

# Joint EFSPI / PSI Meeting: Advances in treatment of Missing Data. Pattern Mixture Models: An introduction

**Dr James Roger.**  
**London School of Hygiene and Tropical Medicine.**

18 November 2011.

[These Latex/Beamer slides are the intellectual property of the individual presenter and are protected under copyright laws. Used by permission. All rights reserved.]

# Acknowledgements

Acknowledgements to

James Carpenter (LSHTM)

Michael Kenward (LSHTM)

## Example application

- Placebo and lowest active dose (200 mcg) from a 5 arm asthma clinical trial of budesonide in chronic asthma.

## Example application

- Placebo and lowest active dose (200 mcg) from a 5 arm asthma clinical trial of budesonide in chronic asthma.
- Study found statistically significant dose-response effect for mean PEFR and FEV<sub>1</sub>. Here we consider just FEV<sub>1</sub>.

## Example application

- Placebo and lowest active dose (200 mcg) from a 5 arm asthma clinical trial of budesonide in chronic asthma.
- Study found statistically significant dose-response effect for mean PEF and FEV<sub>1</sub>. Here we consider just FEV<sub>1</sub>.  
This answered a **de jure** question.

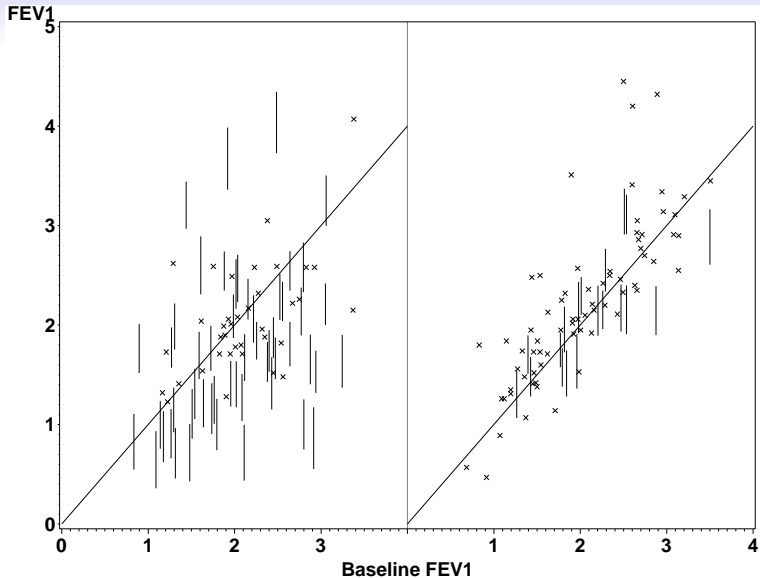
## Example application

- Placebo and lowest active dose (200 mcg) from a 5 arm asthma clinical trial of budesonide in chronic asthma.
- Study found statistically significant dose-response effect for mean PEF and FEV<sub>1</sub>. Here we consider just FEV<sub>1</sub>. This answered a **de jure** question.
- **Deviation** is withdrawal from treatment or other protocol violation.

## Example application

- Placebo and lowest active dose (200 mcg) from a 5 arm asthma clinical trial of budesonide in chronic asthma.
- Study found statistically significant dose-response effect for mean PEF and FEV<sub>1</sub>. Here we consider just FEV<sub>1</sub>.  
This answered a **de jure** question.
- **Deviation** is withdrawal from treatment or other protocol violation.
- Proportion deviating is much higher in the placebo arm.

# Final visit FEV<sub>1</sub> \* Baseline: imputed (MAR)



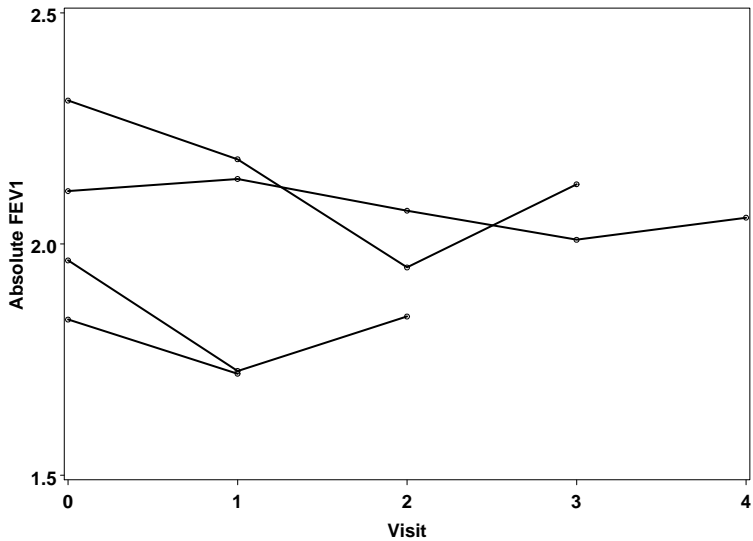
## Mean Fev<sub>1</sub> by Pattern

Dropout pattern	Placebo arm					Number	Percent
	Mean FEV <sub>1</sub> (litres) measured at week						
	0	2	4	8	12		
1	2.11	2.14	2.07	2.01	2.06	37	40
2	2.31	2.18	1.95	2.13	—	15	16
3	1.96	1.73	1.84	—	—	22	24
4	1.84	1.72	—	—	—	16	17
All patients (Mean)	2.11	1.97	1.98	2.04	2.06	90	100
All patients (Std.)	0.57	0.67	0.56	0.58	0.55		
	Budesonide 200 mcg						
1	2.03	2.22	2.23	2.24	2.23	71	78
2	1.93	1.91	2.01	2.14	—	8	9
3	2.28	2.10	2.29	—	—	8	9
4	2.24	1.84	—	—	—	3	3
All patients (Mean)	2.03	2.17	2.22	2.23	2.23	90	100
All patients (Std.)	0.65	0.75	0.80	0.85	0.81		

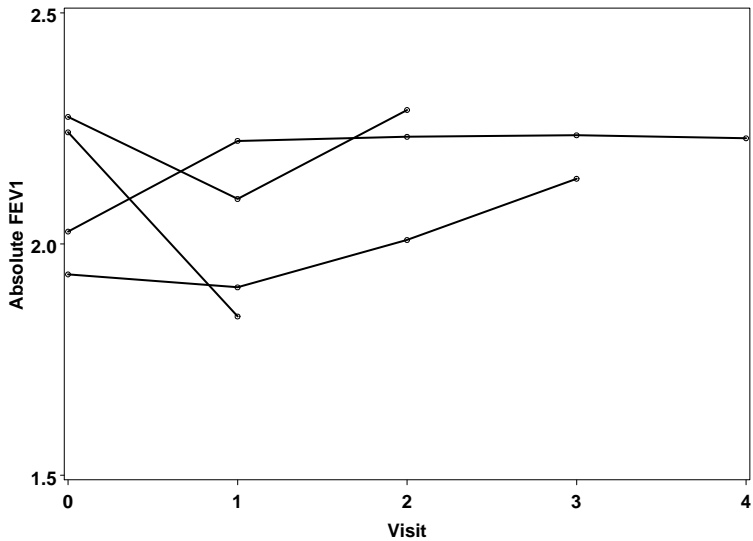
## Example: Primary endpoint for each Pattern

- Plot the primary endpoint within each **Deviation Pattern**, for each treatment arm.

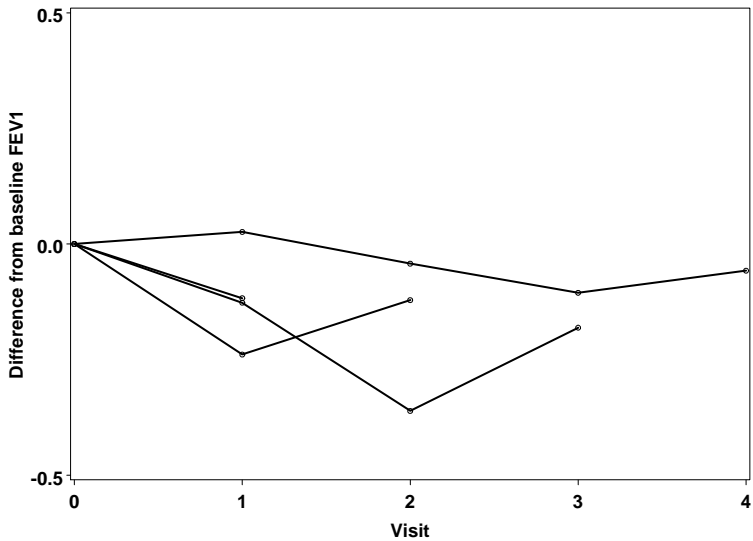
Treatment=Placebo



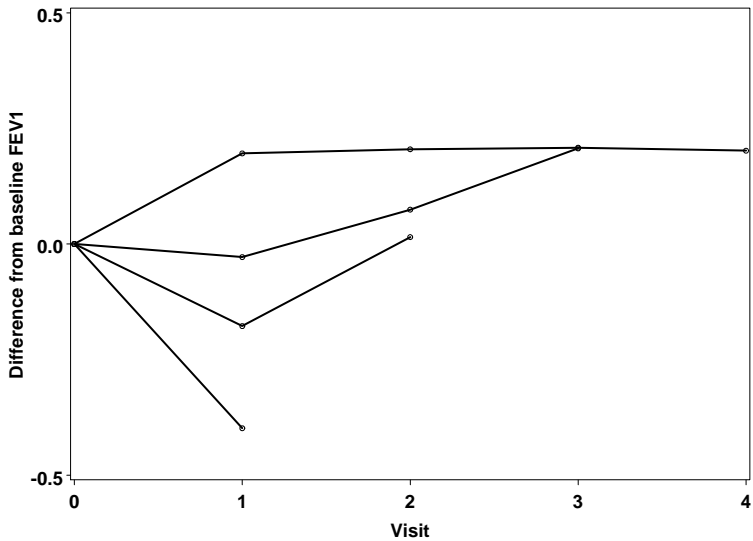
Treatment=Budesonide 200 mcg



Treatment=Placebo



Treatment=Budesonide 200 mcg



# Beyond “Randomised-arm MAR”

Two main directions.

- **De jure:** Allow in some way, other than MAR, for bias due to selection of those with missing data.

## Beyond “Randomised-arm MAR”

Two main directions.

- **De jure:** Allow in some way, other than MAR, for bias due to selection of those with missing data.
- **De facto:** Allow for switching to alternative treatment regime. Both of these are MNAR, but motivation and form is different.

## Beyond “Randomised-arm MAR”

Two main directions.

- **De jure:** Allow in some way, other than MAR, for bias due to selection of those with missing data.
- **De facto:** Allow for switching to alternative treatment regime. Both of these are MNAR, but motivation and form is different.
- Guidance and recommendations from Regulators, suggest routine collection of post-withdrawal data leading to **Retrieved Dropout** analysis.  
But this requires post treatment withdrawal regime to match the required estimand.

# Beyond “Randomised-arm MAR”

Two main directions.

- **De jure**: Allow in some way, other than MAR, for bias due to selection of those with missing data.
- **De facto**: Allow for switching to alternative treatment regime. Both of these are MNAR, but motivation and form is different.
- Guidance and recommendations from Regulators, suggest routine collection of post-withdrawal data leading to **Retrieved Dropout** analysis.  
But this requires post treatment withdrawal regime to match the required estimand.
- Can we build an analysis that predicts what would happen without requiring withdrawn patients to continue on a restricted treatment regime?

## Probability models: Three main approaches

Y is data (observed and unobserved) R is indicator of whether Y is observed or not.

- Pattern mixture models

$$[Y|R][R]$$

# Probability models: Three main approaches

Y is data (observed and unobserved) R is indicator of whether Y is observed or not.

- Pattern mixture models

$$[Y|R][R]$$

- Selection models

$$[R|Y][Y]$$

These first two are equivalent, but expressed very differently.

# Probability models: Three main approaches

Y is data (observed and unobserved) R is indicator of whether Y is observed or not.

- Pattern mixture models

$$[Y|R][R]$$

- Selection models

$$[R|Y][Y]$$

These first two are equivalent, but expressed very differently.

- Latent variable models  $[Y|Z][R|Z][Z]$

These do not seem to impose assumptions, but they do.

Problem is that it is not clear what those assumptions are.

## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- For any specific pattern of dropout (Fixed R) there is a separate distribution of both observed and unobserved data for each pattern.

## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- For any specific pattern of dropout (Fixed  $R$ ) there is a separate distribution of both observed and unobserved data for each pattern.
- Most interest has been in having separate models for the observed section.  
They can be modelled on the observed data.

## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- For any specific pattern of dropout (Fixed  $R$ ) there is a separate distribution of both observed and unobserved data for each pattern.
- Most interest has been in having separate models for the observed section.  
They can be modelled on the observed data.
- Then semi-flexible sharing of distributions between patterns implies conditional distributions for the post-deviation distribution.  
[Little, 1993 onwards, laid out the concept of **identifying restrictions**.]

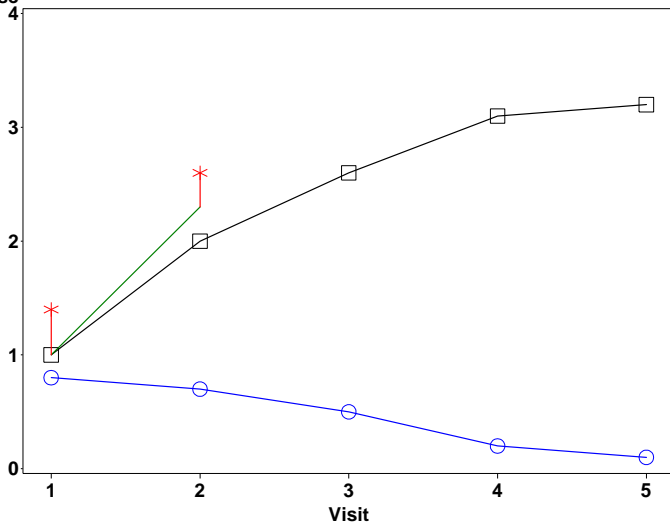
## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

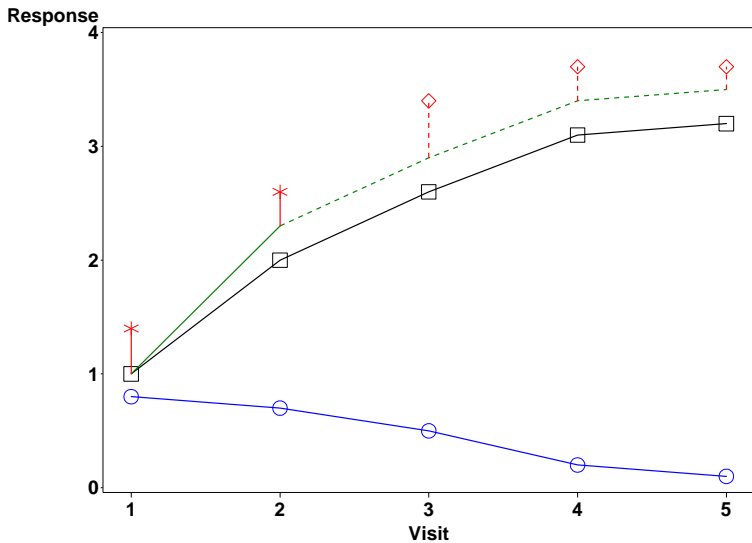
- For any specific pattern of dropout (Fixed  $R$ ) there is a separate distribution of both observed and unobserved data for each pattern.
- Most interest has been in having separate models for the observed section.  
They can be modelled on the observed data.
- Then semi-flexible sharing of distributions between patterns implies conditional distributions for the post-deviation distribution.  
[Little, 1993 onwards, laid out the concept of **identifying restrictions**.]
- Implementation is then possible via Multiple imputation using the conditional distribution for each subject, or by direct likelihood.  
[See Thijs et al (2002) for details of the implementation.]

## Observed Residuals from possible pattern mixture model

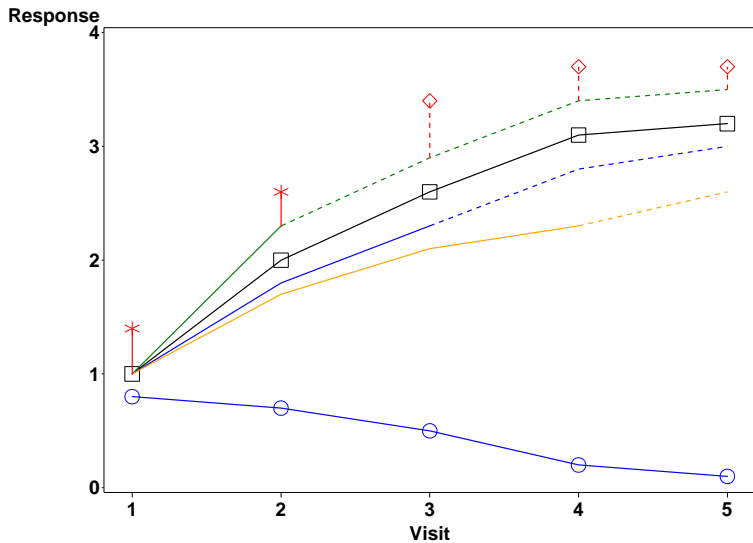
Response



## Projection using possible pattern mixture



## The pattern mixtures for this as-randomized arm



## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- Major limitation is interpretation of different **identifying restrictions**.

## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- Major limitation is interpretation of different **identifying restrictions**.
- Assumption is that data continue quite reasonably, but selection of those who are unobserved is related to the unobserved outcome.

## Pattern mixture models

Distribution of Y (Observed and Unobserved) conditional upon pattern of Missingness.

- Major limitation is interpretation of different **identifying restrictions**.
- Assumption is that data continue quite reasonably, but selection of those who are unobserved is related to the unobserved outcome.
- So most of the current Pattern Mixture models answer the **de jure** and not the **de facto** question.

## Pattern mixture models

Distribution of  $Y$  (Observed and Unobserved) conditional upon pattern of Missingness.

- Major limitation is interpretation of different **identifying restrictions**.
- Assumption is that data continue quite reasonably, but selection of those who are unobserved is related to the unobserved outcome.
- So most of the current Pattern Mixture models answer the **de jure** and not the **de facto** question.
- We feel that bias due to selection in answering the on-treatment question is often well controlled by conditioning in MAR analyses. The real question is rather to provide a **de facto** answer.

## Pattern mixture models

Distribution of Y (Observed and Unobserved) conditional upon pattern of Missingness.

- Major limitation is interpretation of different **identifying restrictions**.
- Assumption is that data continue quite reasonably, but selection of those who are unobserved is related to the unobserved outcome.
- So most of the current Pattern Mixture models answer the **de jure** and not the **de facto** question.
- We feel that bias due to selection in answering the on-treatment question is often well controlled by conditioning in MAR analyses. The real question is rather to provide a **de facto** answer.
- Note how multivariate Normal for each pattern implies the marginal distribution for treatment arm is not Multivariate Normal.

## Pattern mixture. Answering de facto.

- Interested in distribution of Unobserved  $Y$  conditional upon Observed  $Y$  and observed covariates.
- At **Deviation** something dramatic happens, such as end of active treatment, or return to standard therapy.

## Pattern mixture. Answering de facto.

- Interested in distribution of Unobserved  $Y$  conditional upon Observed  $Y$  and observed covariates.
- At **Deviation** something dramatic happens, such as end of active treatment, or return to standard therapy.
- So propose the post-deviance distribution by copying from other arms rather than the same arm as in MAR.

## Pattern mixture. Answering de facto.

- Interested in distribution of Unobserved  $Y$  conditional upon Observed  $Y$  and observed covariates.
- At **Deviation** something dramatic happens, such as end of active treatment, or return to standard therapy.
- So propose the post-deviance distribution by copying from other arms rather than the same arm as in MAR.

We specify this through the means (here are some **examples**):

- Increment in mean from visit to visit copies that in the reference arm
- Mean shifts dramatically to that in the reference arm.
- Mean stays constant after deviation.

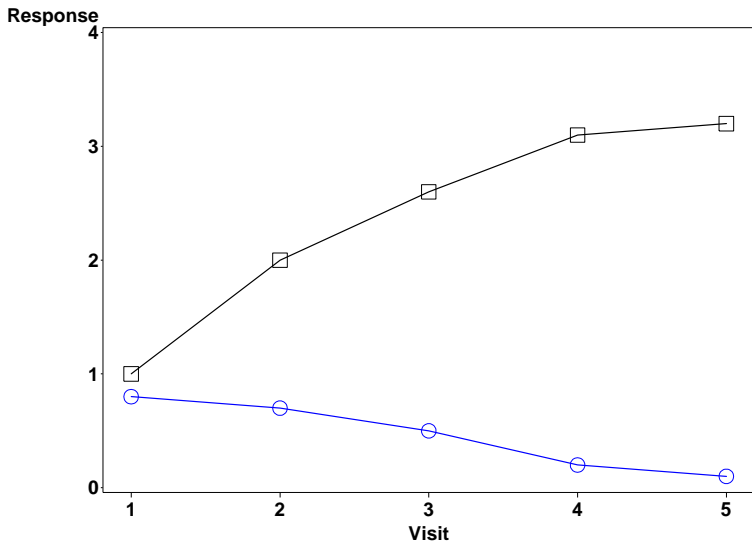
## Pattern mixture. Answering de facto.

- Interested in distribution of Unobserved Y conditional upon Observed Y and observed covariates.
- At **Deviation** something dramatic happens, such as end of active treatment, or return to standard therapy.
- So propose the post-deviance distribution by copying from other arms rather than the same arm as in MAR.  
We specify this through the means (here are some **examples**):
  - Increment in mean from visit to visit copies that in the reference arm
  - Mean shifts dramatically to that in the reference arm.
  - Mean stays constant after deviation.
- The modelled distribution before deviation is kept simple.

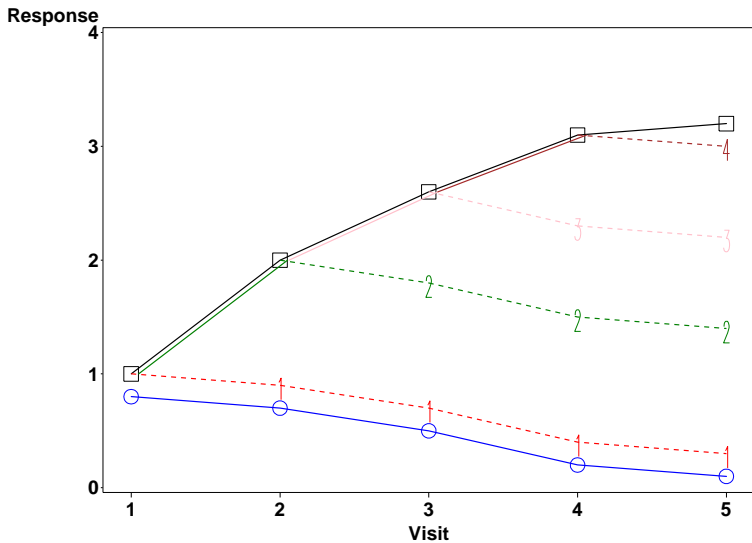
## Pattern mixture. Answering de facto.

- Interested in distribution of Unobserved Y conditional upon Observed Y and observed covariates.
- At **Deviation** something dramatic happens, such as end of active treatment, or return to standard therapy.
- So propose the post-deviance distribution by copying from other arms rather than the same arm as in MAR.  
We specify this through the means (here are some **examples**):
  - Increment in mean from visit to visit copies that in the reference arm
  - Mean shifts dramatically to that in the reference arm.
  - Mean stays constant after deviation.
- The modelled distribution before deviation is kept simple.
- Implemented via Multiple Imputation.

## Completers profiles for as-randomised and reference arms

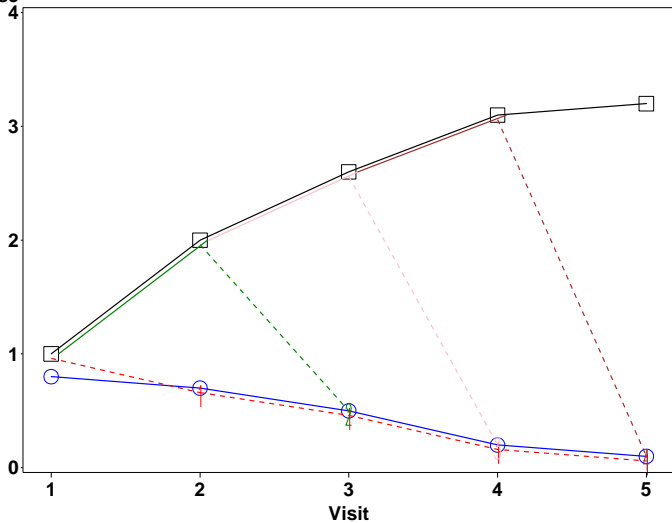


## Copy Increment in Reference



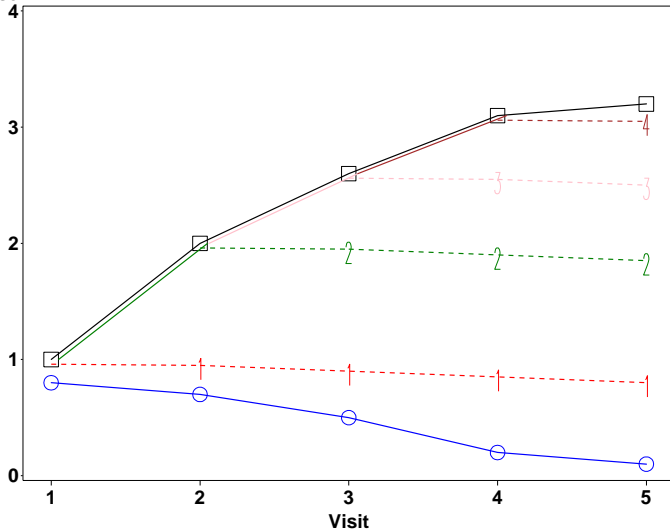
## Jump to Reference

Response

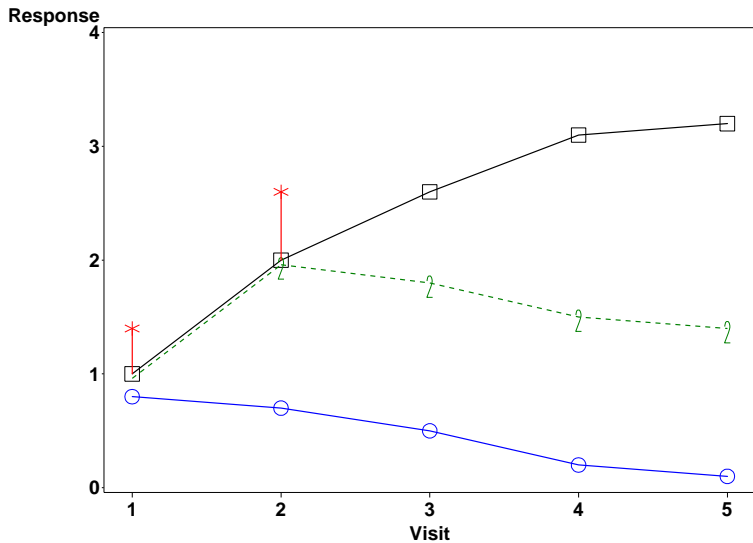


## Constant decline

Response

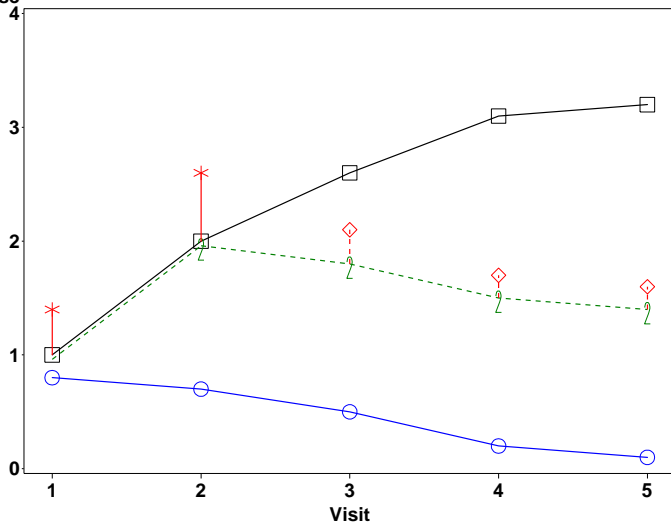


## Observed Residuals for an individual subject

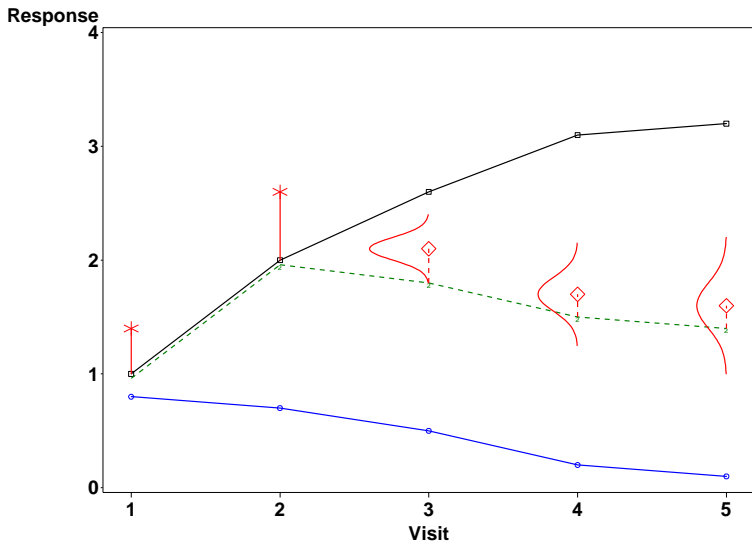


## Means for conditional distribution for missing data

Response



## Samples drawn from conditional distribution for missing data



## Implemented through Multiple Imputation.

- Bayesian Multivariate repeated measures model for observed data.
- Sample from the posterior distribution of the MV parameters, one set for each imputed data set.
- For each subject build conditional distribution for their missing data and take a single sample for each imputation.
- Carry out standard analysis (e.g. univariate ANCOVA).
- Combine results using Rubin's formula.

# Specifying a Pattern Mixture model with MI.

- There are three separate models being used in the process.
  - 1 Model for parameter estimation based on observed data, which may differ from pattern to pattern
  - 2 Model for deriving conditional distribution for each pattern  
This uses the estimated parameter values.
  - 3 Analysis model for the outcome.
- These do not need to be coherent with each other.
- Methods described above have (1) and (2) coherent, but (3) is not because the analysis model ignores differences between the patterns.

# Specifying a Pattern Mixture model with MI.

- There are three separate models being used in the process.
  - 1 Model for parameter estimation based on observed data, which may differ from pattern to pattern
  - 2 Model for deriving conditional distribution for each pattern  
This uses the estimated parameter values.
  - 3 Analysis model for the outcome.
- These do not need to be coherent with each other.
- Methods described above have (1) and (2) coherent, but (3) is not because the analysis model ignores differences between the patterns.
- Advantage of Multiple Imputation is that with no missing data, analysis is automatically “correct”.

# MI with non-coherent missingness

- This **de facto** approach was described at PSI conferences in 2008 and 2009.
- Implemented using MI and MIANALYZE procedures in SAS and MIXED procedure.
- Recent interest has spawned a series of papers and proposals. Most have (1), (2) and (3) all non-coherent.

B Ratitch, & M. O'Kelly, *Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures*, SUGI 2011.

# Copy Reference

- This approach we originally called **Copy Control**.
  - ① Parameter estimation is simple Multivariate Normal based on observed data (MAR effectively).
  - ② Missingness model has all incomplete patterns identical to the complete pattern for the Reference arm.
  - ③ Analysis model is simple ANCOVA for final time point.
- All three models are not coherent.

# Copy Reference

- This approach we originally called **Copy Control**.
  - ① Parameter estimation is simple Multivariate Normal based on observed data (MAR effectively).
  - ② Missingness model has all incomplete patterns identical to the complete pattern for the Reference arm.
  - ③ Analysis model is simple ANCOVA for final time point.
- All three models are not coherent.
- The missing data is imputed assuming that the subject has not responded to treatment at all, . . .  
. . . while the model parameter estimates for the active arm have included this subject's observed data.

# Implementation in SAS using proc MI

We developed a SAS macro that does the following:

- Use MI procedure to draw sample from the bayesian posterior for the multivariate Normal distribution parameters.  
Separately for each arm.  
Discard the imputed values.
- Use the IML procedure to obtain the conditional distribution for post-deviance values for each subject and draw one observation per imputation.
- Analyze each imputed data set using ANOVA.
- Use the MIANALZE procedure to provide combined results.

[Available from the registered user's area at  
<http://missingdata.lshtm.ac.uk>]

## Implementation in SAS NOT using proc MI

The use of the MI procedure restricts the possible models. It forces covariates, such as baseline, to be crossed with treatment in the parameter estimation model.

## Implementation in SAS NOT using proc MI

The use of the MI procedure restricts the possible models. It forces covariates, such as baseline, to be crossed with treatment in the parameter estimation model.

We are developing a SAS macro that does the following:

- Fit model (1) using the MCMC procedure and draw independent sample from posterior.
- Use DATA step and FCMP procedure to derive conditional distribution for missing data. Simulate missing data for each imputation.
- Analyze each imputed data set using ANOVA.
- Use the MIANALZE procedure to provide combined results.

[Still under development.]

## Back to the Example.

- **De jure:** Several different assumptions about model.
- **De facto:** Several different assumptions for post-withdrawal profile.

# Sensitivity Analysis for Treatment difference

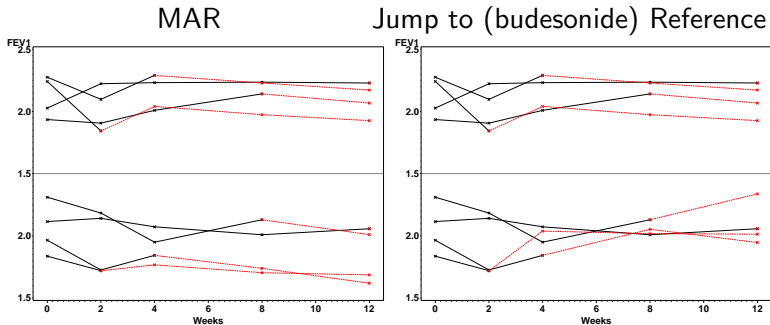
Analysis	Estimate	Std. Err.	DF (model)	t-statistic	P-value
<b>De jure</b>					
ANCOVA (Completers), joint variance	0.247	0.101	105	2.46	0.0155
ANCOVA (Completers), separate variances	0.247	0.101	70.6	2.46	0.0164
Mixed model, joint covar. matrix	0.283	0.094	131	3.02	0.0030
Mixed model, separate covar. matrices	0.346	0.104	72.8	3.34	0.0013
Multiple imputation (proc MI)	0.336	0.106	132.1	3.18	0.0019
Macro MAR	0.334	0.107	130.6	3.13	0.0022
<b>De facto</b>					
Macro, Last Mean Carried Forward	0.296	0.102	141.6	2.90	0.0043
Macro, Jump to Reference (Active)	0.141	0.119	102.7	1.18	0.2390
Macro, Jump to Reference (Placebo)	0.264	0.108	135.5	2.46	0.0153
Macro, Copy Reference (Active)	0.252	0.087	139.4	2.88	0.0046
Macro, Copy Reference (Placebo)	0.295	0.105	146.5	2.82	0.0055
Macro, Copy Increment in Ref. (Active)	0.295	0.103	139.7	2.87	0.0048
Macro, Copy Increment in Ref. (Placebo)	0.323	0.104	139.6	3.12	0.0022

[Based on 1000 imputations.]

## What if the conclusions are different?

- Are they asking the same question?
- Explore why different post-deviance assumptions lead to different conclusions.
- Is it just different levels of variation or a difference in size of effect?
- For each subject with missing data take their projected mean, conditional upon their previous experience and covariates by averaging over multiple imputations.
- Plot by pattern within treatment arm just like for raw summary. Can be used to compare to single imputation methods as well. Ignores the variability aspect.

# Use averages of the imputed data to generate plots.



Top: Budesonide  
Bottom: Placebo

# Conclusions

- We suggest that Sensitivity analysis needs to address the **de facto** question, rather than simply regarding data as unobserved.
- Relatively easy to compute using Multiple Imputation.
- Can use different assumptions for different type of withdrawal.
- Definition of the post-deviation profile is easy to explain.
- Able to respond to different disease scenarios.

## “The down side”!

- No agreed way of describing what we are doing.
- A series of ad hoc methods appearing in the literature based around computing solution rather than model specification.

## References

- Little, R.J.A (1993) Pattern-mixture models for multivariate incomplete data. **Journal of the American Statistical Association**, 88:125–134.
- Little, R. J. A., Yau, L. (1996) Intent-to-treat analysis for longitudinal studies with drop-outs. **Biometrics**, 52:471–483.
- Thijs, H. et al (2002) Strategies to fit pattern-mixture models. **Biostatistics**, 3:245–265.

# Questions