The Estimands Academy for Trial Teams

"Bringing estimands to *life* through real case studies"

Webinar 9: Aligning Estimation with the Clinical Question of Interest

Monday 23rd November 2026 3-4:30 pm UK /4-5:30 pm CET/10-11:30 am EST

EFSPI European Federation of Statisticians in the Pharmaceutical Industry
Representing Statistical Associations in Europe

efpia European Federation of Pharmaceutical Industries and Associations

The Estimands Academy for Trial Teams Webinar Series

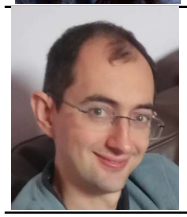
Available at (https://www.psiweb.org/vod/Index/ and @youtube/EFPIA)

Scan QR code for the EIWG training page with all links todate

EFSPI/EFPIA EIWG - Estimand Trainings

	Webinar Title
1	PIONEERing estimands in Clinical Research
2	Estimands in Oncology - How and Why
3	Estimands from trial planning to publication in medical journals: The ETHOS trial
4	Proposing Estimands from Different Perspectives
5	Estimands framework in action, the Alzheimer's disease case
6	Applying the Estimand Framework to Clinical Pharmacology Trials with a Case Study in Bioequivalence
7	Enhancing Clinical Study Reporting with the Estimand Framework
8	Making Estimands Work - Practical Approaches with Implementation in Data Standards

Introductions of Chairperson, Presenters and Q&A Panelists

	Name and job title	Company
	**Anna Robertson (Chairperson)** Statistical Lead	Pfizer
	**James Bell (Presenter, panelist)** Contract Methodology Statistician	elderbrook linking minds
	**Sue McKendrick (Presenter, panelist)** Statistical Science Director	PPD Part of Thermo Fisher Scientific; ThermoFisher SCIENTIFIC
	**David Wright (Panelist)** Head of Statistical Innovation	AstraZeneca

Disclaimer: The views expressed by the presenters and panelists do not necessarily represent the views of their employers or the entire Estimand Implementation Working Group (EIWG).

The Estimands Academy for Trial Teams



Please type your questions as we go in the **Q&A tab** (not the chat)



Aligning Estimation with the Clinical Question of Interest

Scan **QR code** for the Estimand Implementation Working Group (EIWG) **Training page** with links to all our **recordings.** This recording will be made available there too.

Acknowledgements - Sincere Thanks to:

- EFPIA/EFSPI for sponsoring and promoting the webinar.
- Estimand Implementation Working Group (EIWG) members for lively discussion and comments on the slides.
- Boehringer-Ingelheim Pharma GmbH for sponsoring James' work with EIWG
- Estimation Subteam of the EIWG (besides James Bell and David Wright):
  - Christian Bressen Pipper (Novo Nordisk)
  - Thomas Drury (GSK)
  - Lorenzo Guizzaro (EMA)
  - Oliver Keene (KeeneONStatistics)
  - Marian Mitroiu (Biogen)
  - Tobias Muetze (Novartis)
  - Khadija Rantell (MHRA)
  - Marcel Wolbers (Bayer)

Learning Outcomes

By the end of this webinar, participants should understand:

- **what alignment means** with respect to estimation of an estimand
- how different intercurrent event strategies determine **which observations are relevant** for estimation
- why we should consider **patient journeys** and the associated intercurrent events and missing-data patterns
- the **plausibility of different outcome trajectories** and relevance of different missing data approaches

Agenda

Introductions, Acknowledgements and Learning Outcomes

From Clinical Question to Aligning Estimation with Case Study Introduction

- What Alignment Means?

Challenges in Handling Missingness

- Assessing patient journeys, their intercurrent events and missing-data patterns

Which Data are Relevant for Different Strategies?

Treatment-policy Estimation in Practice

Practical Lessons for Trial Teams

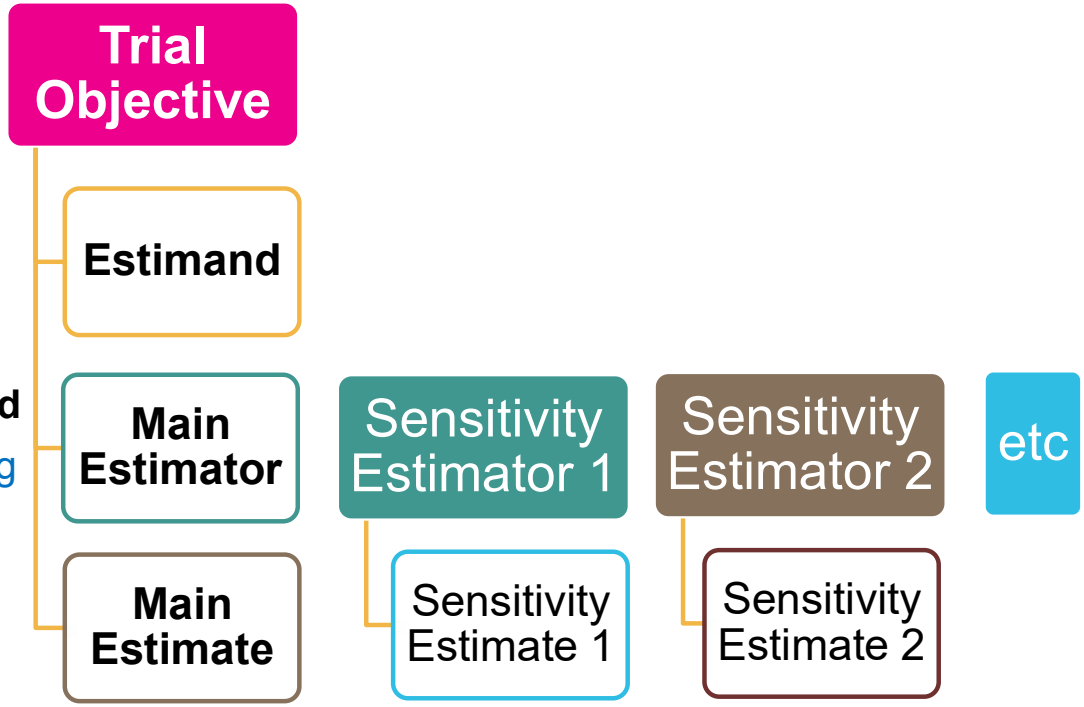
Q & A

The Estimand Framework – Estimation should Align to Estimands

Estimand = what (treatment effect)
we want to estimate

Estimator = analysis set + statistical method
includes censoring rules, missing data handling

Estimate = Numerical result



This talk will focus on the challenge of
aligning the estimation approach with the estimand

Biosimilar Case Study Example

Note: Regulations around biosimilars have changed recently – this is an example based on demonstrating **therapeutic equivalence** (assessing %change from baseline in bone mineral density versus margins +/- 1.45%).

This case study is presented for **illustrative purposes**.

In future development, equivalence based on PK/PD may be sufficient for regulatory approvals.

Start with the Target: What is your clinical question?

What will be your population-level summary,

between what treatments,

of what endpoint,

in which patient or healthy population,

incorporating strategies for intercurrent events?

Biosimilar Case Study Example

What will be the difference in means between

Biosimilar and Prolia (denosumab) (subcutaneous injection every 6 months) in

%change from baseline (%CfB) in lumbar spine bone mineral density (BMD) after 52 weeks in

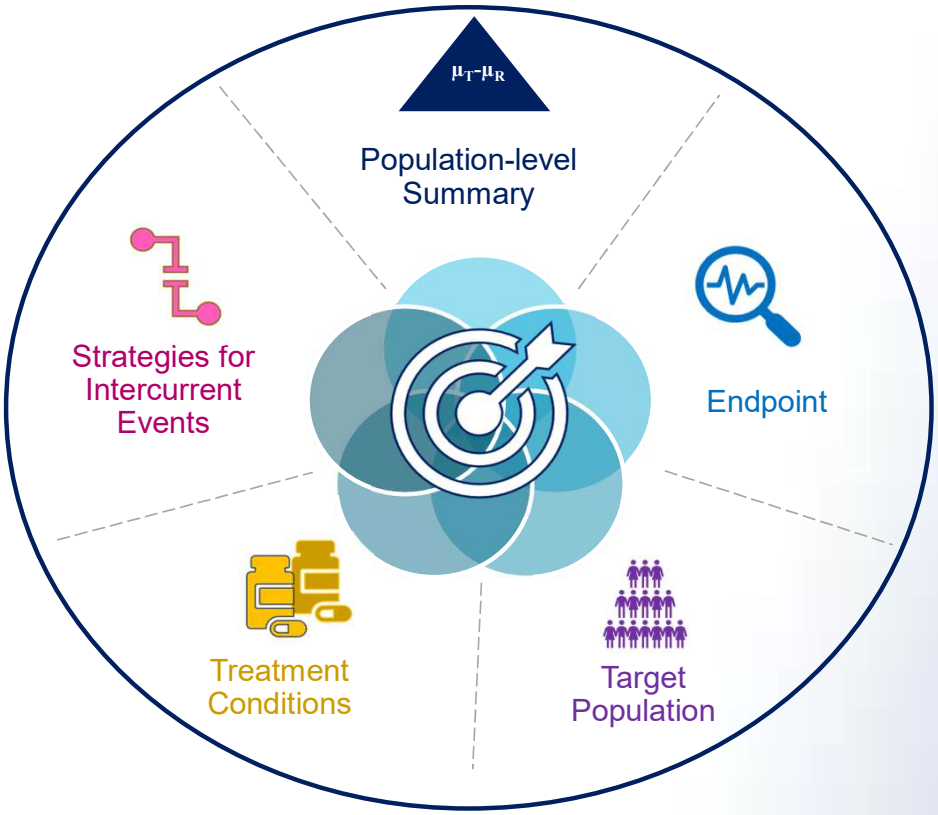
postmenopausal women with osteoporosis (treatment naïve)

using treatment policy strategy;

(including effects of treatment discontinuation, dosing deviations, and taking other osteoporosis medications)?

Alignment first requires a precise clinical question of interest

Align the Design and the Estimand to the Clinical Question



Thinking about the estimand attributes:-

- μ_T-μ_R — If you want a difference in means don't estimate a difference in medians!
- Treatment conditions are clinically relevant, including required reference product!
- Endpoint is clinically relevant, outcome of interest is well defined
- Study sample reflects the target population; disease confirmed? treatment assigned/dosed?
- Strategies for intercurrent events ---This can get tricky (see next slides) ---

Examples of What is or is not an Intercurrent Event

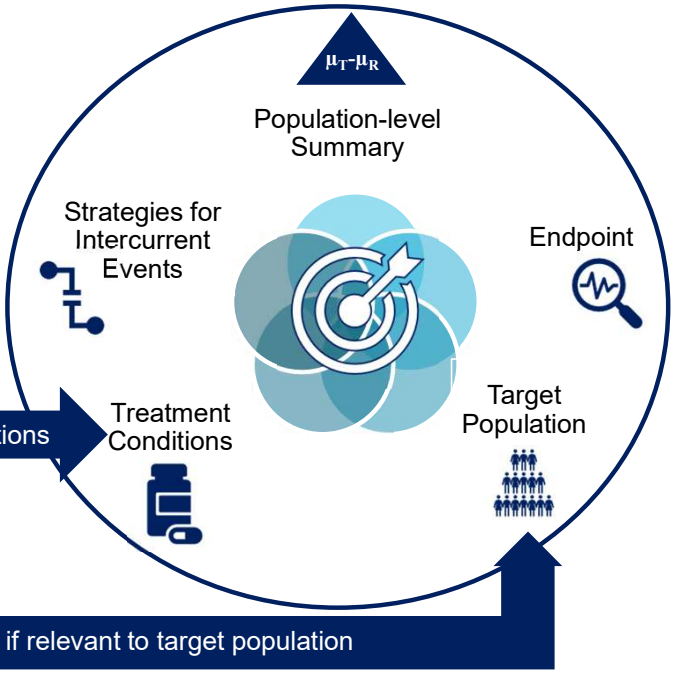
Intercurrent Event

- Treatment Discontinuation
- Rescue medication
- Unscheduled treatment switch or dose adjustment
- Terminal event (death or amputation)

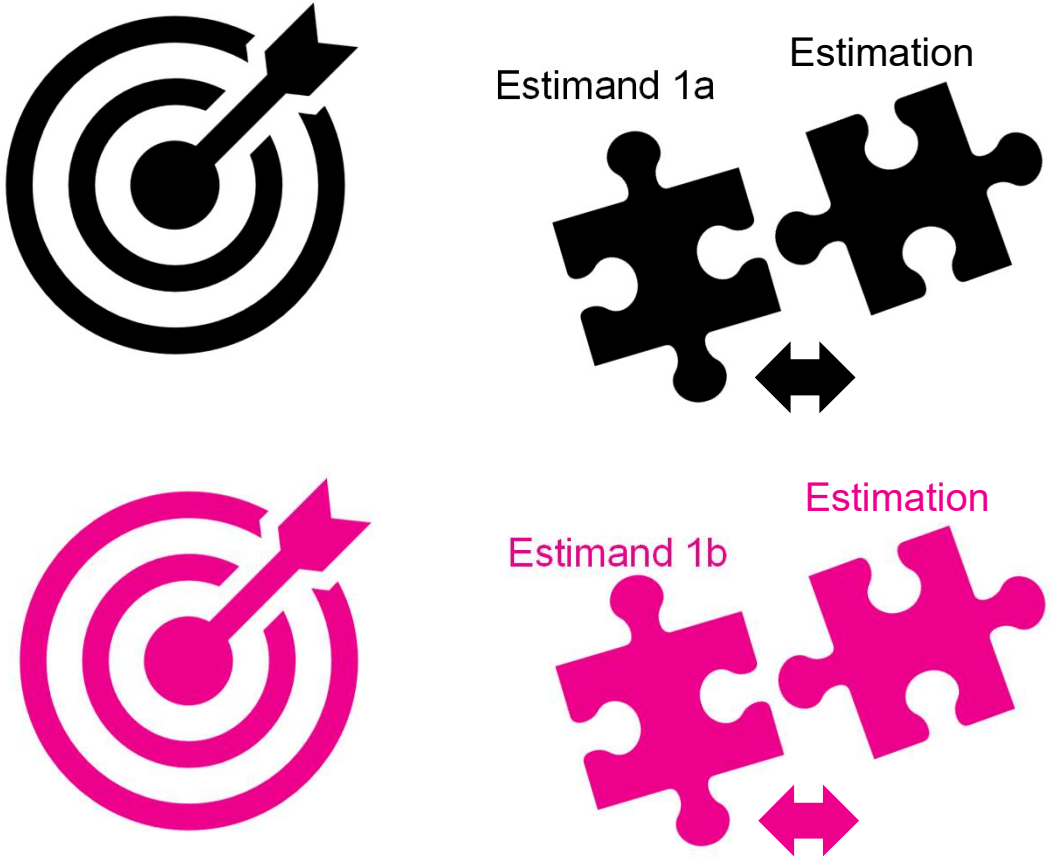
Not an Intercurrent Event

- Missing data issues
  - Missed visit *
  - Equipment failure *
  - Withdrawal from study
- Quality issues
  - Error in measurement
  - Timing deviation
- Scheduled dose titration – part of treatment regimen
- Event before treatment assignment

* **Missing data are sometimes linked with intercurrent events but they are distinct concepts.**



What does Alignment of Estimation Mean?



Alignment

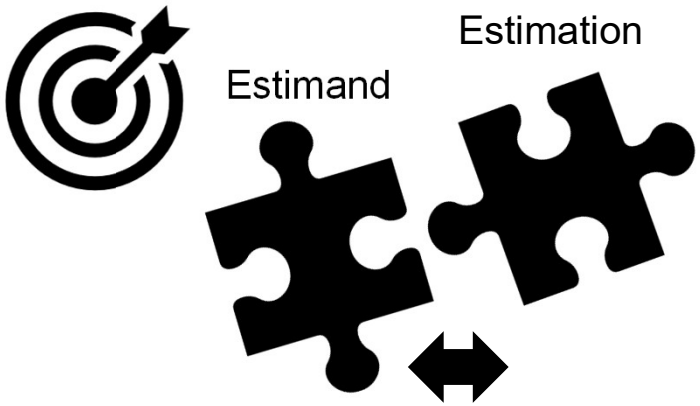
"[An aligned estimator]…

is able to provide an estimate on which **reliable interpretation** can be based" – ICH E9(R1)

- Trial design
- Data collection & retention
- Estimation & sensitivity methods

An unbiased estimator under clinically plausible assumptions

Practical Meaning of Alignment of Estimation





Different intercurrent event strategies produce **different targets**

Tricky Part of Alignment

"Estimation should reflect intercurrent event handling strategies" **Mallinckrodt et al. 2019**

- Missing data handling depends on strategy
- Relevance of data following intercurrent events differs

Alignment therefore comes back to two familiar questions:

- How to handle missing data?
- Which data are relevant?

Agenda

Introductions, Acknowledgements and Learning Outcomes

From Clinical Question to Aligning Estimation with Case Study Introduction

- What Alignment Means?

Challenges in Handling Missingness

- Assessing patient journeys, their intercurrent events and missing-data patterns

Which Data are Relevant for Different Strategies?

Treatment-policy Estimation in Practice

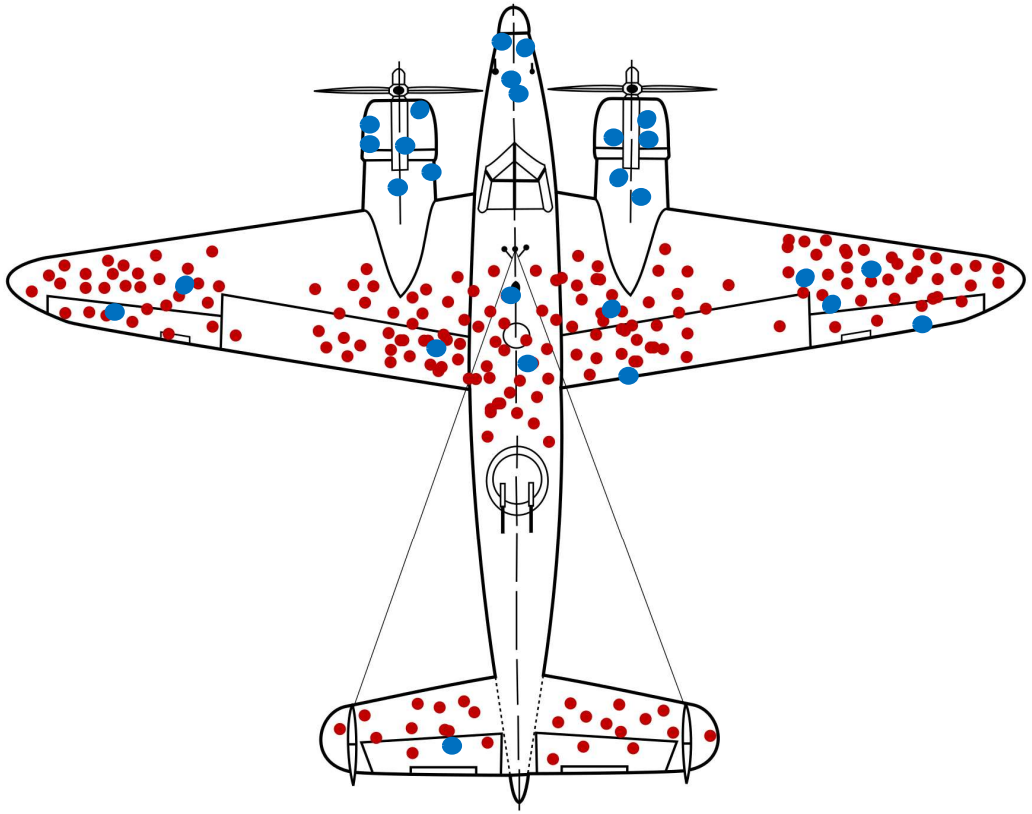
Practical Lessons for Trial Teams

Q & A

Thinking about the missing-data process

Red dots show bullet locations on a World War II aircraft that returned.

Blue dots represent missing aircraft: the bullets we would have seen on aircraft that did not return.



SOURCE: **Martin Grandjean / McGeddon / Cameron Moll** File:Survivorship-bias.svg - Wikimedia Commons

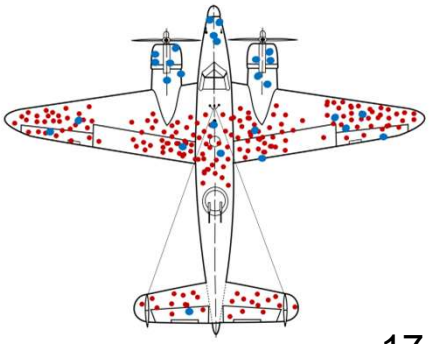
QUIZ Question: Thinking about the missing-data process

What does this example tell us about the missing-data process?

(A) The missing aircraft are a random sample of all aircraft

(B) Missingness can be ignored because we have many observed aircraft

(C) Whether an aircraft is observed depends on characteristics associated with its outcome



Red dots show bullet locations on a World War II aircraft that returned.
Blue dots represent missing aircraft: the bullets we would have seen on aircraft that did not return.

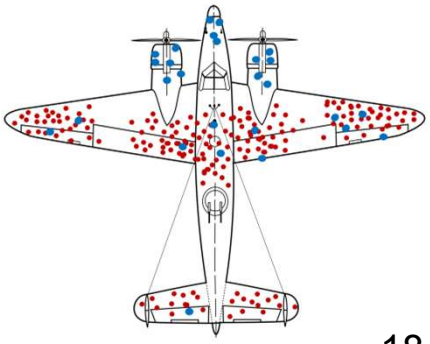
ANSWER QUIZ Question: Thinking about the missing-data process

What does this example tell us about the missing-data process?

(A) The missing aircraft are a random sample of all aircraft

(B) Missingness can be ignored because we have many observed aircraft

(C) Whether an aircraft is observed depends on characteristics associated with its outcome



Red dots show bullet locations on a World War II aircraft that returned.
Blue dots represent missing aircraft: the bullets we would have seen on aircraft that did not return.

Missing Data Definition

Missing data Data that would be meaningful for the analysis of a given estimand but were not collected. They should be distinguished from data that do not exist or data that are not considered meaningful because of an intercurrent event.

Missing data definition, ICH E9(R1) addendum

Sometimes data after intercurrent events are not relevant

Missingness mechanisms are assumptions about dependence

MCAR

Missing completely at random

Missingness is unrelated to observed or unobserved data (treatment, patient characteristics or outcomes)

MAR

Missing at random

Missingness is related to observed information but not to unobserved information

MNAR

Missing not at random

Missingness is related to unobserved outcomes even after accounting for observed information



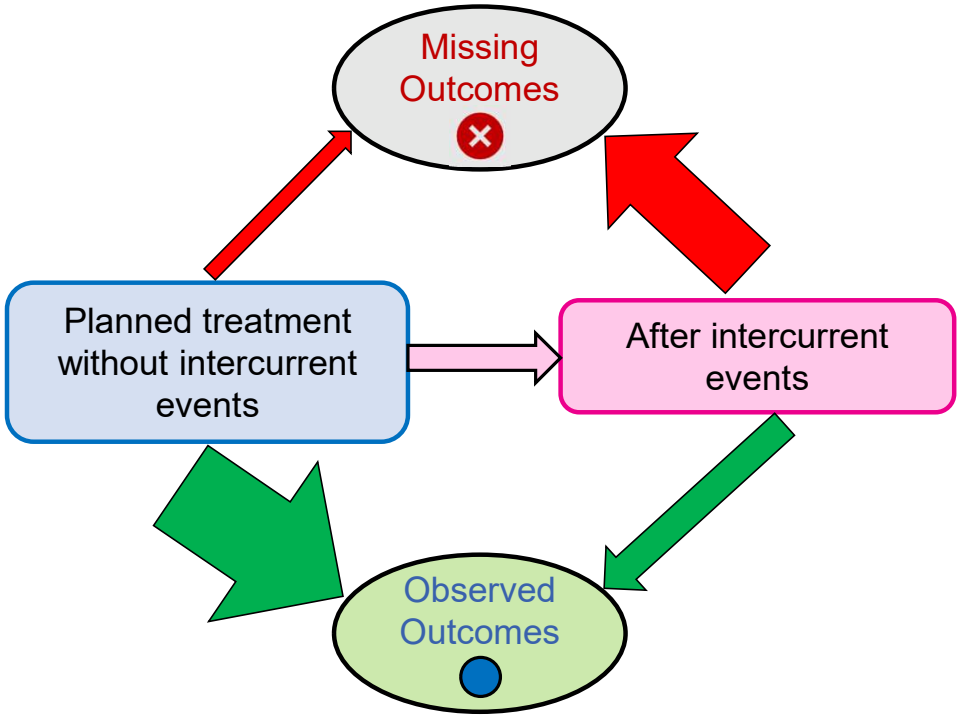
We need to make assumptions about how missingness depends on observed and unobserved information.

Intercurrent Events and Missingness

Missing data in clinical trials occurs disproportionately after intercurrent events

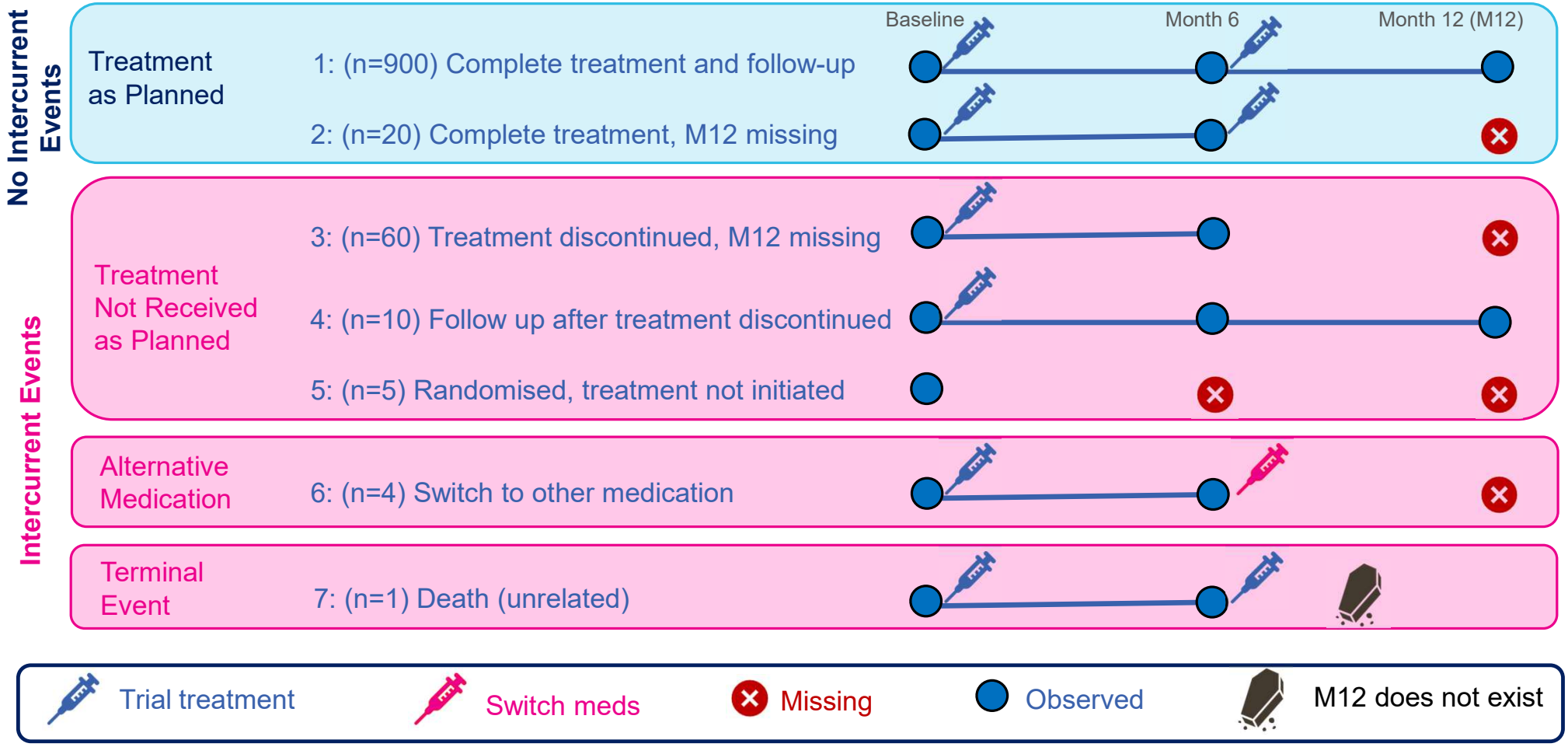
Unobserved patients may differ from those observed →
Complex missing data problem

Intercurrent events affect subsequent outcomes



Intercurrent events cause MNAR unless an analysis ‘knows’ about them

Illustrative Example of Different Patient Journeys (N=1000)



Disproportionate Missingness after Intercurrent Events

Same Month-12 target bone mineral density (BMD) endpoint. Very different missingness patterns.

Treatment proceeds as planned, without intercurrent events



920 participants

900 observed
Month-12 BMD observed

20 missing
Month-12 BMD missing

Only 2.2% (20/920) **missing**

Intercurrent Events

Dosing deviations

Treatment discontinuation

Use of other medications

80 participants

1 death
M12 outcome does not exist

10 observed
Month-12 BMD observed

69 missing
Month-12 BMD missing

86.3% (69/80) **missing**

Missingness is not evenly distributed — it is strongly associated with the patient journey.

Intercurrent events may affect both the subsequent outcome and the probability that we observe it.

Agenda

Introductions, Acknowledgements and Learning Outcomes

From Clinical Question to Aligning Estimation with Case Study Introduction

- What Alignment Means?

Challenges in Handling Missingness

- Assessing patient journeys, their intercurrent events and missing-data patterns

Which Data are Relevant for Different Strategies?

Treatment-policy Estimation in Practice

Practical Lessons for Trial Teams

Q & A

Treatment-policy Estimand requires Follow-up to Week-52

What will be the difference in means between

Biosimilar and Prolia (denosumab)* (subcutaneous injection every 6 months) in

%change from baseline (%CfB) in lumbar spine bone mineral density (BMD) after 52 weeks in

postmenopausal women with osteoporosis (treatment naïve)

* including effects of discontinuation of treatment, dosing deviations, and other osteoporosis medications?

Intercurrent Event Data
Analysis methods need information about intercurrent events

Week-52 BMD remains relevant after intercurrent events

Estimation of a Hypothetical Estimand

Unchanged 3 attributes

What will be the difference in means between μ_T-μ_R

Biosimilar and Prolia (denosumab) (subcutaneous injection every 6 months) in

%change from baseline (%CfB) in lumbar spine bone mineral density (BMD) after 52 weeks* in

postmenopausal women with osteoporosis (treatment naïve)

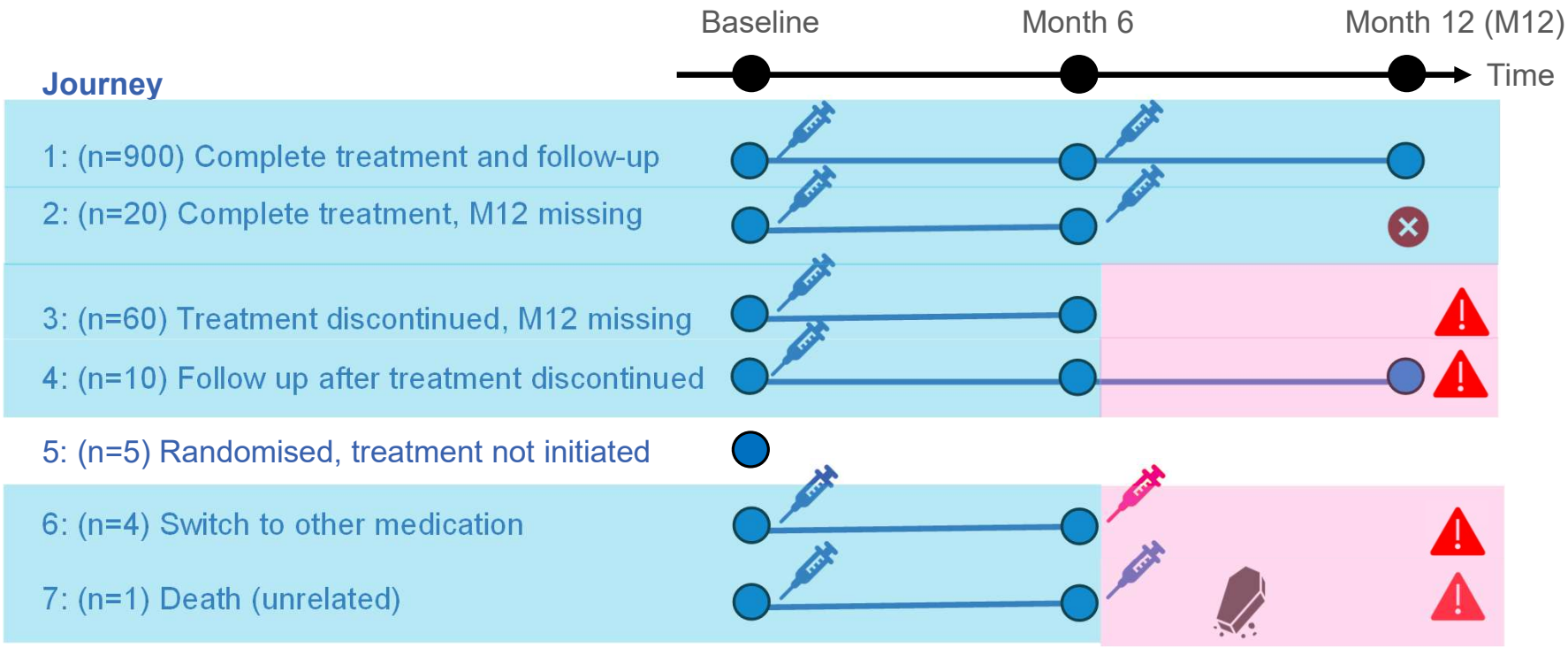
* as though they received both doses without any dosing deviations and without receipt of other osteoporosis medication (except supplements)?

Hypothetical scenario unobservable – requires imputation or prediction (standard methods)

Observed data after intercurrent events may not be relevant

Note: for simplicity, we will only consider a single strategy and not mix strategies

Which Data are Relevant for Estimating the Hypothetical Estimand?



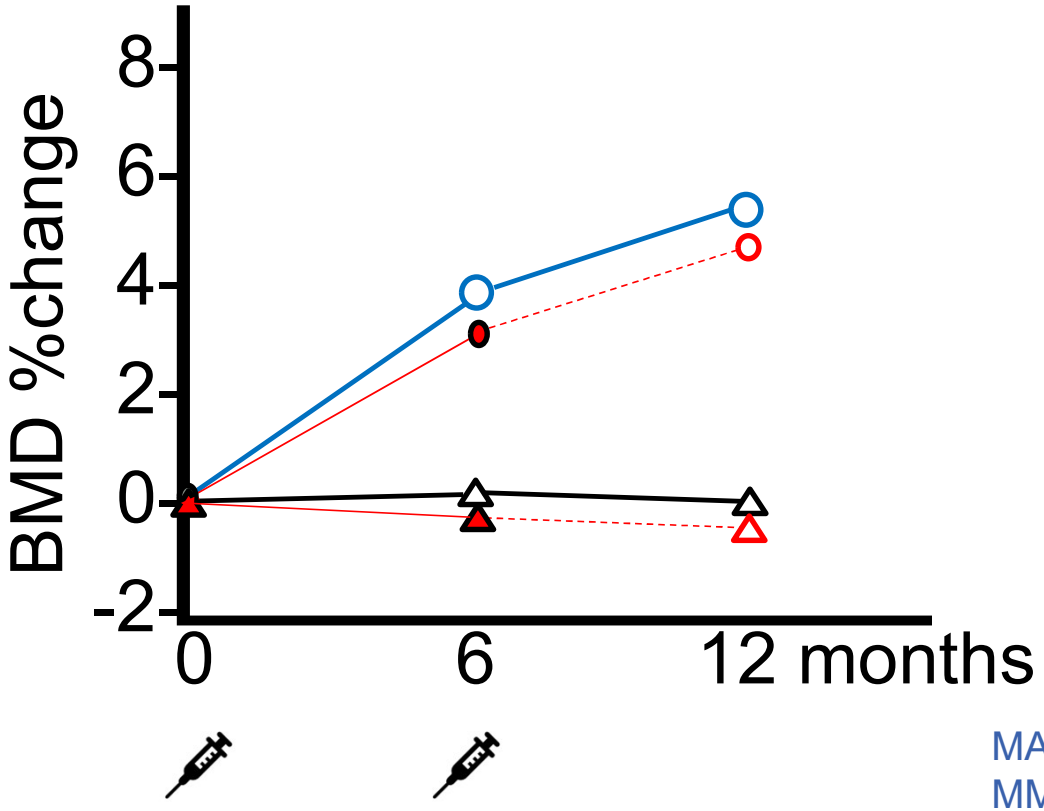
Relevant observation contributes to analysis

Relevant value is missing

Unobservable under the hypothetical scenario

Under MAR, MMRM predicts missing outcomes consistent with the observed treatment-arm trajectory

Usually Test vs Reference (placebo) are very different profiles …



Test mean

Reference mean

Last observed values (Placebo, Test) before withdrawal from study

Predicted value from MMRM (Placebo, Test)

MAR=missing at random
MMRM =Mixed Model with Repeated Measures

Estimation of a While-on-Treatment Estimand

Unchanged 2 attributes

What will be the difference in means between

Biosimilar and Prolia (denosumab) (subcutaneous injection every 6 months) in

rate of change* from baseline (slope) in lumbar spine bone mineral density (BMD) up to 52 weeks in

postmenopausal women with osteoporosis (treatment naïve)

* prior to missed dose or dosing deviations or receipt of other osteoporosis medication (except supplements)?

Less missing data

Observed data after intercurrent events are not relevant

JA4

Estimation of Composite Estimand

What will be the difference in means between μ_T-μ_R

Biosimilar and Prolia (denosumab) (subcutaneous injection every 6 months) in

composite endpoint* of % change from baseline in lumbar spJA2e bone mineral density (BMD) after 52 weeks in

postmenopausal women with osteoporosis (treatment naïve),

* where the endpoint takes a **failure value of zero** for discontinuation of treatment, dosing deviatiorJA3 and use of other osteoporosis medications

Less missing data

Observed data after intercurrent events are not relevant

Slide 30

JB1 What is the composite value when you have a % change from baseline? 0% change would represent no improvement (or worsening)
zxBIB Bell,Dr.,James (MED BDS) , 2026-09-14T09:59:00.060

SM1 0 %cfb=0 represents no improvement
McKendrick, Sue, 2026-09-21T14:47:17.984

JA2 The * does not have a footnote
Anzures Cabrera, Judith, 2026-10-01T13:13:10.412

JA3 Suggest "...zero (non responder) for discontinuation..."
Anzures Cabrera, Judith, 2026-10-01T13:20:19.384

JA4 I got a little lost in this slide, because I missed the % from change from baseline and thought that you were applying a composite strategy to a continuous endpoint. Presumably you will mention that you are now moving from a continuous endpoint to a binary one?
Anzures Cabrera, Judith, 2026-10-01T13:23:36.476

Estimation of Principal Stratum Estimand

$\mu_T - \mu_R$ What will be the difference in means between

Biosimilar and Prolia (denosumab) (subcutaneous injection every 6 months) in

%change from baseline in lumbar spine bone mineral density (BMD) after 52 weeks in

postmenopausal women with osteoporosis * (treatment naïve)

* who would complete treatment of both products without use of other osteoporosis medications

Patient stratum assignments unknown

Occurrence of the intercurrent event under both treatments is not jointly observed in a parallel-group trial.

Summary: the Estimand Strategy Determines which Data are Relevant

Strategy	Meaning of Strategy	Comments on Missing Data
Treatment policy	Outcome after intercurrent event is still of interest	Actual **post-intercurrent event outcomes remain relevant**. Missingness conditional on patient journey.
Hypothetical	Scenario in which the intercurrent event would not occur.	Hypothetical outcomes without the intercurrent event are **unobservable**; additional 'missing'.
While on treatment	Effect summarised over the period **prior to the intercurrent event.**	Post-intercurrent event observations **irrelevant**; Less missing data for the endpoint of interest.
Composite variable	Intercurrent event is incorporated into the endpoint definition.	Post-intercurrent event observations **irrelevant**. Composite endpoint assigned a failure value leading to less missing **composite endpoint data**
Principal stratum	Population defined by those in whom intercurrent event would/would not occur	Principal stratum assignments may be unknown in parallel group study where only one treatment observed; estimation relies on additional assumptions.

The estimand determines which data are relevant.
The estimation approach should reflect plausible missing data assumptions

Agenda

Introductions, Acknowledgements and Learning Outcomes

From Clinical Question to Aligning Estimation with Case Study Introduction

- What Alignment Means?

Challenges in Handling Missingness

- Assessing patient journeys, their intercurrent events and missing-data patterns

Which Data are Relevant for Different Strategies?

Treatment-policy Estimation in Practice

Practical Lessons for Trial Teams

Q & A

Issues in Treatment Policy Estimation

Treatment effect includes the impact of intercurrent events

- Treatment discontinuation
- Dosing deviations
- Use of other medications

Post-intercurrent event Month 12 outcomes remain relevant

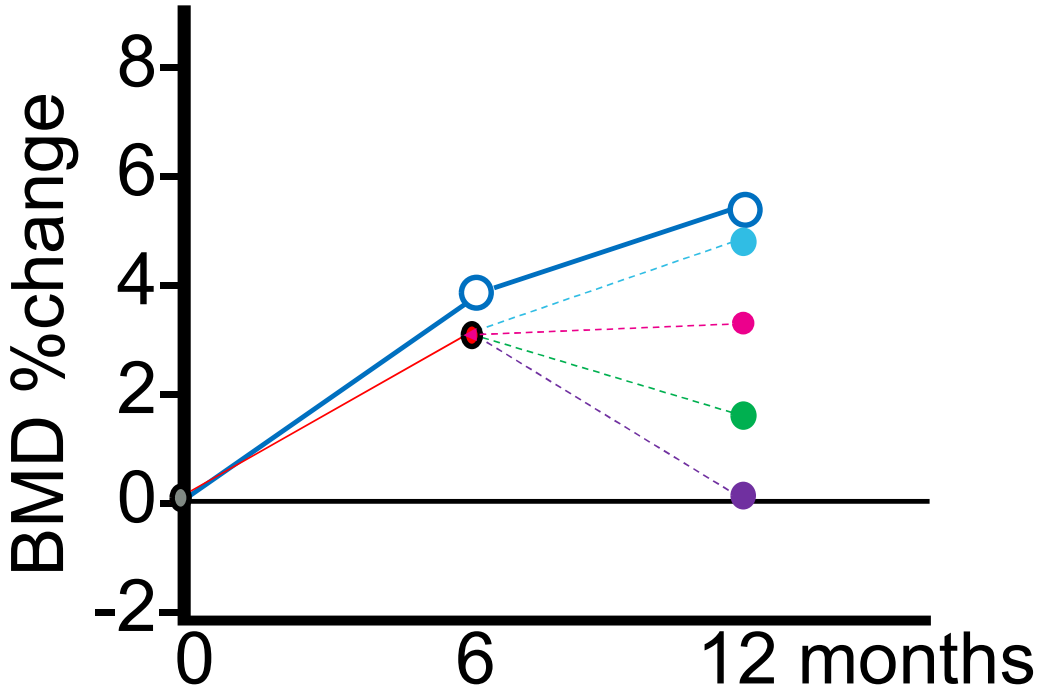
Why estimation is difficult

- Missingness may be more likely after intercurrent events
- Unobserved patients may differ from those observed
- Intercurrent events affect subsequent outcomes

Treatment-policy estimation would be simple with complete follow-up, but disproportionate missingness makes estimation difficult.

Poll: Which Trajectory is Most Clinically Plausible?

If a patient receives one dose of denosumab (but not the 6-month dose)



Possible trajectories

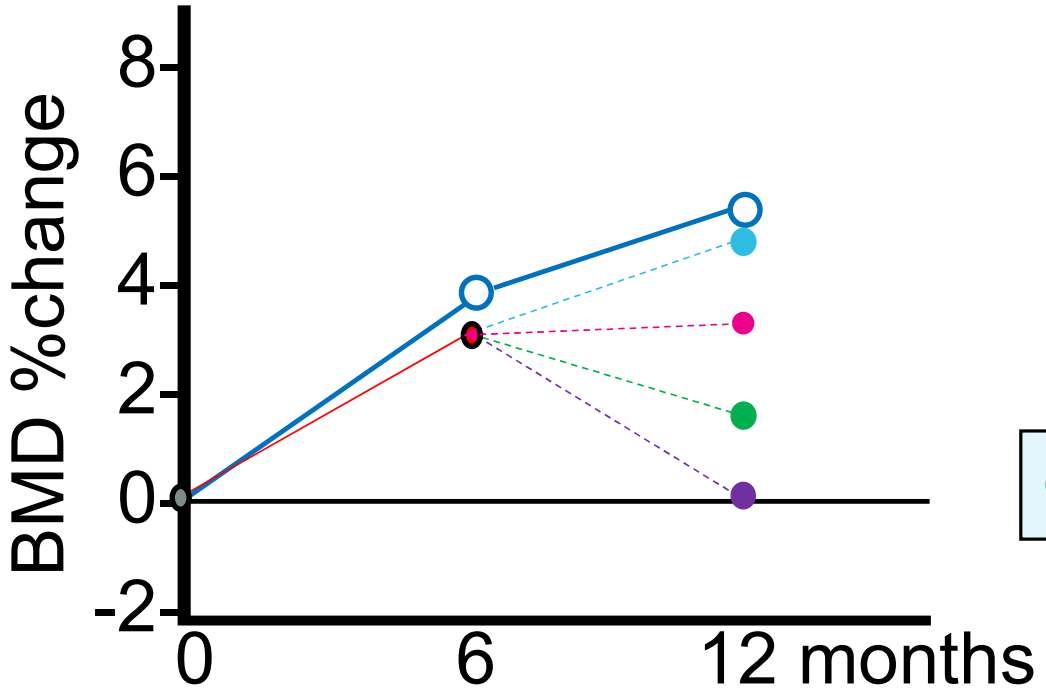
- (A) Continues on the fully treated trajectory
- (B) Maintenance of the effect
- (C) Slow loss of effect
- (D) Loss of all effect / return to baseline

○ Biosimilar mean ● Last observed value on individual, then missing 12 month data

Missing data assumptions need clinical justification.
Answer depends on treatment persistence, subsequent therapy and disease course.

Our Thoughts: Which Trajectory is Most Clinically Plausible?

If a patient receives one dose of denosumab (but not the 6-month dose)



Possible trajectories

- (A) Continues on the fully treated trajectory
- (B) Maintenance of the effect
- (C) Slow loss of effect
- (D) Loss of all effect / return to baseline

○ Biosimilar mean ● Last observed value on individual, LFU=lost to follow-up

Denosumab is long-acting monoclonal antibody. A gradual loss of effect is clinically more plausible than either full maintenance or an immediate return to baseline.

Which Trajectory is Most Clinically Plausible? (Case Study Assumptions for Treatment-Policy)

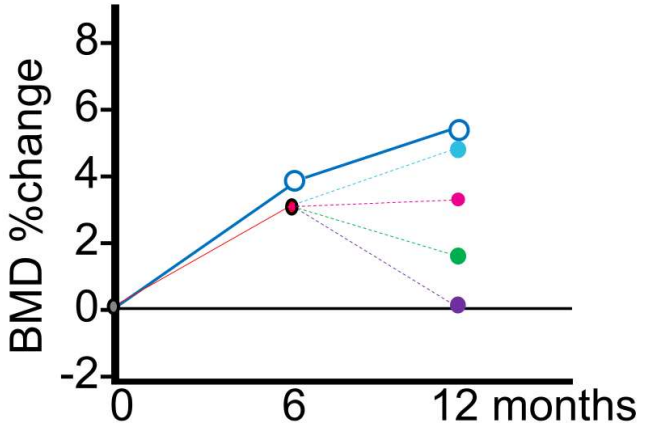
(i) Treatment discontinued after single dose (without alternative treatments) => C

(ii) Complete treatment / ongoing effect => A

(iii) Switch to other medication => A or B?

(D) Loss of all effect / return to baseline

Unlikely here (given Month 6 data), more likely if treatment was never initiated and the underlying condition is stable. Could be assumed as conservative approach?



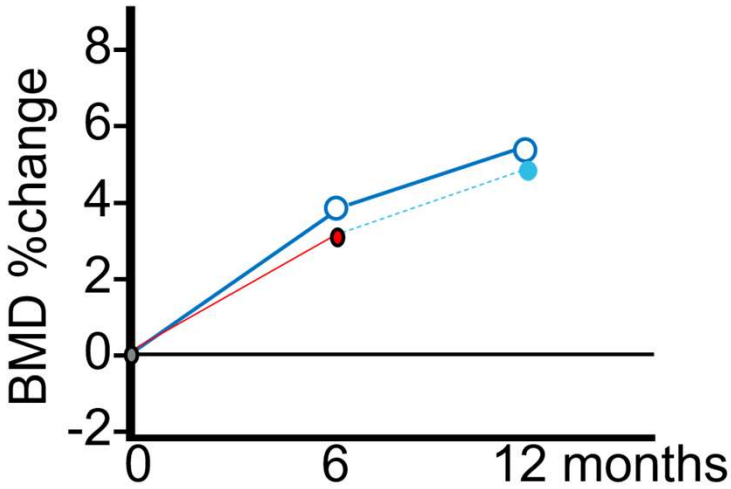
(A) Continues on the fully treated trajectory
(B) Maintenance of the effect
(C) Slow loss of effect
(D) Loss of all effect / return to baseline

Clinical context determines plausibility. When post-intercurrent event data are sparse, assumptions are influential and should be stress-tested.

Treatment Policy: When are MAR Analyses Credible?

MAR credible when treatment proceeds as planned, without intercurrent events





Probably not credible when disproportionate missingness after intercurrent events

WARNING: MMRM does not distinguish pre- and post- intercurrent event trajectories

- → clinically implausible → biased → misaligned
- Effectively assumes 'mostly on-treatment'
- May be acceptable if very few intercurrent events or missingness expected (low bias)

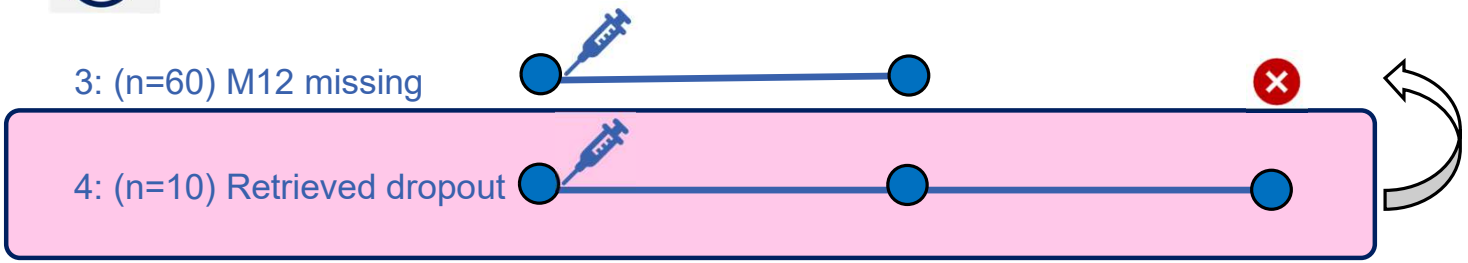
MAR= missing at random
MMRM = mixed model with repeated measures

Treatment Policy Estimation: Retrieved Dropout Method?

- Idea: Missing outcomes after intercurrent events may be like observed outcomes after intercurrent events
  - Retrieved dropout imputation

Assume Missing at Random (MAR) conditional upon occurrence (and timing) of intercurrent event

After treatment discontinued, can the retrieved-dropouts be utilised?



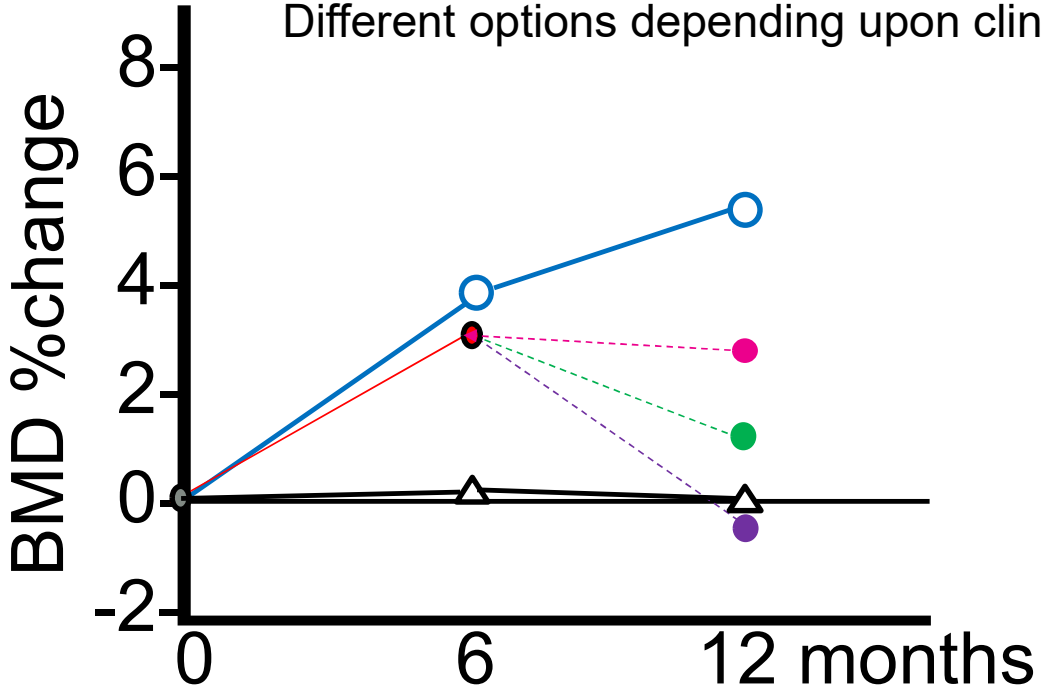
Insufficient retrieved dropouts => estimation difficulties & variance inflation

Reference-Based Multiple Imputation Approaches

- If insufficient retrieved dropout data is expected, what else can be done?
- In most cases, treatment policy is applied to changes (+/-) of treatment
- We can assume what happens to outcomes when treatments change
- Treatment discontinuation → some form of loss of treatment effect
  → Impute missing data using (negative) control arm
- Reference-Based Multiple Imputation

Reference-Based Multiple Imputation Approaches

Different options depending upon clinically plausible assumptions for trajectories



- **Copy Increment to Reference** / Maintenance of effect
- **Copy Reference** / Slow loss of effect
- **Jump to Reference** / Immediate loss of all effect

△ Placebo mean (if available)

○ Biosimilar mean

● Last observed value on individual

Pros and Cons of Treatment Policy Estimation Approaches

Approach	When it is most credible	Strengths	Main risks / cautions
Standard MAR (e.g. MMRM)	• Few or low impact IcEs • Little post-IcE missingness	• Simple and familiar • Usually precise • Reliable to implement (stable)	• Often anti-conservative bias
Retrieved Dropout (i.e. impute using observed post-IcE data)	• Post-IcE outcomes are observed for enough similar patients • Retrieved patients represent missing post-IcE patients	• Directly uses relevant post-IcE data • Often least biased, best aligned with treatment policy • Weakest assumptions	• Needs sufficient retrieved outcomes • Unstable or very imprecise if few retrieved dropouts • Difficult to pre-specify and implement as primary analysis
Reference Based (e.g. jump-to-reference, copy reference)	• Predictable loss-of-effect clinical course after IcE • Treatment discontinuation • Placebo reference arm	• Usually stable and precise • Different versions (assumptions) possible	• Strong assumptions about treatment effect → bias • Chosen trajectory must be clinically justified • Struggles without placebo control • Debate over variance estimation

IcE =Intercurrent event; MAR=missing at random; MMRM= mixed model with repeated measures

Agenda

Introductions, Acknowledgements and Learning Outcomes

From Clinical Question to Aligning Estimation with Case Study Introduction

- What Alignment Means?

Challenges in Handling Missingness

- Assessing patient journeys, their intercurrent events and missing-data patterns

Which Data are Relevant for Different Strategies?

Treatment-policy Estimation in Practice

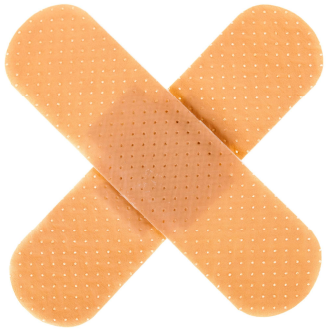
Practical Lessons for Trial Teams

Q & A

Better to Fix the Estimand and Trial Design Up Front

Alignment Starts with the Estimand and Trial Design

- Identify which outcomes the estimand requires.
- Design trial to target estimand and collect those outcomes whenever possible.
- Anticipate intercurrent event and missingness patterns.
- Use clinically plausible assumptions where relevant data are missing.



- **Estimation is only a sticky-plaster for mismatches between trial design and estimand; minimise the gap!**
- **Statistical modelling cannot replace relevant data the trial was never designed to collect.**

Summary Flow: From Estimand to Aligned Estimation



Reflect intercurrent event handling strategies

Goal: unbiased estimator under clinically plausible assumptions answering the intended clinical question

References

- Estimation alignment
  - Mallinckrodt et al. (2019). Aligning Estimators with Estimands in Clinical Trials: Putting the ICH E9(R1) Guidelines into Practice. *Therapeutic Innovation and Regulatory Science*, https://doi.org/10.1177/2168479019836979

- Treatment policy estimation
  - Bell et al. (2025). Estimation methods for estimands using the treatment policy strategy; a simulation study based on the PIONEER 1 Trial. *Pharmaceutical Statistics*; 24(2) https://doi.org/10.1002/pst.2472
  - Carpenter et al. (2013). Analysis of Longitudinal Trials With Protocol Deviation: A Framework for Relevant, Accessible Assumptions, and Inference via Multiple Imputation. *Journal of Biopharmaceutical Statistics* 23(6): 1352–1371, https://doi.org/10.1080/10543406.2013.834911.
  - Wolbers et al. (2022). Standard and reference-based conditional mean imputation. *Pharmaceutical Statistics*; 21(6): 1246-1257. https://doi.org/10.1002/pst.2234

Thank you for listening

Q&A

Please use the Q&A feature in Zoom for your questions and comments

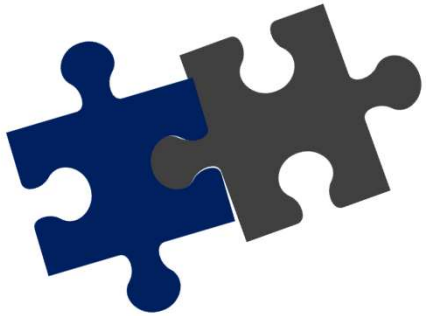
The Estimands Academy for Trial Teams

Aligning Estimation with the Clinical Question of Interest – 23 Nov 2026

- Chairperson: Judith Anzures-Cabrera (Astellas)

Panel

- James Bell (Elderbrook Solutions)
- Sue McKendrick (PPD, part of Thermo Fisher Scientific)
- David Wright (AstraZeneca)

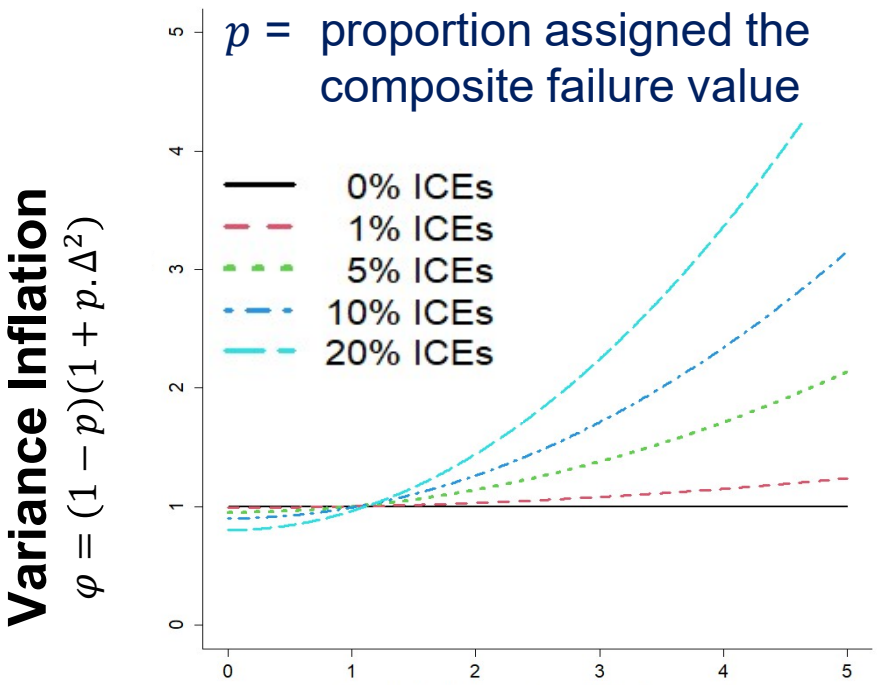


EFSPI European Federation of Statisticians in the Pharmaceutical Industry
Representing Statistical Associations in Europe

efpia European Federation of Pharmaceutical Industries and Associations

Back up Slide

Composite Estimands: Failure Values Affect Estimation



Δ = distance of failure value from the outcome mean, in SDs

Extreme failure values can substantially inflate variance

The farther the failure value lies from the outcome distribution, the greater the impact —

especially when composite failure value is assigned frequently.

The failure value is part of the estimand — choose a value that is both clinically meaningful and statistically workable.