# Implementation of the Treatment Policy Strategy for Continuous Longitudinal Endpoints: A Comparison of Estimation Methods

*Thomas Drury – GSK*

*Estimation Workstream*
*Estimands Implementation Working Group (EIWG)*

*08th December 2022*

# Treatment Policy Estimation

## EIWG

- Talk is by **Estimation Workstream** of the **EIWG**

  - James Bell (Elderbrook Solutions)
  - Amel Besseghir (Clinchoice)
  - Christian Bressen Pipper (LEO Pharma)
  - Thomas Drury (GSK)
  - Lorenzo Guizzaro (EMA)
  - Tobias Muetze (Novartis)
  - Khadija Rantell (MHRA)
  - Marcel Wolbers (Roche)
  - David Wright (AstraZeneca)

  Disclaimer: The talk reflects the collective opinions of the members of the workstream and are not necessarily the views of our respective organisations


- Acknowledgements:

  - Ian White (UCL) – prior collaboration with GSK looking at multiple imputation models.
  - James Roger (LSHTM) – discussions about estimation of univariate vs. multivariate models.

# Treatment Policy Estimation

## What Is Needed?

- Treatment-policy includes intercurrent events (ICEs) within the treatment effect of interest
  - i.e. treatment changes (e.g. treatment discontinuation, use of rescue therapy) are part of the treatment regimens being compared.

- Its estimation requires continued data collection regardless of ICE occurrence
  - Nonetheless, missing data is almost inevitable

- Estimation of treatment policy in the presence of missing data is difficult
  - Treatment status within arms is heterogeneous (unlike other strategies)
  - ICEs highly correlated with subsequent missingness

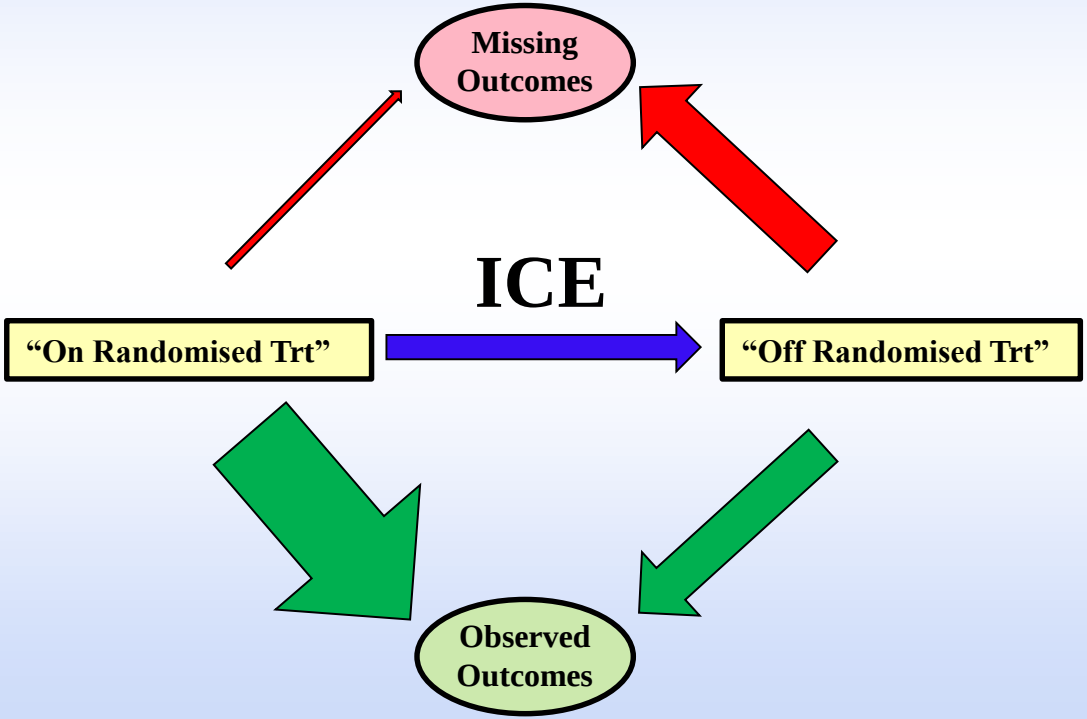
# Treatment Policy Estimation

## Missing Data

- Missing data in clinical trials is disproportionately “off randomised treatment” (off-trt)

- Observed patients are ‘different’ to unobserved patients
→ Complex missing data problem

- Conditioning on patients’ trt state to solve missingness issues is necessary (but not sufficient) for unbiased estimation



# Treatment Policy Estimation

## Aims

- Aims for simulation study:
  - Evaluate estimation methods for treatment policy handling in continuous outcome data
    - Use as realistic simulation as possible
    - Key criteria: bias and variability
  - Quantify impact of increasing missingness on estimation methods
    - How much missingness is too much?
    - When do methods ‘break’?
  - Investigate whether (restricted) maximum likelihood estimation (MLE) can be used
  - Investigate whether Jump-to-Reference is useful for treatment policy estimation
  - Compare both to Multiple Imputation methods – suggested by previous researchers.

# Treatment Policy Estimation

## Estimation Methods

| Approach | MI Model | MLE Model | Implied Clinical Assumption |
|---|---|---|---|
| **Standard** | **MI1** | **MMRM1** | No distinction between on- and off- treatment |
| **Retrieved dropout ('time independent')** | **MI2** | **MMRM2** | Immediate off-treatment effects |
| **Retrieved dropout ('time dependent')** | **MI3** | **MMRM3** | Time-dependent off-treatment effects |
| **Jump-to-reference** | **J2R** | - | Off-treatment assumed to be reference arm effect |

- Focus on retrieved dropout (RD) methods
- All models target same estimand: treatment policy for all ICEs

# Treatment Policy Estimation

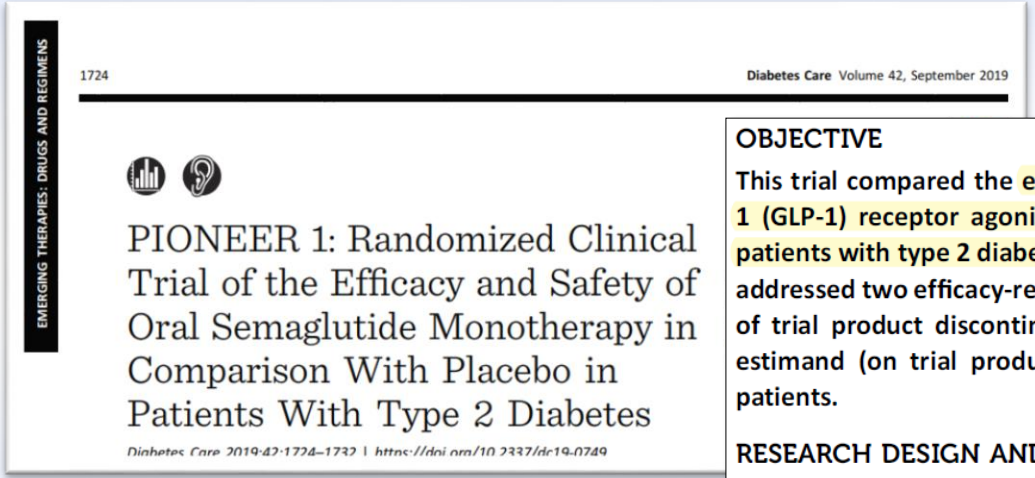
## Estimation Methods

| Approach | MI Estimation | MLE Estimation |
|---|---|---|
| **Standard ('1 models')** | • **No** off-rand-trt indicators<br>• **Standard (sequential regression) MI** | • **No** off-rand-trt indicators<br>• **Standard MMRM** |
| **Retrieved dropout ('time independent', '2 models')** | • **Binary** off-rand-trt indicators<br>• **Residual-based MI** | • **Binary** off-rand-trt indicators<br>• **off-trt*trt*visit interaction MMRM**<br>• Combine on-, off- rand-trt estimates<br>• Apply binomial variance correction |
| **Retrieved dropout ('time dependent', 3' models)** | • **Ordinal** off-rand-trt patterning<br>• **Standard MI with pattern covariates** | • **Ordinal** off-rand-trt patterning<br>• **pattern*trt*visit interaction MMRM**<br>• Combine on-, off-trt estimates<br>• Apply multinomial variance correction |
| **Jump-to-reference** | • **Binary** off-trt indicators<br>• **Residual-based MI**<br>• **Use control group as reference set** | N/A |

- For simulations where models '2' and '3' did not fit, a fixed step-down procedure was used:
  - '3' → 'minimally merged patterns 3' →'2' → '1'
- All estimation methods implemented in SAS, primarily using standard procedures (MIXED, MI, BGLIMM)

# Treatment Policy Estimation

## Motivation: PIONEER 1



EMERGING THERAPIES: DRUGS AND REGIMENS



PIONEER 1: Randomized Clinical Trial of the Efficacy and Safety of Oral Semaglutide Monotherapy in Comparison With Placebo in Patients With Type 2 Diabetes

Diabetes Care 2019;42:1724–1732 | https://doi.org/10.2337/dc19-0749


**OBJECTIVE**

This trial compared the efficacy and safety of the first oral glucagon-like peptide 1 (GLP-1) receptor agonist, oral semaglutide, as monotherapy with placebo in patients with type 2 diabetes managed by diet and exercise alone. Two estimands addressed two efficacy-related questions: a treatment policy estimand (regardless of trial product discontinuation or rescue medication use) and a trial product estimand (on trial product without rescue medication use) in all randomized patients.

**RESEARCH DESIGN AND METHODS**

This was a 26-week, phase 3a, randomized, double-blind, placebo-controlled, parallel-group trial conducted in 93 sites in nine countries. Adults with type 2 diabetes insufficiently controlled with diet and exercise were randomized (1:1:1:1) to once-daily oral semaglutide 3 mg, 7 mg, 14 mg, or placebo. The primary end point was change from baseline to week 26 in $HbA_{1c}$. The confirmatory secondary end point was change from baseline to week 26 in body weight.

# Treatment Policy Estimation

## PIONEER 1 Model

- Simulate longitudinal HbA1c through a multivariate normal distribution
  - Baseline and five post-baseline visits
- 400 patients randomized 1:1 between treatment and control
- Model parameters chosen to emulate PIONEER 1 trial
  - On-treatment means and off-treatment means
  - Variances
  - Intercurrent event rate
  - Visit schedule

# Treatment Policy Estimation

## PIONEER 1 Model

- **Single intercurrent event**
  - Emulating discontinuation of randomized treatment and simultaneous use of rescue medication
  - Intercurrent event occurrence conditional on prior outcomes & baseline
- **Two scenarios for post-discontinuation behavior**
  - Immediate change in HbA1c
  - Linear change in HbA1c
- **Missing data can only occur after the intercurrent event**
  - Conditional only upon discontinuation (and # visits since)
  - Monotone, with no intermediate missingness

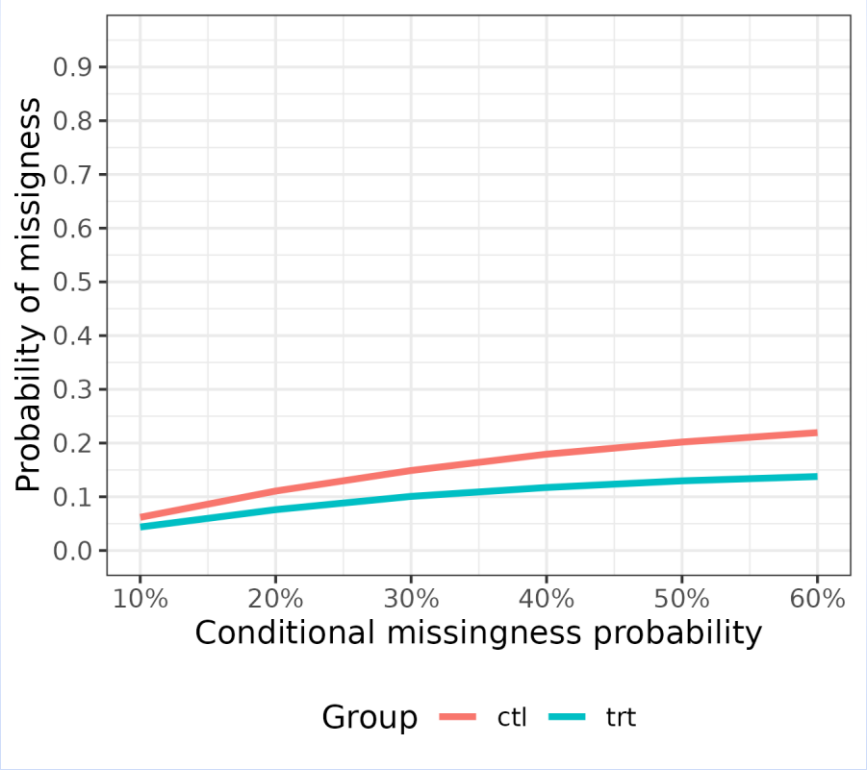
# Treatment Policy Estimation

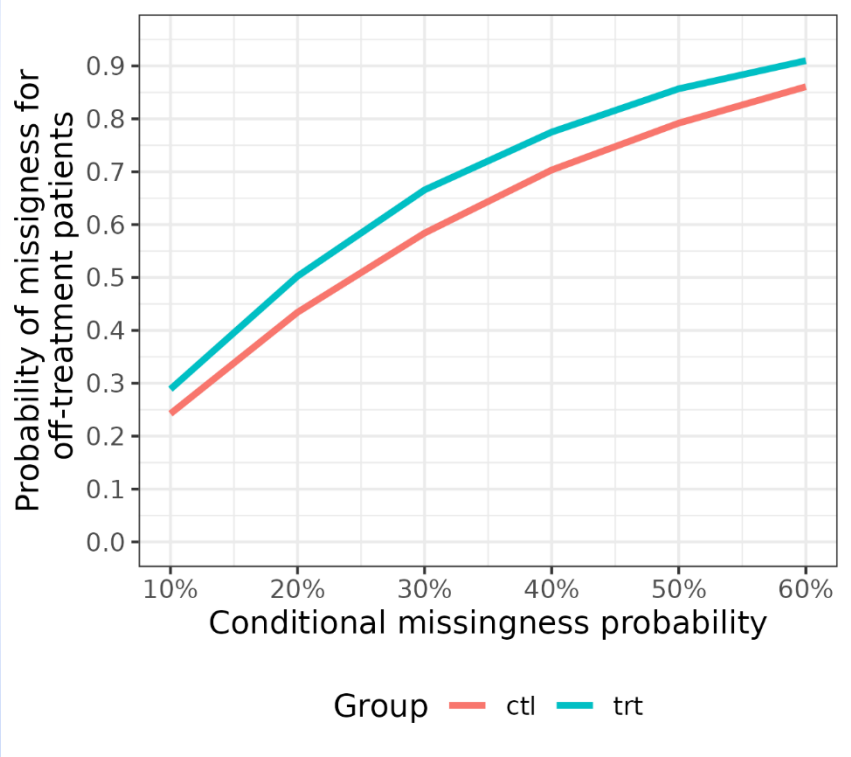
## Simulation Study

- For each setting, 1000 trials are simulated
- Here: focus on results for estimate $\hat{\theta}$ of mean difference between treatment and control **at final visit**
- Operating characteristics of interest:
  - **Relative bias**: $RelBias(\hat{\theta}, \theta) = \frac{E[\hat{\theta}]-\theta}{\theta}$
  - **Expected standard error**: $E[se(\hat{\theta})]$
- Varying Parameter: Probability of missing values at study visit $i$, conditional on having discontinued the randomized treatment and not having missing value prior to visit $i$

# Treatment Policy Estimation

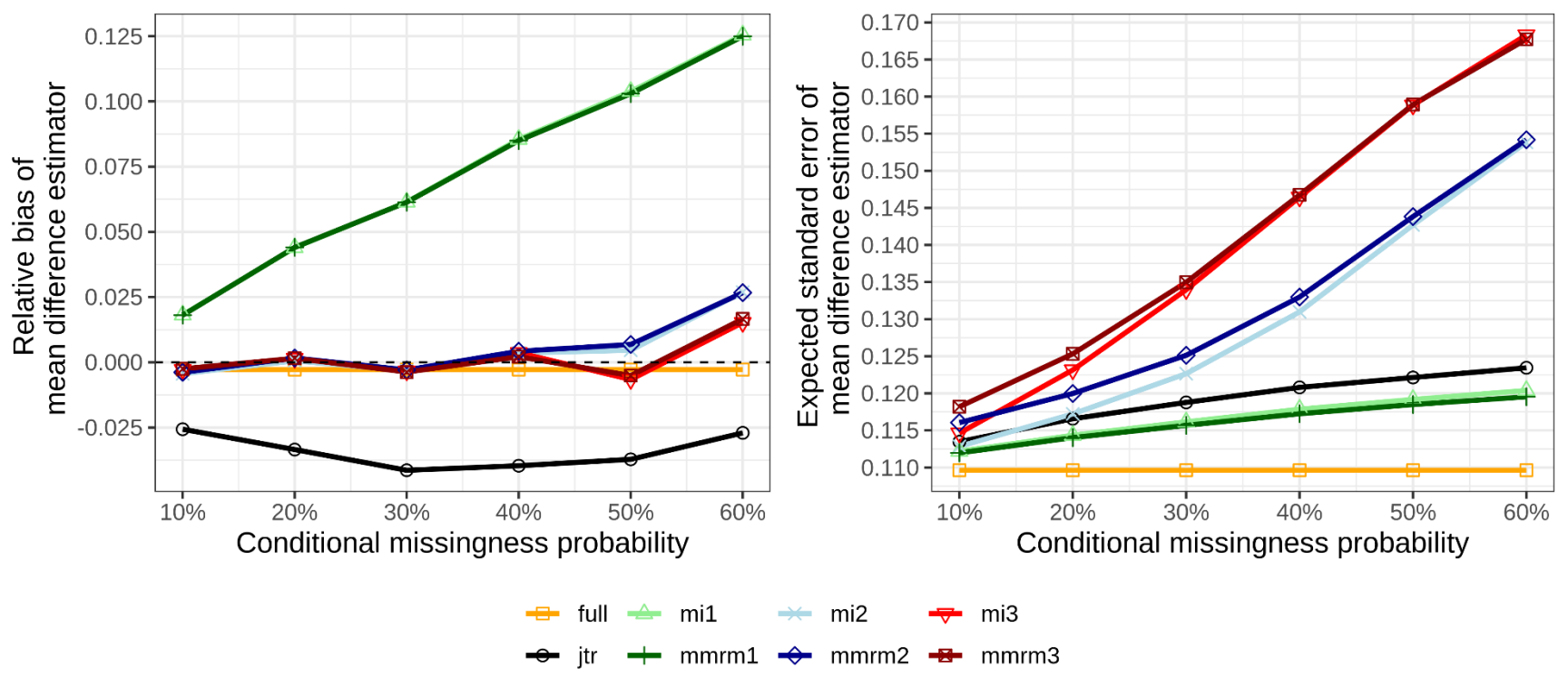
# Probability of Missingness



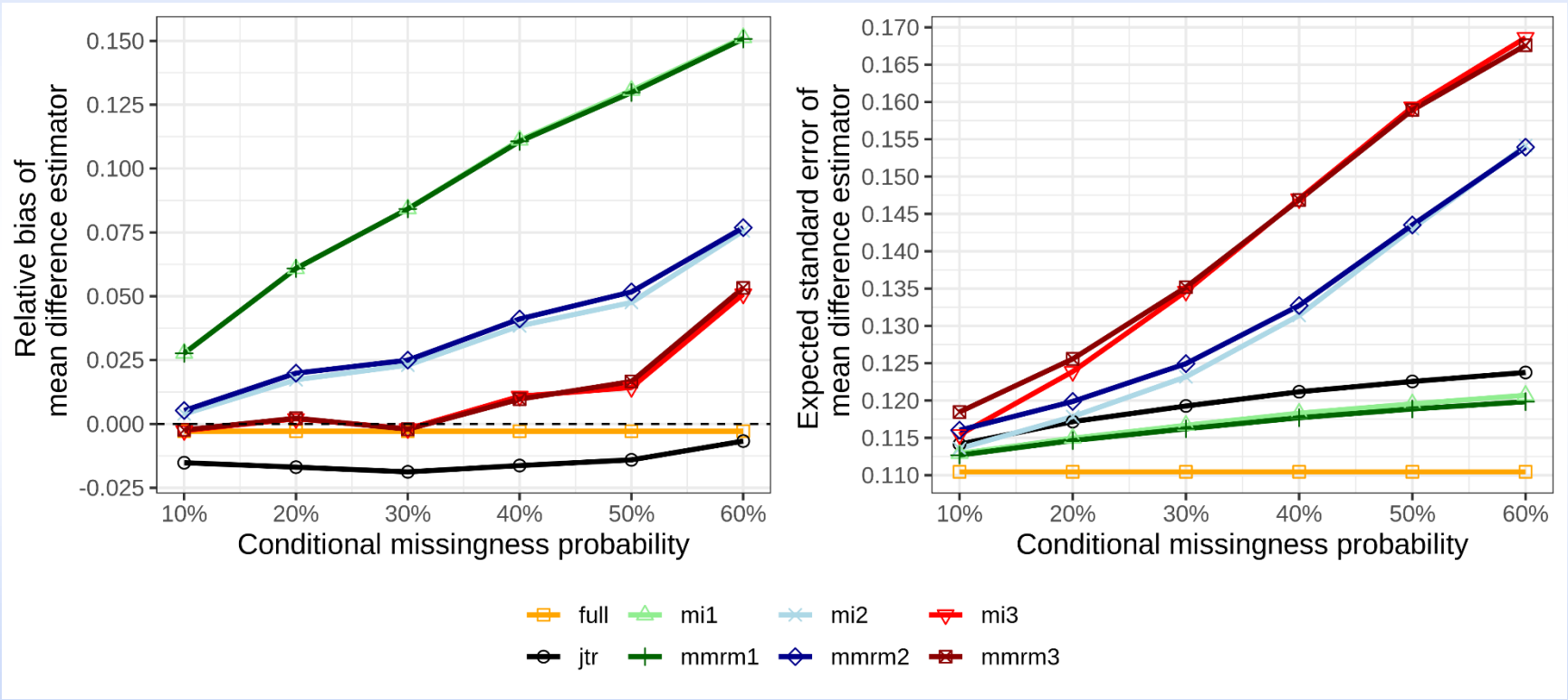
# Treatment Policy Estimation

## Bias & Variance: Immediate Change



# Treatment Policy Estimation

## Bias & Variance: Linear Change



# Treatment Policy Estimation — Recommendations

- **MI and MLE methods appear relatively interchangeable; pick the approach you prefer**
  - MLE logistical advantages? MI variance advantages?
- **Always have step-down / back-up approaches pre-specified for RD methods**
  - Simpler RD methods and/or reference based approaches
- **MMRM1/MI1 methods are not appropriate**; heavily biased unless minimal missing data
- **Use the simplest RD model appropriate for your trial** to minimise variance inflation
  - For straightforward fast-off symptomatic treatments, MMRM2/MI2 are sufficient
  - More complex off-treatment effects need MMRM3/MI3 (but awkward, high variance)
- **Sensitivity analysis is essential**: careful trade-offs between assumptions and statistical properties
- **Time-dependent covariates not problematic here** as MMRM2/3 sum effects across them

# Treatment Policy Estimation Conclusions

- **Treatment policy estimands are difficult to unbiasedly estimate**
- **MLE can replicate MI RD methods** using off-trt TDC interactions with treatment
- Simple, off-trt behaviour with low missingness → 'MMRM2'/'MI2' RD
- Complex off-trt behaviour with low missingness → 'MMRM3'/'MI3' RD
- Reference-based approaches probably best bias/variance trade-off when high missingness
- **Set up trial conduct procedures to collect as much post-ICE data as possible**
  - Statistical methods work best with 50%+ retrieval and start breaking down below 40%

Thank you for your attention!

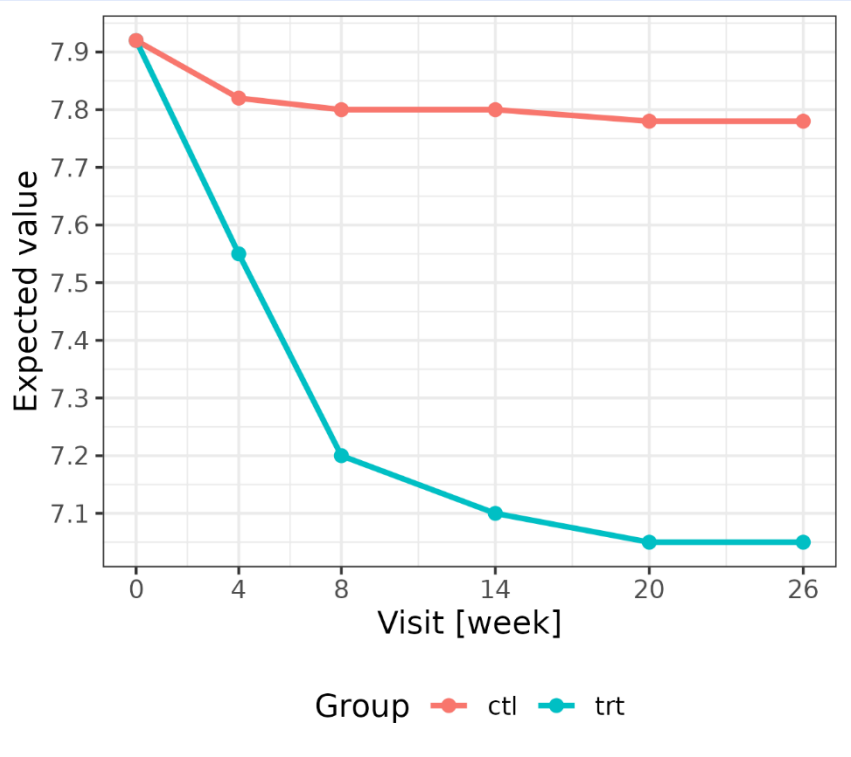
# Backup

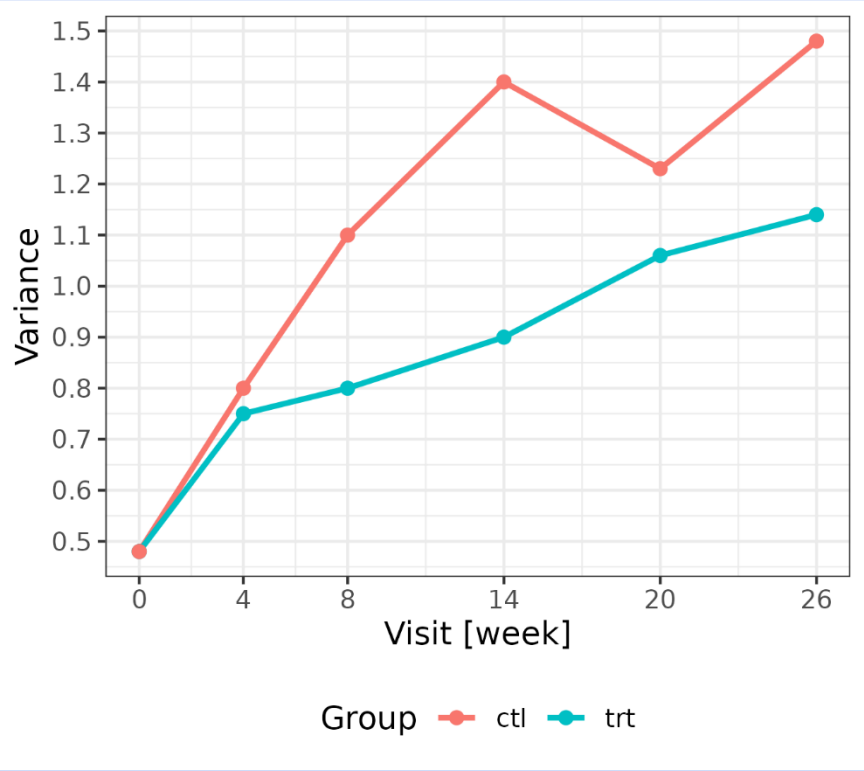
# Treatment Policy Estimation — Limitations

- Generally realistic trial data simulated, but:
  - Based on summary, not patient level, data
  - Single intercurrent event type, combining discontinuation and rescue
    - ‘Average effect’ modelled on real data
  - Simulated missingness is MCAR conditional on discontinuation
    - Does not extend to MAR, but most methods fail this easier test
  - Other than missingness amount and discontinuation mechanism, only one ‘scenario’
    - Expect generalisability of results, but cannot prove it

- Not simulated type I error yet
  - Not expected to be an issue if global null hypothesis is identical trajectories

# Mean and Variance under Randomized Treatment



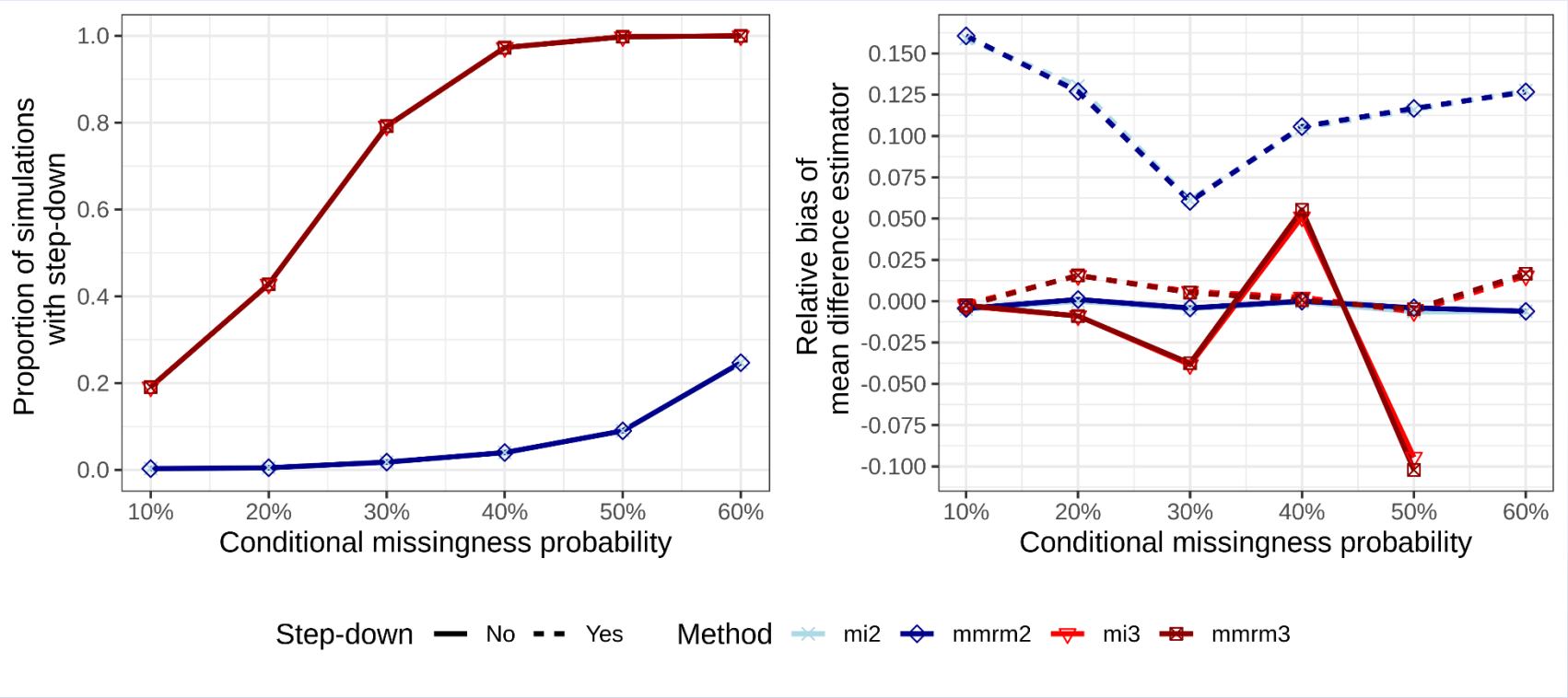
# Post Intercurrent Event Behavior

- Immediate change in HbA1c
  - For control group: drop of HbA1c by 0.6
  - For treatment group: drop of HbA1c by 0.2
  - Note that this represents an *improvement* upon going off-trt (high use of rescue therapy)

- Linear change in HbA1c
  - Let $x$ be the number of visits since intercurrent event
  - For control group: decrease of HbA1c by $\min(x, 3) \cdot 0.8$
  - For treatment group: decrease of HbA1c by $\min(x, 3) \cdot 0.25$

# Treatment Policy Estimation

## Step-down & Bias: Immediate Change



# Treatment Policy Estimation

## Step-down & Bias: Linear Change