PSI's Pharmaceutical Statistics Journal Club

15th October 2015

Missing Data

Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models

Authors: Bohdana Ratitch, Michael O'Kelly, and Robert Tosiello

Pharmaceutical Statistics, Volume 12, Issue 6, Pages 337-347, November/December 2013

SPEAKER: BOHDANA RATITCH (QUINTILES)

Motivation and Regulatory Context

- "*Guideline on Missing Data in Confirmatory Trials*" (2010), European Medicines Agency (EMA)
- "The Prevention and Treatment of Missing Data in Clinical Trials" (2010), report by National Research Council (NRC) panel commissioned by Food and Drug Administration (FDA)

- Emphasis on prevention of missing data
  - not just a statistician's problem that can be fixed during analysis
- Inevitability of some amount of missing data
  - necessity for an upfront planning of a comprehensive analysis strategy

Motivation and Regulatory Context

Most assumptions about missing data are 'untestable', but

- must be 'reasonable' in a context of a given study (EMA)
- must be 'thoroughly justified' (NRC)
- 'assumptions need to be expressed in as transparent a manner as possible so that researchers and practicing clinicians are able to assess their validity' (NRC)
- 'there is no single. . . method' for handling missing data (EMA and NRC)
  - primary analysis should be 'unlikely to be biased in favor of the experimental treatment to an important degree' (EMA)
  - 'Sensitivity analyses are . . . important to assess the degree to which the treatment effects rely on the assumptions used' (NRC)
  - 'an upfront investigation of different missing data handling methods under different assumptions, ranging from pessimistic to realistic to optimistic. . . should be carried out' (EMA).

What We Did and Didn't Address

- We were mainly addressing the issue of missing data from subjects who discontinue permanently (monotone missing) and in the context of superiority testing
- We do not discuss discontinuation from treatment vs discontinuation from study and do not discuss merits of using post-treatment discontinuation data in analysis
  - It's important, but we see it as a separate topic
- Confession: our paper does not contain the word "estimand"
- Proposed analysis strategies are presented mainly as sensitivity analyses for several specific departures from a Missing at Random (MAR) assumption
  - MAR implies: subjects who discontinue would have similar outcomes to subjects in their arm, given observed data prior to discontinuation
  - However, similar strategies can be applied to assess departures from other "base" assumptions
- Our presentation was in terms of continuous endpoints, but similar ideas apply to other types
- We did not aim to make comparisons with other frameworks, e.g., selection or shared parameter models

Pattern Mixture Models: Motivation

- One of our objectives: simplicity of implementation while allowing for clinically plausible and transparent assumptions
- PMMs can offer a straightforward interpretation as imposing selected assumptions about missing values on clearly-defined categories of subjects (patterns).

Pattern Mixture Models: General Form

Y_{obs} : observed data; Y_{mis} : missing data; X: covariates; R : response/missingness indicators

$$p(Y_{obs}, Y_{mis}, R|X) = p(R|X) \times p(Y_{obs}, Y_{mis}|R, X)$$

- Parameters of $p(Y_{obs}, Y_{mis}| R, X)$ cannot be estimated from the available data alone
- Separate observed data distribution and a predictive distribution of missing data given observed data

$$p(Y_{obs}, Y_{mis}, R|X) = p(R|X) \times p(Y_{obs}|R, X) \times p(Y_{mis}|Y_{obs}, R, X)$$

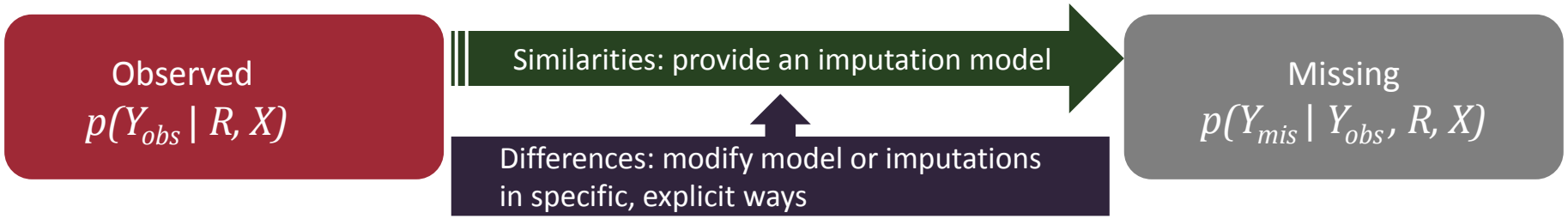
- Impose explicit restrictions/assumptions on $p(Y_{mis} | Y_{obs}, R, X)$

Little RJA. Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association 1993; 88: 125–134.

Pattern Mixture Models: Identifying Restrictions

$$p(Y_{obs}, Y_{mis}, R|X) = p(R|X) \times p(Y_{obs}|R,X) \times p(Y_{mis}|Y_{obs},R,X)$$

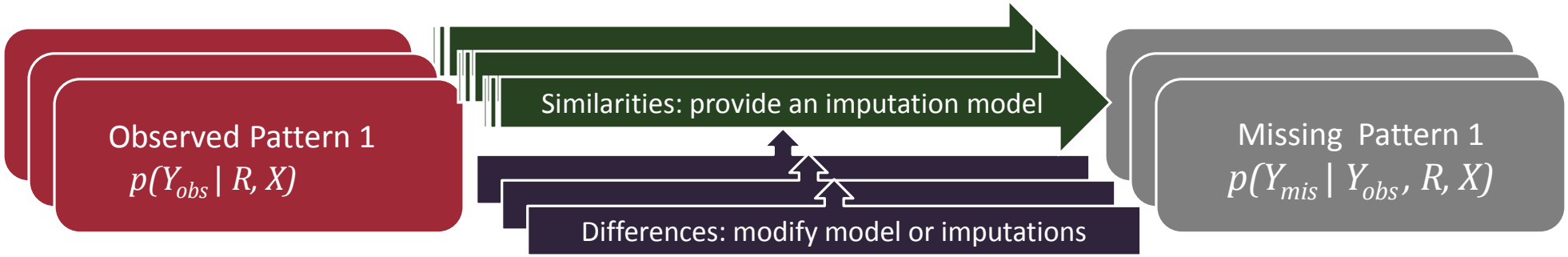
- Restrictions/assumptions on $p(Y_{mis} \mid Y_{obs}, R, X)$ can be expressed in terms of similarities and differences between $p(Y_{mis} \mid Y_{obs}, R, X)$ and $p(Y_{obs} \mid R, X)$



Thijs H, Molenberghs G. Strategies to fit pattern-mixture models. *Biostatistics* 2002; **3**(2):245–265.

Pattern Mixture Models: Identifying Restrictions

- Multiple patterns can be defined and different links between them postulated



- Patterns can be defined based on a combination of factors, e.g.,:
  - Treatment arm
  - Completion/time of discontinuation
  - Reason for discontinuation
- Multiple Imputation methodology can be used for implementation

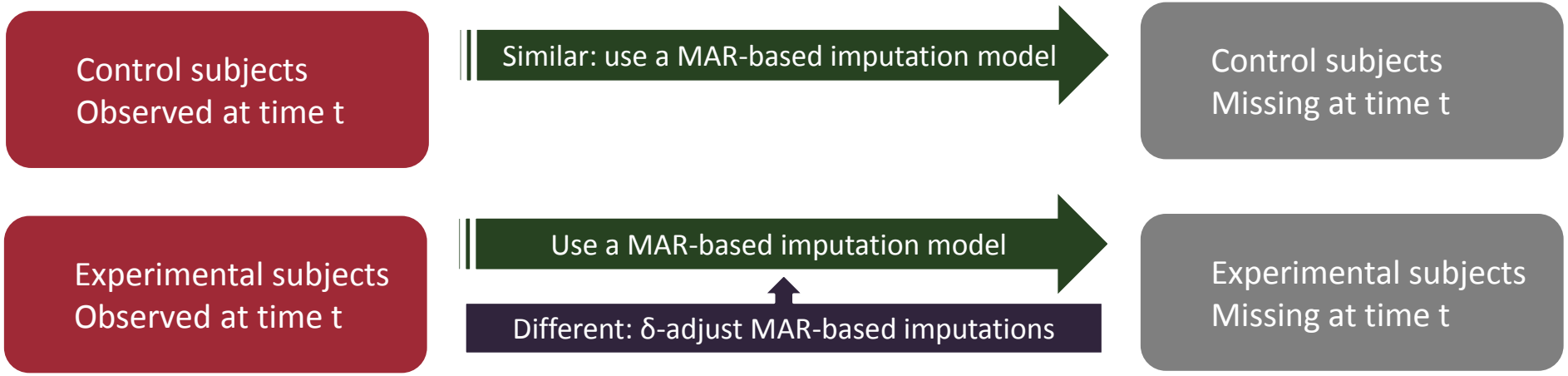
Multiple Imputation (MI)

- MI (Rubin, 1987) is a principled method that accounts for uncertainty of imputations
- Each missing value is replaced with multiple imputed values
- Imputed values are generated by an imputation model estimated based on observed data
  - Uncertainty about the imputation model is accounted for through using multiple draws of model parameters from a Bayesian posterior predictive distribution for each imputed dataset
- Each imputed dataset is analyzed separately using an analysis model
  - Imputation model can be different from analysis model – very helpful for PMMs
- Results are combined using Rubin's rules taking into account uncertainty (variability) of imputations
- "Standard" MI operates under MAR, but MI can be used under departures from MAR

Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley and Sons, Inc.: New York, 1987.

Imputation with Delta-Adjustment

- Assumption: Subjects who discontinue at a given time-point would have, on average, their unobserved efficacy score after discontinuation worse by some amount δ compared with the observed efficacy score of subjects who continue to the next time-point.
  - δ can have a clear and direct clinical meaning in terms of efficacy measure
  - It may be reasonable to assume worsening only in the experimental arm
  - Discontinued subjects from the control arm can be assumed to have the same evolution of the disease as control subjects who stay on study



Imputation with Delta-Adjustment: Variants

- **First visit after withdrawal***:
  δ is applied just once, at the first visit after withdrawal for each subject and imputed δ-adjusted value is then used as predictor for the imputation of subsequent time-points
- **All visits after withdrawal, sequentially**:
  δ is applied at each visit, from withdrawal up to the end of the study, in addition to using previously δ-adjusted values as predictors
- **All visits after withdrawal, after performing a complete MAR imputation**:
  δ is applied at each visit, from withdrawal up to the end of the study, after all values are imputed based on MAR

- Delta-adjustment can be applied as a mean shift (dampening) or mean scale for continuous endpoints
- Delta can vary, e.g., depending on time of discontinuation
- Applicable to other types of endpoints with appropriate interpretation of δ

*National Research Council. Panel on Handling Missing Data in Clinical Trials. *The prevention and treatment of missing data in clinical trials*. The National Academies Press: Washington, DC, 2010.

Tipping Point Analysis

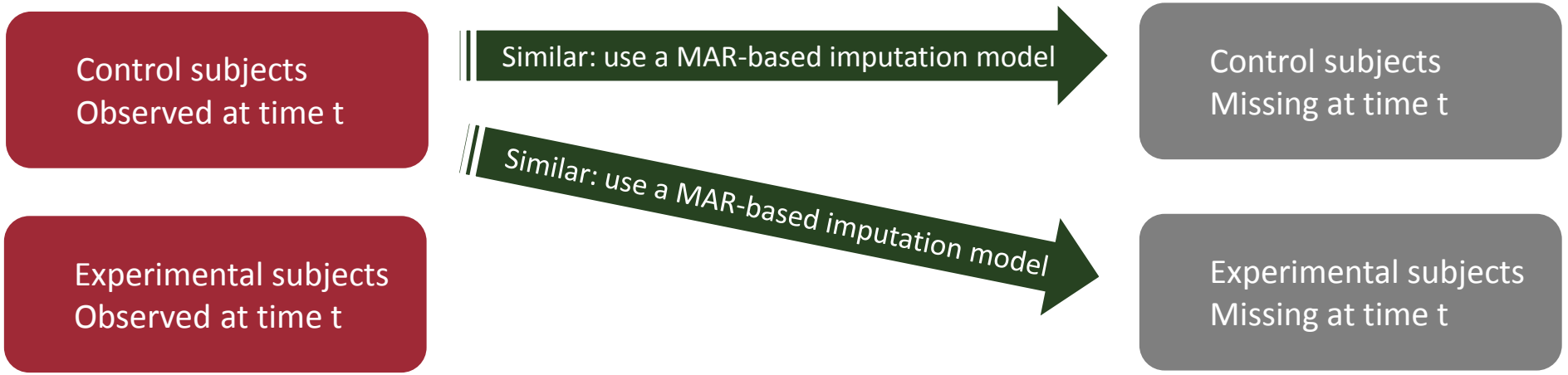
- Perform a sequence of analyses assuming progressively worse values/adjustments for discontinued subjects
- Find a “tipping point” – assumptions/adjustments under which the result is no longer significant (statistically and/or clinically)
- Apply clinical judgment about the plausibility of the assumptions underlying the tipping point

  For example:
  - All discontinued subjects in the experimental arm would need to have HAM-D-17 total score worse by 4 points compared to completers to lose statistical significance
    - Not a plausible assumption – primary analysis results can be considered robust
  - All discontinued subjects in the experimental arm would need to have HAM-D-17 total score worse by 1 point compared to completers to lose statistical significance
    - A plausible assumption – primary analysis results may be questionable

Yan X, Lee S, Li N. Missing data handling methods in medical device clinical trials. Journal of Biopharmaceutical Statistics 2009; 19:1085–1098.

Control-Based Imputation

- Assumption:
  - Subjects from the experimental treatment arm, after withdrawal (no longer receiving the experimental treatment), will exhibit the same future evolution of the disease as subjects on the control treatment (who are also not exposed to the experimental treatment).
  - Subjects who discontinue from the control arm will evolve in the same way as control subjects who remain in the study.



Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics 1996; 52:1324–1333.

Control-Based Imputation

- This approach will tend to provide a reduced estimate of treatment effect.
  - In some contexts, it can be considered as “the worst clinically plausible” scenario.
- This strategy does not push the analysis into a more extreme scenario where subjects who discontinue from the experimental arm could be doing worse than control subjects.
- Our implementation uses control group outcomes to model outcomes of those who withdraw from the experimental arm; differences from group means in pre-discontinuation data (serving as predictors) are taken with respect to control means
  - Other variants of control-based imputation have been proposed in the literature (Carpenter, Roger, &Kenward; 2013)

Carpenter J, Roger J, Kenward M. Analysis of longitudinal trials with protocol deviation:- a framework for relevant, accessible assumptions and inference via multiple imputation. Journal of Biopharmaceutical Statistics, 2013; 23(6):1352-71

Imputation Based on Reasons for Discontinuation

- Assumptions can vary depending on reasons for discontinuation, e.g., withdrawal due to Adverse Events (toxicity) vs. other reasons
  - See paper for details

Example Dataset

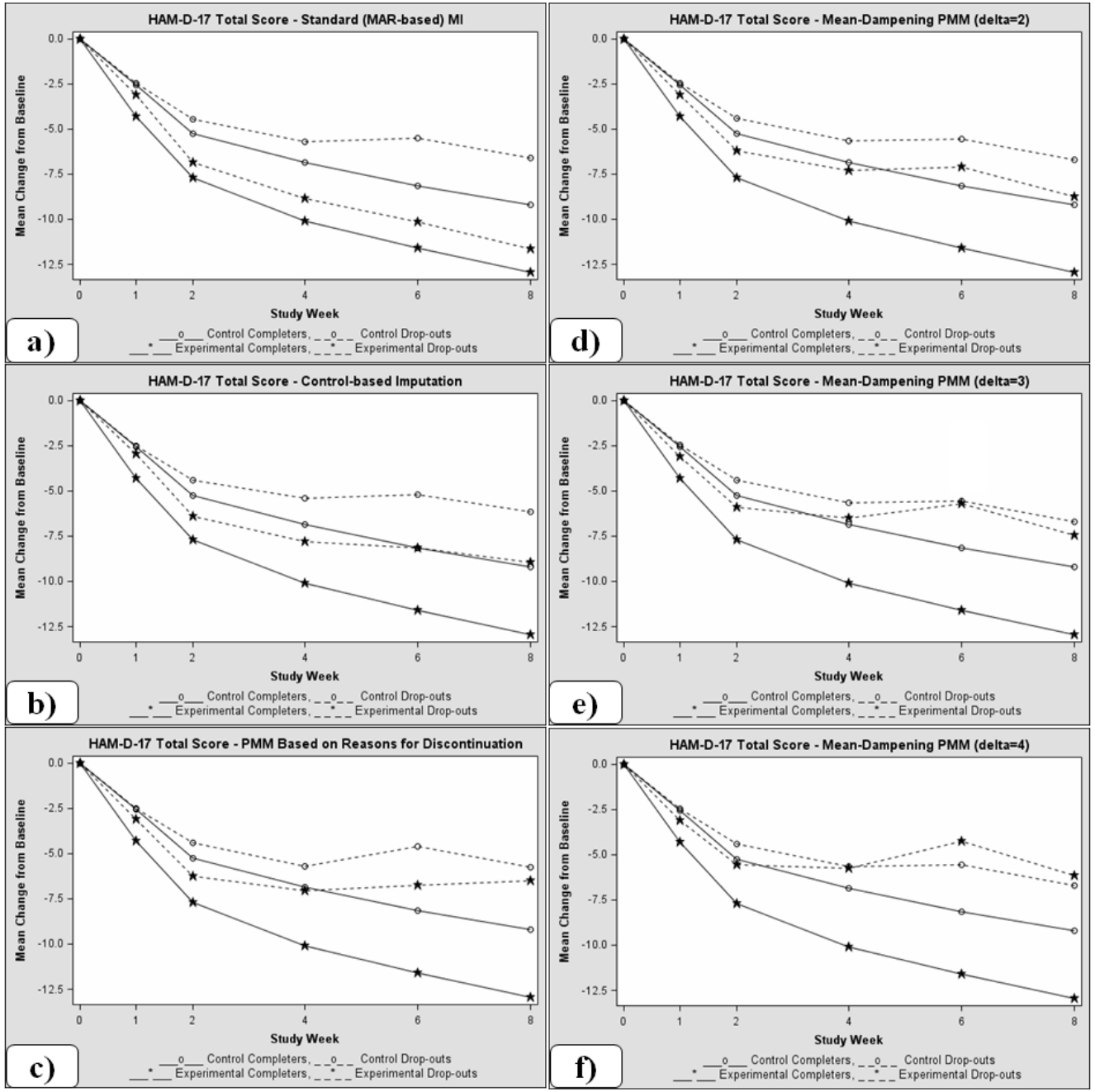
- Dataset patterned after a clinical trial in major depressive disorder
- 2 treatment arms: placebo and experimental
- Primary endpoint: Hamilton Depression Scale, HAM-D-17 total score (lower values are better)
- Time points: Baseline, Weeks 1, 2, 4, 6, and 8
- Study completion rates: Placebo = 77%; Experimental = 63%
- Discontinuations due to AEs: Placebo = 4%; Experimental = 12 %
- Secondary parameter, Brief Psychiatric Rating Scale, was a significant predictor of drop-out

Results

Results from Analyses Using Standard MI and PMM-based Approaches Change from Baseline to Week 8 in HAM-D-17 Total Score		
Analysis	**LS Mean Difference of Experimental vs. Placebo (CIs)**	**p-value**
Standard MI (MAR-based) [a]	-4.096 (-6.143, -2.050)	0.0001
Mean-Dampening PMM: δ=2 [d]	-2.995 (-5.157, -0.834)	0.0066
Mean-Dampening PMM: δ=3 [e]	-2.504 (-4.713, -0.295)	0.0263
Mean-Dampening PMM: δ=4 [f]	-2.009 (-4.273, 0.256)	0.0821
Control-Based PMM [b]	-3.196 (-5.277, -1.114)	0.0026
PMM Based on Reasons for Discontinuation [c]	-2.395 (-4.602, -0.187)	0.0335

- Treatment difference often assumed at design stage is 2 points
- In general, it's not expected that all sensitivity analyses will maintain statistical significance
- Control-based PMM remains statistically and clinically significant
- Change of 1/2 SD (~2 points) in HAM-D-17 total score is considered meaningful for individual subject*, yet with δ=2 and 3, statistical and clinical significance holds
- Tipping point (based on loss of statistical significance) is reached at δ=4, thus results can be considered robust

*Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Medical Care 2003; 41(5):582–592.



Summary

- PMMs enable us to identify directly a predictive distribution (or set of predictive distributions) for the missing outcomes
- Assumptions about missing outcomes used in PMMs can be formulated in a transparent way and facilitate a dialogue between statisticians and clinicians
- We need to select justifiable assumptions for primary analysis, then assess robustness through sensitivity analyses that postulate specific, clinically interpretable departures from primary assumptions

- Acknowledgments: James Roger, Gary Koch, paper reviewers

- Discussion
  - Questions
  - Comments
  - Experience within your organizations
  - Regulatory experience

Extra Slides

Imputation Based on Reasons for Discontinuation

- In some therapeutic areas (e.g., symptom relief in chronic diseases), no therapeutic benefit is attributed to subjects discontinuing due to AEs (toxicity).
- "Hybrid" imputations used in the past:
  - Return-to-baseline (via BOCF) for withdrawals due to AE
  - Other methods for withdrawals due to other reasons
- A clinically reasonable approach, but criticized if implemented using single imputation (BOCF).

Imputation Based on Reasons for Discontinuation

- Can be implemented using multiple imputation within PMM framework