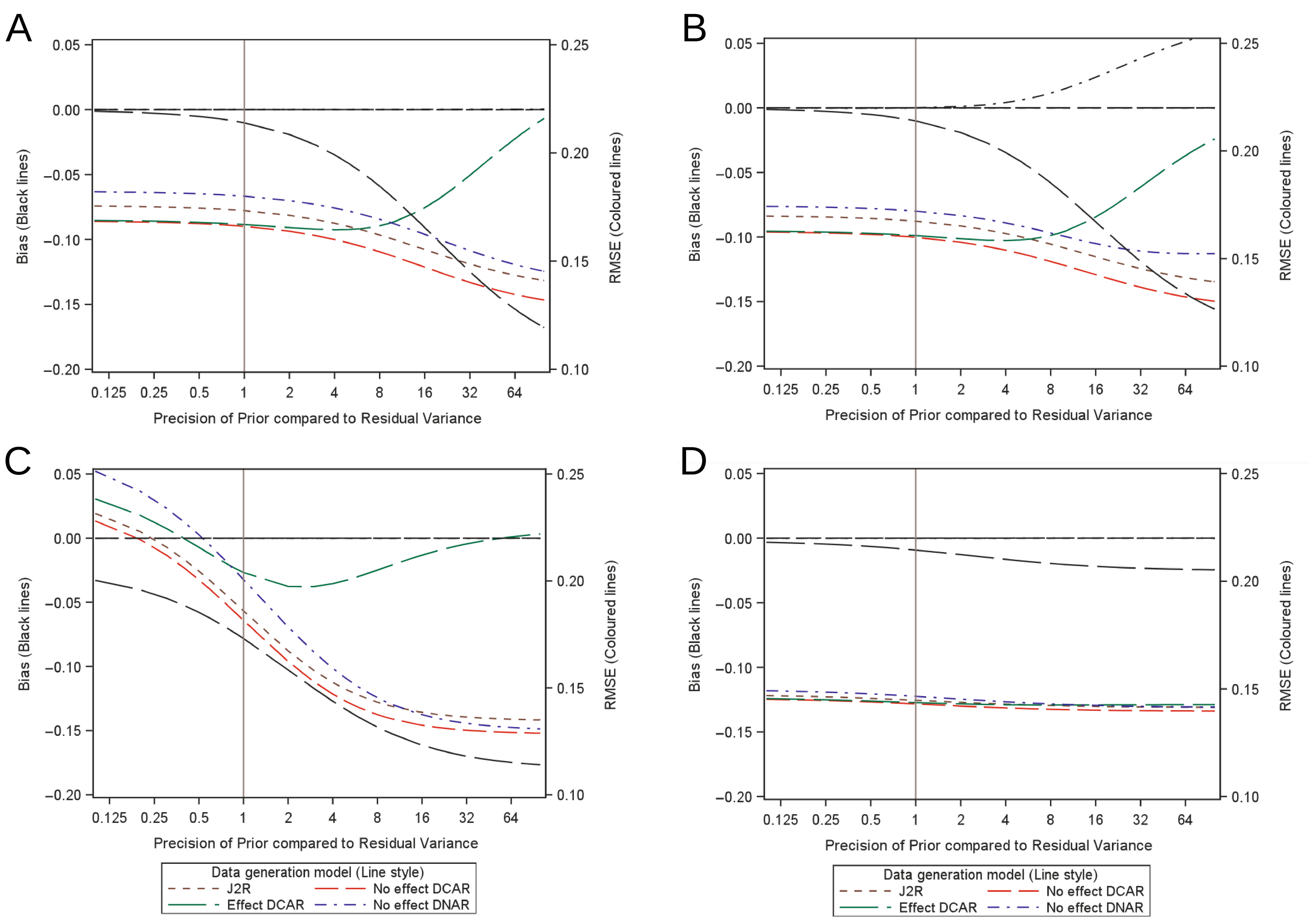


Fig. 1: Analytical bias and RMSE exploration with n=100 per arm and a single outcome for different data generation models and dispersion of prior for additional parameter. A=40% treatment withdrawal & 50% missing both arms. B=20% reference and 40% active treatment withdrawal & 50% missing both arms. C=20% treatment withdrawal & 90% missing both arms. D=5% treatment withdrawal & 50% missing both arms.

| Method | Covered dataset | | | | Perforated dataset | | | |
|---|---|---|---|---|---|---|---|---|
| | TE | SE | TE | SE | TE | SE | TE | SE |
| **Reference based MI** | | | | | | | | |
| -J2R | 2.18 | 1.13 | | | 2.17 | 1.13 | | |
| -J2R + observed off-treatment data | 2.28 | 1.05 | | | 2.39 | 1.05 | | |
| **Retrieved dropout MI** | Historic | | Current | | Historic | | Current | |
| -Outcomes missing for patient visits with non-estimable imputation parameters[a] | 2.32 | 1.10 | 2.34 | 1.07 | 2.45[a] | 1.08[a] | 2.74[a] | 1.05[a] |
| -Non-estimable imputation parameters set to zero[b] | 2.32 | 1.10 | 2.34 | 1.07 | 2.44[b] | 1.10[b] | 2.39[b] | 1.06[b] |
| -Priors for non-estimable imputation parameters using Bayesian MVN model[c] | 2.31 | 1.10 | 2.35 | 1.06 | 2.75[c] | 2.39[c] | 2.52[c] | 1.05[c] |
| **Retrieved dropout reference-based centred MI varying prior variance** | | | | | | | | |
| -Var=1 | 2.28 | 1.05 | 2.29 | 1.05 | 2.38 | 1.04 | 2.42 | 1.04 |
| -Var=10 | 2.31 | 1.06 | 2.33 | 1.06 | 2.40 | 1.08 | 2.49 | 1.05 |
| -Var=40 | 2.32 | 1.08 | 2.35 | 1.06 | 2.41 | 1.16 | 2.51 | 1.05 |
| -Var=160 | 2.32 | 1.09 | 2.36 | 1.06 | 2.44 | 1.45 | 2.52 | 1.05 |
| -Var=1000 | 2.32 | 1.10 | 2.36 | 1.06 | 2.63 | 2.08 | 2.52 | 1.05 |

Table 1 : Results of applying RBI with J2R, RD MI and novel method to two data sets inspired by anti-depression study. TE=mean treatment policy difference. SE = Standard error. For novel method J2R used as core RB model, RD model is either ‘historic’ (accounts for full compliance history, defined by treatment and last-on-treatment i.e. Pattern * Treatment * Visit*OffT means) or ‘current’ (accounts for compliance at each visit defined by treatment and being off treatment at visit i.e. Treatment * Visit*OffT means). RD models have convergence issues for the perforated data: a) Those patient visits which cannot be imputed are removed. Implemented using %MISTEP macro in SAS. b) In SAS Proc MI sets un- estimable parameters to zero before imputing . c) Prior N(0, 1000) used for all fixed effects as some parameters non- estimable, implemented using Proc BGLIMM in SAS

## Results:

*Bias and RMSE:*

When we trust the core model (i.e., additional parameters zero: J2R, No effect DCAR), the precision of the prior's variance has little impact on bias & RMSE (Fig.1:A,B).

If the core model is not true (Effect DCAR, No effect DNAR) the prior variance may want to be more conservative to not introduce bias (Fig.1:A,B).

When the amount of observed off-treatment data is small we expect the prior variance to have more impact (Fig.1:C).

With little missing data the impact of the prior variance will be small (Fig.1:D).

*Anti-depression trial:*

Increasing the prior variance for the additional parameters in the extended model increases the variance for the estimated treatment effect by a small amount for the perforated data set.

## Conclusions:

Reference-base centred multiple imputation provides a useful tool for estimation of treatment policy strategies with missing data.

More extensive simulation studies are required to further investigate the methods performance under different conditions.

For details see:

Imperial College London

ICTU
Imperial Clinical Trials Unit

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

UKRI

MRC Clinical Trials Unit

UCL

imperial.ac.uk