Whether to use MMRM as primary estimand.

James Roger
London School of Hygiene & Tropical Medicine, London.

PSI/EFSPI European Statistical Meeting on Estimands.
Stevenage, UK: 28 September 2015.

Acknowledgements

- Mouna Akacha, Michael O'Kelly, James Carpenter & Mike Kenward for many discussions.
- All those project teams at GSK and elsewhere who have allowed me into their stories.
- But all the views expressed here are my own.

- These slides are designed to be self explanatory (so rather extensive).

[Declaration: James Roger and Livedata (UK) Ltd have a consultancy agreement with AstraZeneca. James Roger was employed by Johnson & Johnson and GlaxoSmithKlyne, and is a director of Livedata (UK) Ltd.]

Outline

- The salient features of MMRM.
- How the estimand is usually explained.
- Alternative estimands that it can possibly estimate.
- A personal view of the way forward.

What is MMRM

Mixed Models Repeated Measures.

- Not really a Mixed model but a useful acronym.

What is MMRM

Mixed Models Repeated Measures.

- Not really a Mixed model but a useful acronym.

Multivariate Normal distribution across visits within patient.

- Fully parameterised “Unstructured” covariance matrix.
  - Usually the same for each arm.
- Fixed effects linear model
  - Treatment by Visit interaction.
  - Other baseline covariates that are often crossed with visit.

MMRM: Why is MMRM important? (1)

- Flexible model for quantitative data measured repeatedly across multiple visits.
- Allows for correlation between observations within a subject.
- If data are complete and all covariates are crossed with visit (saturated model) then it is equivalent to a univariate analysis at each visit.

MMRM: Why is MMRM important? (2)

- Easy to fit in conventional software:
  - Fast iterative maximum likelihood solution.
  - Simple conjugate priors for Bayesian modelling.
  - But issues if there are more parameters than data (at last visit)!

MMRM: Why is MMRM important? (2)

- Easy to fit in conventional software:
  - Fast iterative maximum likelihood solution.
  - Simple conjugate priors for Bayesian modelling.
  - But issues if there are more parameters than data (at last visit)!
- Nice distributional properties:
  - Conditional model is MV Normal.
    e.g. distribution at second visit given known value at first visit.
  - Marginal model is MV Normal.
    e.g. distribution at second visit when value at first visit is not known.
  - But margin over mixture of Normals is not Normal.
    e.g. Marginal distribution from Pattern mixture models is not Normal.

MMRM: Why is MMRM important? (2)

- Easy to fit in conventional software:
  - Fast iterative maximum likelihood solution.
  - Simple conjugate priors for Bayesian modelling.
  - But issues if there are more parameters than data (at last visit)!
- Nice distributional properties:
  - Conditional model is MV Normal.
    e.g. distribution at second visit given known value at first visit.
  - Marginal model is MV Normal.
    e.g. distribution at second visit when value at first visit is not known.
  - But margin over mixture of Normals is not Normal.
    e.g. Marginal distribution from Pattern mixture models is not Normal.
- Last but not least . . .
  Popular way to handle “missing data”.

So what does it estimate?

The conventional answer

Illustration

To illustrate the issues I am going to assume, . . .

- an active treatment with primarily AE related withdrawals
- a control treatment with withdrawals mostly due to lack of efficacy.

. . . and see the implication.

Early withdrawal of randomized treatment

There are two different important impacts of early withdrawal from treatment.

1. **Selection:** Those who withdraw in one arm are a different population from . . .
   - those who **do not** withdraw in this arm
   - those who **do** withdraw in the other arm
   - those who **do not** withdraw in the other arm

   This messes up any analysis simply based on completers.

Early withdrawal of randomized treatment

There are two different important impacts of early withdrawal from treatment.

1. **Selection:** Those who withdraw in one arm are a different population from . . .
   - those who **do not** withdraw in this arm
   - those who **do** withdraw in the other arm
   - those who **do not** withdraw in the other arm

   This messes up any analysis simply based on completers.

2. **Switching:** After withdrawal patients receive some alternative treatment, which may or may not be closely aligned with the estimand, and may or may not be recorded within the trial datbase.

Early withdrawal of randomized treatment

1. **Selection:**
   - We can often observe the impact on this and other outcome measures before withdrawal and can try to correct for it.

Early withdrawal of randomized treatment

1. **Selection:**
   - We can often observe the impact on this and other outcome measures before withdrawal and can try to correct for it.

2. **Switching:** We can decide whether to include the impact of patients receiving some alternative treatment within the estimand.
   - **De jure**: We ignore it and are interested in either . . .
     1. outcome if “carry on taking treatment”,
     2. or perhaps restrict to population that can take either treatment.
   - **De facto**: We allow for impact of alternative treatments in an ITT fashion. Need to . . .
     - either observe outcome after withdrawal and use it,
     - or hypothesise some pattern of treatment profile after withdrawal.

Early withdrawal of randomized treatment

1. **Selection:**
   - We can often observe the impact on this and other outcome measures before withdrawal and can try to correct for it.

2. **Switching:** We can decide whether to include the impact of patients receiving some alternative treatment within the estimand.
   - **De jure**: We ignore it and are interested in either . . .
     1. outcome if "carry on taking treatment",
     2. or perhaps restrict to population that can take either treatment.
   - **De facto**: We allow for impact of alternative treatments in an ITT fashion. Need to . . .
     - either observe outcome after withdrawal and use it,
     - or hypothesise some pattern of treatment profile after withdrawal.

Here we ignore switching and go down the two **de jure** routes.

Missing Data: MMRM is just a model

MMRM is a sensible model for . . .

- what happens to a quantiative outcome on randomized treatment
- in a homogeneous population
- **before** “something goes wrong” and the patient is withdrawn from treatment.

Missing Data: MMRM is just a model

MMRM is a sensible model for . . .

- what happens to a quantiative outcome on randomized treatment
- in a homogeneous population
- **before** “something goes wrong” and the patient is withdrawn from treatment.

It is not an **estimand** in itself.

- Under a Missing at Random (MAR) assumption it allows one to make statements about what would have happened if patients had in some sense **remained in the study**.

Missing Data: MMRM is just a model

MMRM is a sensible model for . . .

- what happens to a quantiative outcome on randomized treatment
- in a homogeneous population
- **before** "something goes wrong" and the patient is withdrawn from treatment.

It is not an **estimand** in itself.

- Under a Missing at Random (MAR) assumption it allows one to make statements about what would have happened if patients had in some sense **remained in the study**.
  - But this is a Survey Statistician's approach to missing data (the data is out there somewhere if only I could find it).

Missing Data: MMRM is just a model

MMRM is a sensible model for . . .

- what happens to a quantiative outcome on randomized treatment
- in a homogeneous population
- **before** "something goes wrong" and the patient is withdrawn from treatment.

It is not an **estimand** in itself.

- Under a Missing at Random (MAR) assumption it allows one to make statements about what would have happened if patients had in some sense **remained in the study**.
  - But this is a Survey Statistician's approach to missing data (the data is out there somewhere if only I could find it).
  - Here such data are "hypothetical". Not only may it not have happened it may even be impossible.

Missing at Random (MAR)

Missing at Random is best seen in terms of **Selection** models

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\theta, \phi) \;=\; p(\mathbf{R}|\mathbf{y}_{obs}, \mathbf{y}_{mis}, \phi)\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\theta)$$

where **R** is a vector indicating whether data are observed (1) or not (0) at visit i. Here **R** is constrained by monotonicity.

Under MAR the probability of withdrawal at visit i **only** depends on baseline covariates and **previous** observed values.

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\theta, \phi) \;=\; p(\mathbf{R}|\mathbf{y}_{obs}, \phi)\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\theta)$$

Missing at Random (MAR)

Missing at Random is best seen in terms of **Selection** models

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\theta, \phi) \;=\; p(\mathbf{R}|\mathbf{y}_{obs}, \mathbf{y}_{mis}, \phi)\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\theta)$$

where **R** is a vector indicating whether data are observed (1) or not (0) at visit i. Here **R** is constrained by monotonicity.

Under MAR the probability of withdrawal at visit i **only** depends on baseline covariates and **previous** observed values.

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\theta, \phi) \;=\; p(\mathbf{R}|\mathbf{y}_{obs}, \phi)\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\theta)$$

Then we simply estimate parameters based on $p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\theta)$ ignoring the missingness process.

[θ and ϕ are parameters.]

Missing at Random (MAR)

By conditioning on the baseline covariates and previous observed values can we remove the impact of the **selective aspect of patient withdrawal**?

If so, then we have MAR.

Then we can correct for any difference between those withdrawing and those who do not.

How does it work

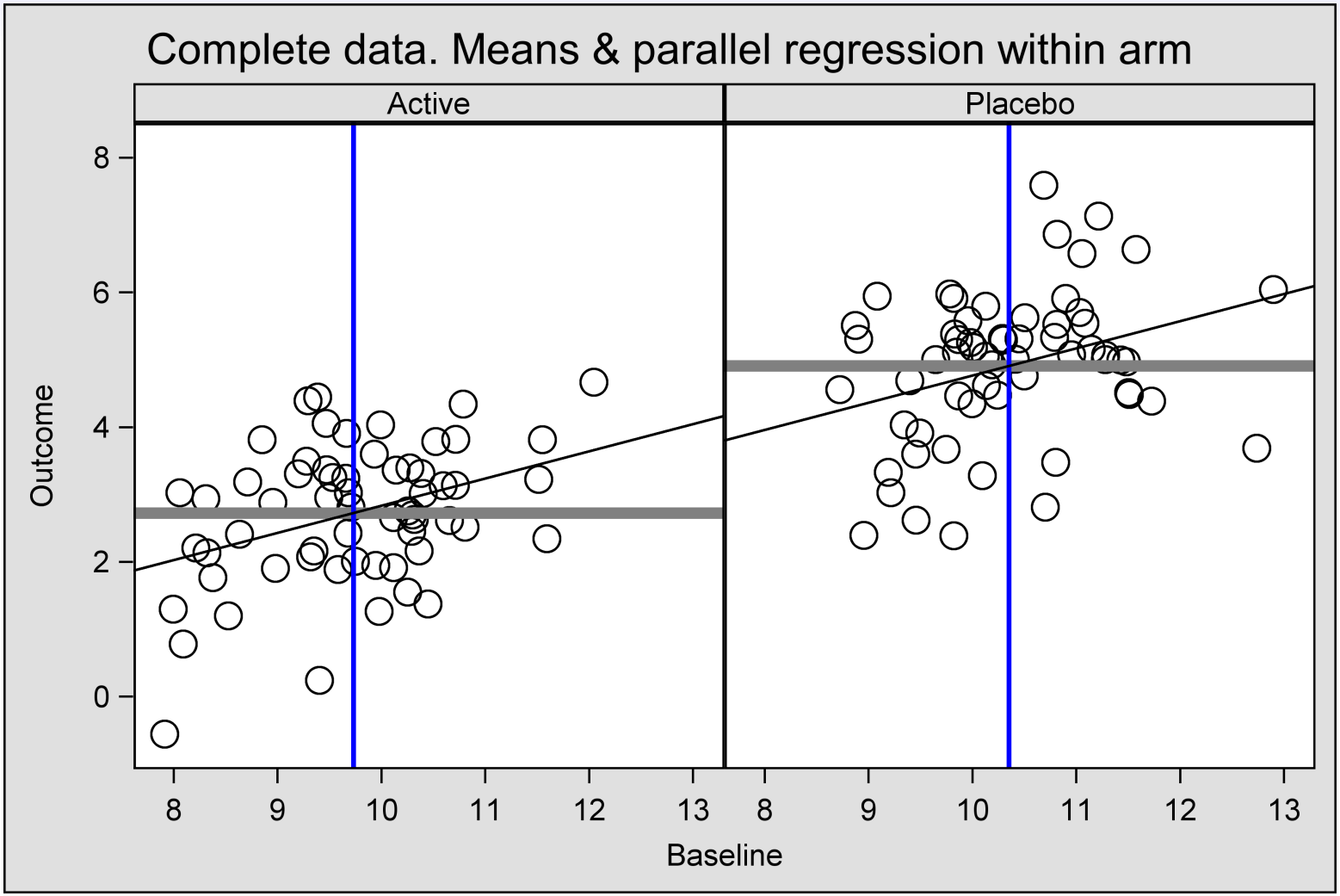
- It uses baseline covariates and pre-withdrawal observed values to correct the treatment difference we see between those who complete.

How does it work

- It uses baseline covariates and pre-withdrawal observed values to correct the treatment difference we see between those who complete.

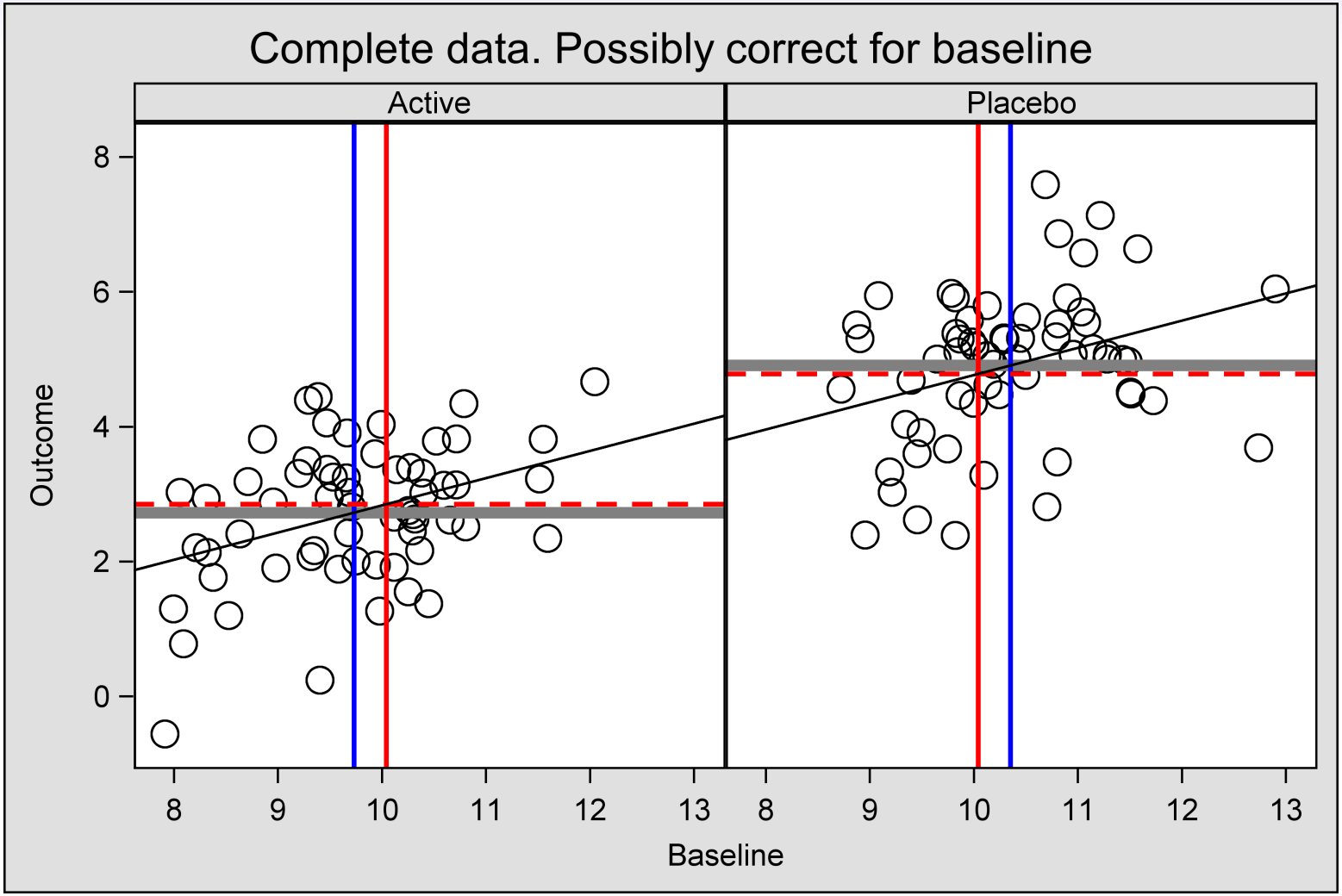
First consider just baseline data with no post-randomization data before outcome.

Complete data: Plot versus baseline.



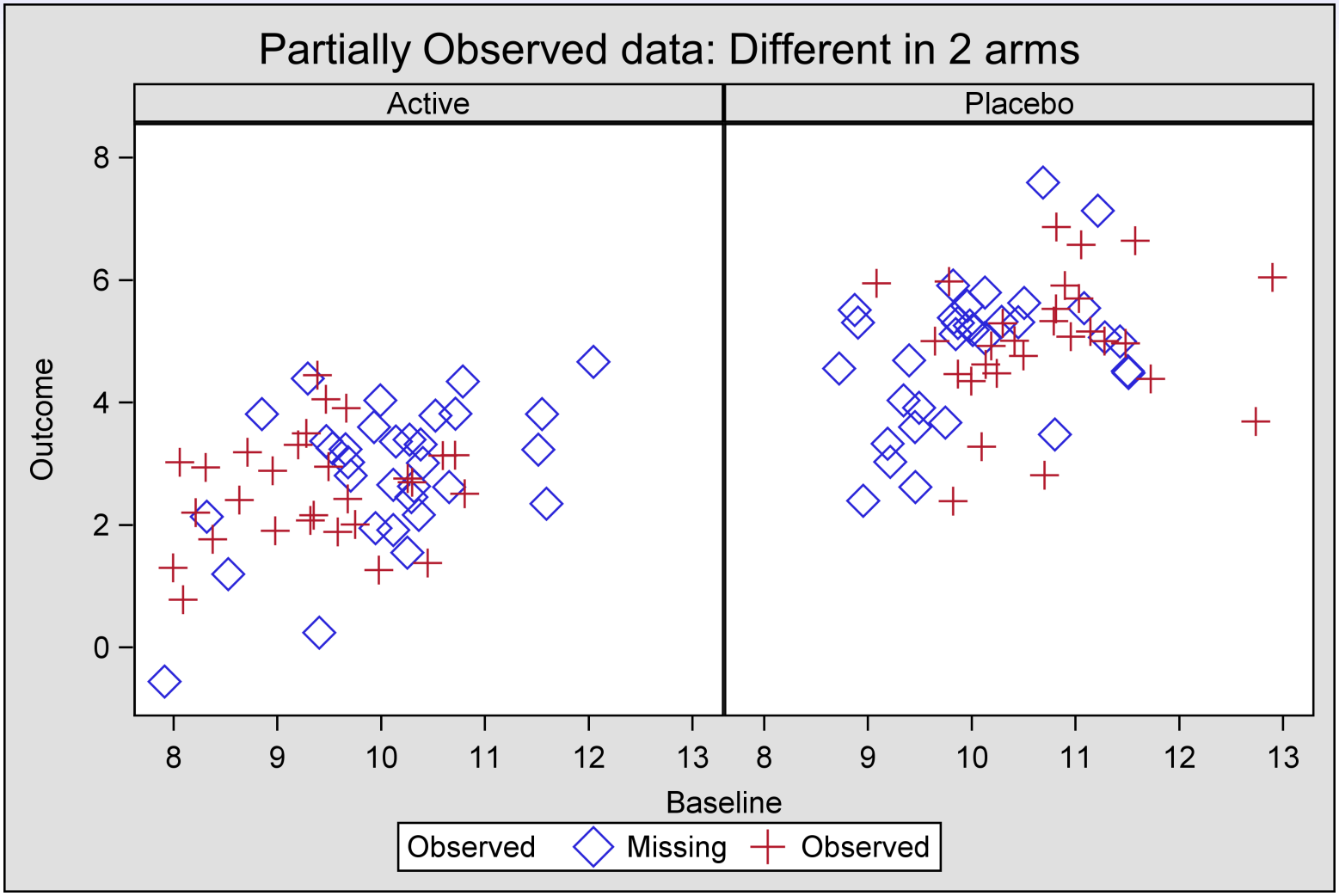
Randomization means baseline means should be similar.

Complete data: Adjust for baseline diff.



Estimated treatment difference is slightly smaller.

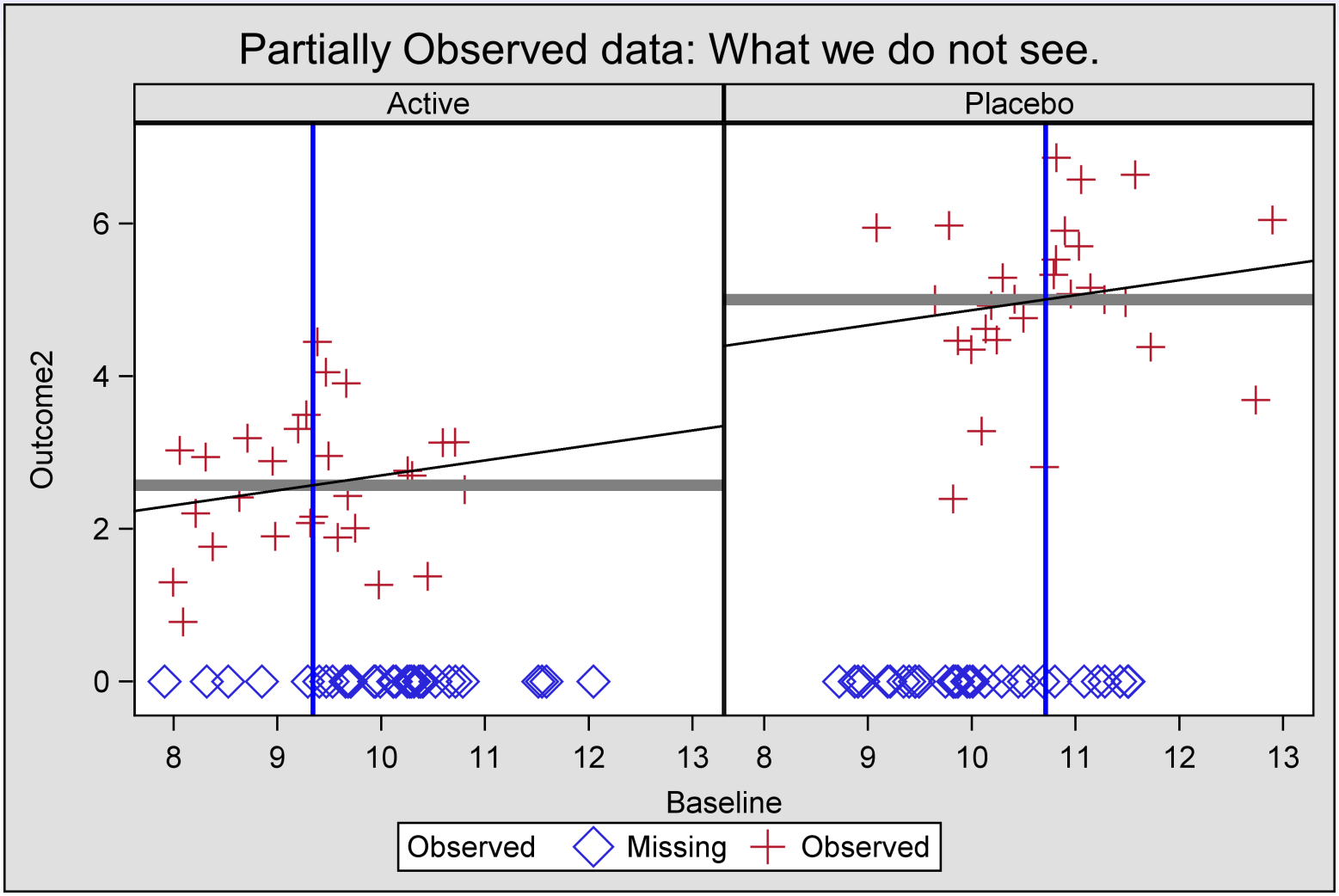
Missing data: Observed are +.



Removed low baseline/outcome from Placebo

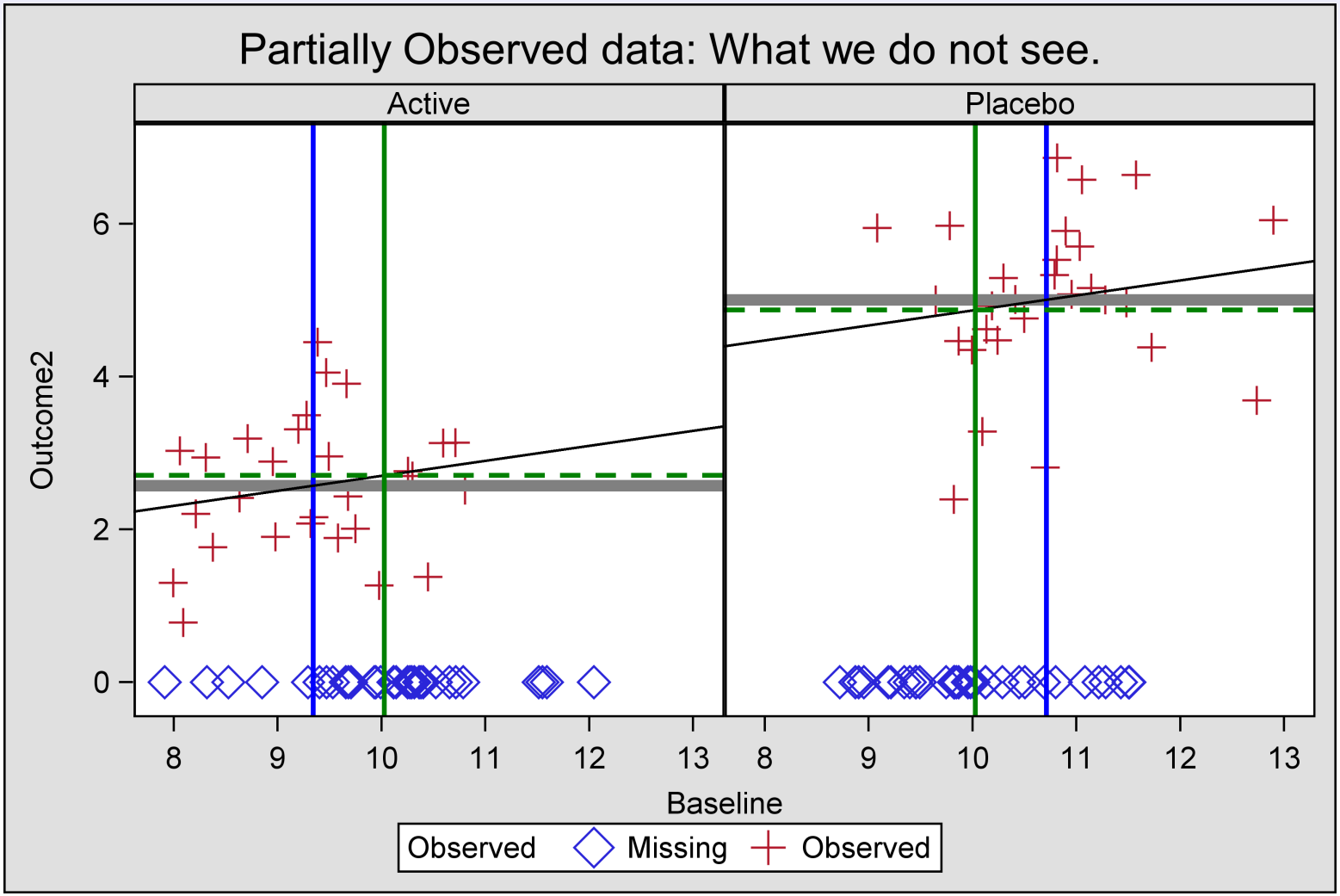
Removed high baseline/outcome from Active.

Missing data.



MAR is also MCAR. Nothing post-randomization to condition on.

Baseline correction.



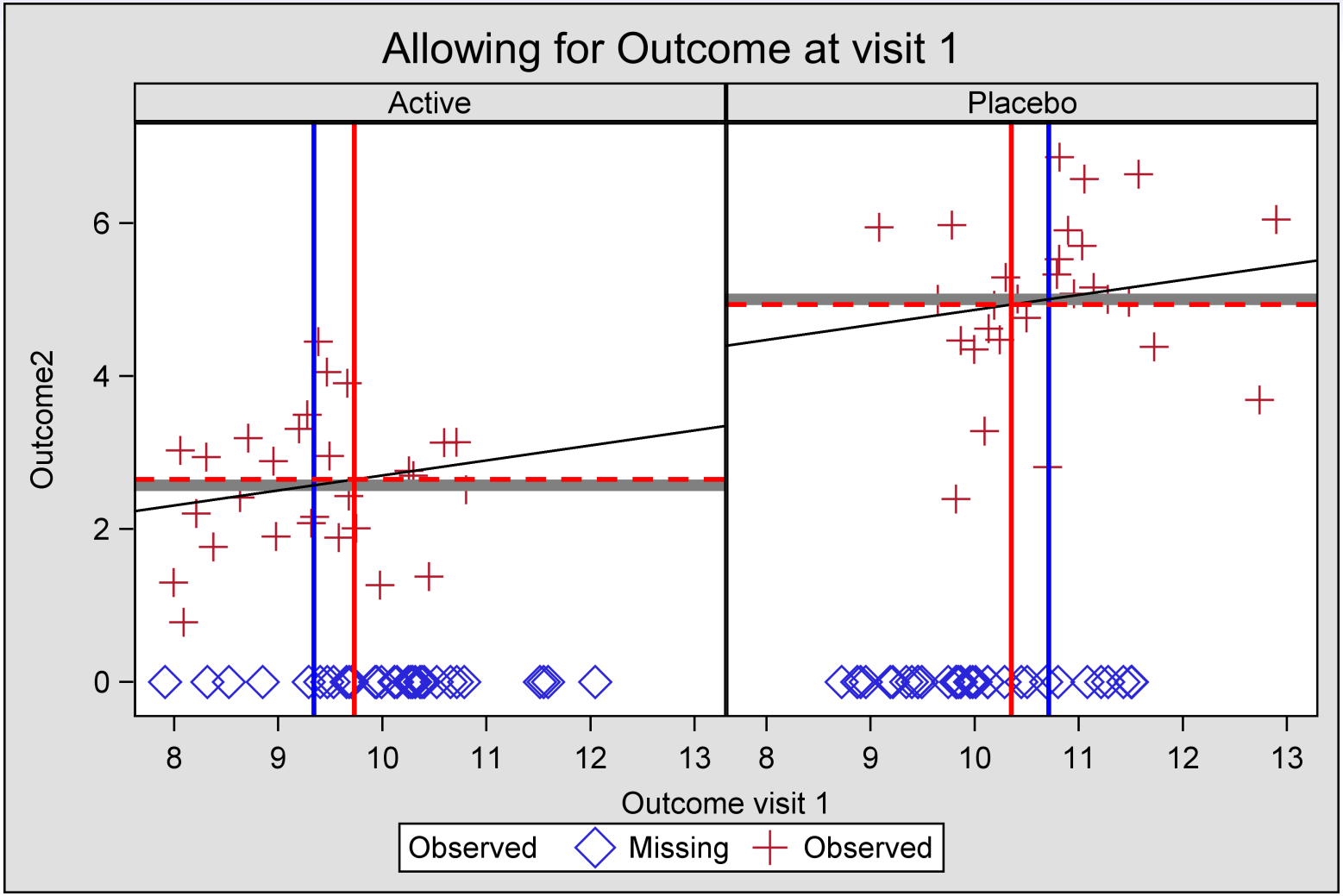
Baseline values for those withdrawn are ignored.

Two visits and no baseline.

- Use the same data layout but now the X-axis is Outcome at first visit rather than baseline.
- This means that we **expect** X-axis mean to be different between arms.

Rather than adjust to the same (mean) X-axis values, now correct each arm to mean of the observed X-axis values within each arm.

Previous outcome data.



Difference between Outcome values at visit 1 is important.

MMRM analysis

MMRM sinply applies these two ideas with

- several baseline covariates and
- several post-randomization observed values.

MMRM analysis

MMRM sinply applies these two ideas with

- several baseline covariates and
- several post-randomization observed values.

- Note how this links directly to how baseline covariates can be included as part of the Repeated Measures series.
  No treatment effect at extra visit.

The conventional answer summarised

Under MAR the MMRM model estimates the mean treatment effect assuming that . . .

- after withdrawal subjects would have continued just like their peers in the same arm who have . . .
  - the same covariates and
  - same observed data (so far).

The conventional answer summarised

Under MAR the MMRM model estimates the mean treatment effect assuming that . . .

- after withdrawal subjects would have continued just like their peers in the same arm who have . . .
  - the same covariates and
  - same observed data (so far).

This links with equivalent ideas of Multiple Imputation (MI) under MAR, where we impute as if patients continue in the same way within arm.

A better way forward?

How can we develop easily interpreted Repeated Measures models?

. . .

Rather than use selection models or pattern mixture models we need to model the outcome and missingness processes as they evolve across visits . . .

Extrapolation Factorizaton (EF)

In contrast to both pattern mixture

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\omega) \;=\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\mathbf{R}, \omega)\; p(\mathbf{R}|\omega)$$

and selection models . . .

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{mis}|\omega) \;=\; p(\mathbf{R}|\mathbf{y}_{obs}, \mathbf{y}_{mis}, \omega)\; p(\mathbf{y}_{obs}, \mathbf{y}_{mis}|\omega)$$

. . . the distribution of the full data (observed and unobserved) is factored into

1. joint distribution of the observed values $\mathbf{y}_{obs}$ and the Response pattern $\mathbf{R}$,
2. that for unobserved or **potential** values $\mathbf{y}_{pot}$ conditional on the previous [effectively same as $\mathbf{y}_{mis}$].

$$p(\mathbf{y}_{obs}, \mathbf{R}, \mathbf{y}_{pot}|\omega) \;=\; p(\mathbf{y}_{pot}|\mathbf{y}_{obs}, \mathbf{R}, \omega_E)\; p(\mathbf{y}_{obs}, \mathbf{R}|\omega_O)$$

[The Extrapolation Factorizaton (EF). See Daniels & Hogan. Chapter 9, section 9.1.1. $\omega = (\theta, \phi)$]

Sequential distribution - 3 visits

*Visit*1 Y_1		*Withdraw*? R_2		*Visit*2 Y_2		*Withdraw*? R_3		*Visit*3 Y_3
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$

$f(Y_i|Y_1, \ldots, Y_{(i-1)}, R_i = k)$ is distribution of outcome conditonal upon history.
$g(R_i|Y_1, \ldots, Y_{(i-1)})$ is probability of continuing to observe at visit i **conditional** upon observing at vist (i-1).
Both will depend on baseline and possibly other previous post-randomization observations - especially on treatment regime.
[Here we ignore possible withdrawal before first visit (No R_1).]

Possible flows

*Visit*1	|	*Withdraw?*	|	*Visit*2	|	*Withdraw?*	|	*Visit*3
Y_1	$\rightarrow$	R_2	$\rightarrow$	Y_2	$\rightarrow$	R_3	$\rightarrow$	Y_3
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$

The part we always observe

*Visit*1 Y_1	$\rightarrow$	*Withdraw*? R_2	$\rightarrow$	*Visit*2 Y_2	$\rightarrow$	*Withdraw*? R_3	$\rightarrow$	*Visit*3 Y_3
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$

We can model these distributions in red and estimate parameters based on data.

The margin distribution for final visit

*Visit*1 Y_1		*Withdraw*? R_2		*Visit*2 Y_2		*Withdraw*? R_3		*Visit*3 Y_3	*Margin*
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 1)$	$f(Y_3, R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_3 = 0)$	$f(Y_3, R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1, R_2 = 0)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2, R_2 = 0)$	$f(Y_3, R_2 = 0)$
									$f(Y_3)$

- If we know all these intermediate distributions then we can derive “least squares means” for the margin for some predefined set of covariate values.
- That is we **can get the treatment effect** by differencing between “means” based on one arm or the other.
- The other terms we may base on post-withdrawal data or by sharing parameters from elsewhere.

Missing at Random (MAR)

*Visit*1 Y_1		*Withdraw*? R_2		*Visit*2 Y_2		*Withdraw*? R_3		*Visit*3 Y_3	*Margin*
$f(Y_1)$	$\rightarrow$	$g(R_2 \mid Y_1)$	$\rightarrow$	$f(Y_2 \mid Y_1)$	$\rightarrow$	$g(R_3 \mid Y_1, Y_2)$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2)$	$f(Y_3) \times g(R_3 = 1)$
		$\downarrow$				$\searrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2)$	$f(Y_3) \times g(R_3 = 0)$
		$\searrow$	$\rightarrow$	$f(Y_2 \mid Y_1)$	$\rightarrow$	$\rightarrow$	$\rightarrow$	$f(Y_3 \mid Y_1, Y_2)$	$f(Y_3) \times g(R_2 = 0)$
									$f(Y_3)$

MAR indicates that $f(Y_2|Y_1, R_2)$ is the same irrespective of R_2 being 0 or 1.
So margin is simply $\int \int f(Y_1) \times f(Y_2|Y_1) \times f(Y_3|Y_1, Y_2) dY_1 dY_2$.
Simply ignore the missingness part of the distribution.

A general framework

- Extrapolation Factoristaion (EF) provides a general framework for deriving estimands based on a flexible approach to the selectiveness of the withdrawal process.

- Can also incorporate **de facto** style estimands. As such it is a suitable way forward for handling the switching issue.
- An MCMC approach can be applied to the model leading to . . .
  - an overall Bayesian approach, or
  - imputation of complete data and a summary across multiple analyses (MI).

  It could possibly allow an estimand to be defined in a subpopulation based on withdrawal process.
  - . . . those who would reamin on treatment in either arm.

So what does MMRM + MAR estimate?

- The title of the talk should have been “Whether to use MMRM + MAR as primary estimand.”
- MMRM has tried to answer the **selection** issue but ignored the **switching issue**.

So what does MMRM + MAR estimate?

- The title of the talk should have been “Whether to use MMRM + MAR as primary estimand.”
- MMRM has tried to answer the **selection** issue but ignored the **switching issue**.
- As a result it answers a hypothetical question. Can it answer more (**de jure** option 2)?

Circumventing the hypothetical nature

- One might restrict to those subjects who can tolerate the treatment.

Circumventing the hypothetical nature

- One might restrict to those subjects who can tolerate the treatment.
  Restrict to those who can tolerate . . .
  - either treatment.
    or
  - the new treatment, as this is our focus (while retaining counterfactual for Control).

Circumventing the hypothetical nature

- One might restrict to those subjects who can tolerate the treatment.
  Restrict to those who can tolerate . . .
  - either treatment.
    or
  - the new treatment, as this is our focus (while retaining counterfactual for Control).
    [If the patient stops treatment then their additional cost may be small!]

Circumventing the hypothetical nature

- One might restrict to those subjects who can tolerate the treatment.
  Restrict to those who can tolerate . . .
  - either treatment.
    or
  - the new treatment, as this is our focus (while retaining counterfactual for Control).
    [If the patient stops treatment then their additional cost may be small!]

MMRM + MAR does not directly answer this question.

How it is often interpreted

- MMRM is sometimes thought of as the treatment effect in those who can accept either treatment regime.
- What additional assumptions do we need to interpret it in this way?
  - A predefined baseline covariate condition that guarantees trial completion in both arms. OR
  - There is effectively no difference between arms in the way previous outcome impacts on the probability of withdrawal.

How it is often interpreted

- MMRM is sometimes thought of as the treatment effect in those who can accept either treatment regime.
- What additional assumptions do we need to interpret it in this way?
  - A predefined baseline covariate condition that guarantees trial completion in both arms. OR
  - There is effectively no difference between arms in the way previous outcome impacts on the probability of withdrawal.

  If not, then we could model the whole thing and draw conclusion.

Conclusions

- MMRM under MAR defines a classic **de jure** estimand.
- MAR before withdrawal and MNAR after withdrawal using modified post-withdrawal RM distribution defines a series of possible **de facto** estimands.
- A full EF sequential specification of the model allows the derivation of a wide range of possible estimands.

Conclusions

- MMRM under MAR defines a classic **de jure** estimand.
- MAR before withdrawal and MNAR after withdrawal using modified post-withdrawal RM distribution defines a series of possible **de facto** estimands.
- A full EF sequential specification of the model allows the derivation of a wide range of possible estimands.
- **The Regulators need to state the kinds of estimand they want.**
  This does not need to be limited by currently available statistical methods.
  Nearly anything is possible.

Abstract (1)

There are two main impacts of early withdrawal on study results; first the potential selection bias caused by those withdrawing being different from the remaining patients, and second the fact that patients may receive alternative treatments after withdrawal. The most common method for handling early withdrawal in clinical studies is MMRM or some other form of missing at random (MAR) based analysis. The motivation behind MMRM is to solve the first issue while addressing an on-treatment question, i.e. what happens if a typical patient completes their assigned treatment. It does this by conditioning on the previous observations and other covariates that may inform on both missingness and outcome.

Abstract (2)

So what is the scientific question of interest, i.e. the estimand that it targets? This cannot be answered without considering the second aforementioned issue which is related to treatment switching or modification after withdrawal. If the design of the study, in terms of treatment after withdrawal from randomized treatment, matches the estimand, then collection of data after treatment withdrawal allows direct analysis. Then later absolute study termination after switching (truly missing data) can be handled via a modified MMRM approach. But when the design of the study after treatment withdrawal does not match the estimand any analysis must depend upon additional unverifiable information or assumptions. This is a whole new area of potential statistical research. Several of the more recent proposals ignore the first issue of selection bias. Any coherent approach must address both issues.

The proposal from Tom Permutt

One might restrict to those subjects who can tolerate the treatment.
Restrict to those who can tolerate . . .

- either treatment
- the new treatment, as this is our focus (while retaining counterfactual for Control).

The proposal from Tom Permutt effectively does neither of these.

- Withdrawals receive worst outcome and then he discards **within each arm** the worst x% of subjects.
- So discarded subjects are different sorts of subject in each arm. Withdrawn for AE versus withdrawn for lack of efficacy, say.
- So this can only **test** if there is any difference between treatments. It does not define an estimand to compare between arms.