A structured framework for assessing sensitivity to missing data assumptions in longitudinal clinical trials

C. H. Mallinckrodt, Q. Lin ,G. Molenberghs

Pharmaceut. Statist. 2013, 12 1–6

Outline

- **Background**
- **Analytic Road Map**
- **Example**

History

- Lots of recent research
- Research fostered updated guidance
- Considerable agreement between NRC guidance PhRMA Expert group, CHMP Points to consider
- We largely agree on what to do – in theory
- Now we need to bring the theory into practice

3 Pillars of Dealing with Missing Data

- **Clear objectives and causal estimands**
- **Limit missing data**
- **Sensible primary analysis supported by sensitivity analyses**
  - **Methods driven by plausible scientific assumptions**
  - **Sensitivity analyses assess robustness to departures from assumptions**

Acknowledgements

- **DIA Scientific Working Group**
- **Programs freely available at missingdata.org.uk**

- **Specific thanks**
  - **Michael O'Kelly** Quintiles Dublin Ireland
  - **Bohdana Ratitch** Quintiles Saint-Laurent, Québec, Canada,
  - **James Roger** London School of Hygiene &Tropical Medicine
  - **Pierre Bunouf** Laboratoires Pierre Fabre, Toulouse, France

Outline

- **Background**
- **Analytic Road Map**
- **Example**

Missing Data in Clinical Trials

- **Efficacy outcomes are seldom MCAR because the observed outcomes typically influence dropout (DC for lack of efficacy)**
- **Trials are designed to observe all the relevant information, which minimizes MNAR data**
- **Hence in the highly controlled scenario of longitudinal confirmatory trials, missing data *may* be mostly MAR**

Modeling Conundrum

- **Can't assume MCAR**
- **We don't have the missing data about which the assumptions are made, Therefore…**
- **Validity of MAR can't be verified; i.e., MNAR can not be ruled out**
- **Key assumptions in MNAR models can't be verified**

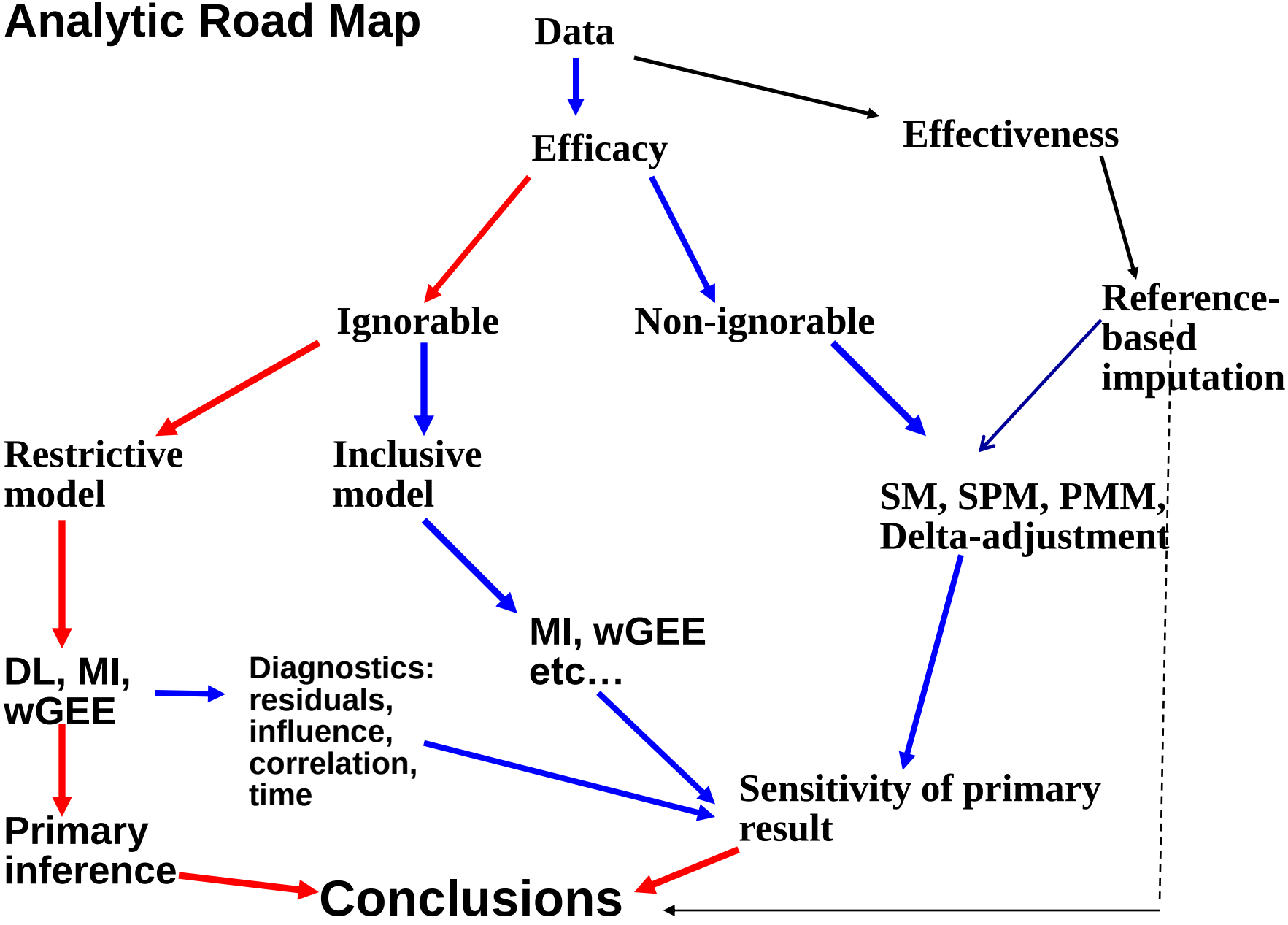
Modeling Considerations

- **Can do better than MAR only via assumptions**
- **Sensitivity to violations of assumptions and model misspecification more severe in MNAR**
- **No individual MNAR analysis is definitive**

General Guidance

- **Strive for validity of MAR primary analysis**
- **Implement MNAR model(s) under varying but plausible assumptions**
- **Compare results from assuming MAR vs. various MNAR implementations**

Analytic Road Map



Selection Models

- **Joint model**
  - **Primary analysis (same or similar model)**
  - **logistic regression for probability of dropout**
  - **Dropout and measurement models linked as primary outcome is predictor of dropout**

Pattern Mixture Models

- **Assesses the outcome variable separately for different groups (patterns), often defined by time of drop-out, and then combines results across groups for final inference**
- **Assesses consequences when distribution of unobserved data ≠ what it would have been if observed**

Controlled Imputations: New Methods in the PMM Framework

- Assumptions can be transparent and debated
- General idea is to create departures from MAR
  - Plausible worst case
  - Progressive stress test
- Based on standard MI
  - Reference-based
  - Delta adjustment

Controlled Imputations: Reference Based

- **Jump to reference**
  - **The statistical behavior of drug treated patients after dropout immediately becomes that of reference patients" (e.g., placebo)**
    - **Use for drugs with short on target half-life**
- **Copy reference**
  - **…gradually transitions to placebo**
    - **Use for drugs with long on target half-life**
- **Copy increment**
  - **After dropout, change for drug = change for placebo**
    - **Use for disease modifying drugs**

Mechanics of Reference-based Imputations

- **Similar to "Standard" MI, except**
  - **Parameters for imputation model obtained from only the reference (control/placebo) group**
  - **Missing data for both reference and drug group are imputed based on the imputation model derived from placebo data**

Contexts for Reference-based Imputations

- **Effectiveness context**
  - **Assumes benefit diminishes / disappears**
  - **Accounts for study effect & placebo effect**
  - **Valid if patients improve or worsen**
  - **Free of confounds in follow up data**
- **Efficacy context**
  - **Worst reasonable case MNAR**
- **Standard software, standard tests**

Controlled Imputations: Delta Adjustment

- Visit-by-Visit
  - Subtract a constant (delta) from visit X imputed value that then further influences imputed values at visit > X
    - First missing visit only
    - All missing visits
    - Progressively increase delta until conclusion from primary analysis is overturned
- After completion of all imputations

Outline

- **Background**
- **Analytic Road Map**
- **Example**

Example Data from Major Depression

- Two real but contrived data sets (n=100/arm)
- Drug arm patients randomly selected from 3 active arms
- Placebo arms mostly as is (with minor replication)
- Nearly identical designs
  - 8-week, double blind, randomized 1:1:1:1
  - Assessments @ weeks 1,2,4,6,8
  - Similar inclusion / exclusion
  - Low dropout from EU study with ext and titration
  - High dropout from US fixed dose, no extension

Example Data

- **High dropout**
  - **Completion rates: placebo 60%, drug 70%**
  - **1000 planned observations, 850 available**
- **Low dropout**
  - **Completion rates: placebo 92%, drug 92%**
  - **1000 planned observations, 961 available**

Example From A Depression Trial

- **Primary objective: test the difference between drug and placebo in mean change to planned endpoint**
  - ***de-jure* (efficacy) hypothesis**
  - **… when taken as directed**
- **Secondary objective:**
  - ***de-facto (effectiveness)* hypothesis**
  - **… as actually taken**

Patient Retention

	Week				
	1	2	4	6	8
High dropout					
Placebo	8	7	12	13	60
Drug	9	6	10	5	70
Low Dropout					
Placebo	2	0	3	3	92
Drug	2	1	2	3	92

Prior to week 8 these values are the number of dropouts. For week 8 the values are the number of completers because week 8 is the last scheduled assessment.

Primary Analysis: Direct Likelihood

```
proc mixed;
    class subject treatment time site;
    model Y  = baseline treatment time site
                  baseline*time treatment*time site*time;
    repeated time /  sub = subject type = un;
    lsmeans treatment*time / cl diff;
    run;
```

Random effects modeled as part of within subject errors.

Full multivariate model: unstructured modeling of time and correlation.

Primary Results: Endpoint Contrast

Data	LSMEAN Change Drug	Placebo	Diff	SE	P value
High	**8.24**	**5.94**	***2.29***	**1.00**	**0.024**
Low	**12.32**	**10.50**	***1.82***	**0.70**	**0.010**

Correlation Sensitivity Results

CORR	AIC	Estimate	StdErr	P Value
High dropout				
UN	4679.82	**2.2928**	1.0024	0.0240
UN EMPIRICAL	4679.82	2.2928	0.9794	0.0202
TOEPH	4684.44	2.1003	0.9148	0.0231
TOEPH EMPIRICAL	4684.44	2.1003	0.9278	0.0239
TOEPH GROUP=TRT	4689.88	1.8207	0.9139	0.0482
UN GROUP=TRT	4692.05	1.9622	1.0059	0.0535
CSH	4735.81	1.8689	0.9330	0.0471
CSH EMPIRICAL	4735.81	1.8689	0.9169	0.0419
CSH GROUP=TRT	4739.34	**1.6973**	0.9323	0.0708
Low dropout				
UN GROUP=TRT	4861.70	**1.8535**	0.7034	0.0092
UN	4867.68	1.8150	0.6995	0.0103
UN EMPIRICAL	4867.68	1.8150	0.6669	0.0071
-	.	.	.	.
CSH	5030.40	**1.7653**	0.7054	0.0132

Range in estimates nearly 7x greater from high dropout

Selection Model Results Summary

Range of outcomes from endpoint contrast in selection model results using "plausible MNAR" inputs

	Endpoint Contrasts	
Data	**Lowest plausible**	**Highest plausible**
High dropout	**1.30**	**3.60**
Low dropout	**1.71**	**2.01**

Range in estimates nearly 8x greater in high dropout

Selection Model Results – *high dropout*

Input Values Ψ_5[1]	Ψ_6[1]	Week 8 LSMEANS Placebo	Drug	Endpoint Contrast	Standard Error	P value
0.2	0.2	4.87	7.33	2.46	1.09	0.023
0.0	0.2	5.60	7.38	1.78	1.05	0.091
-0.2	0.2	6.28	7.41	1.18	1.05	0.282
-0.4[2]	0.2[2]	6.76	7.42	0.66	1.06	0.527
0.2	0.0	4.94	7.97	3.03	1.07	0.005
0.0	**0.0**	**5.63**	**8.00**	**2.37**	**1.04**	**0.022**
-0.2	0.0	6.29	8.04	1.75	1.02	0.087
-0.4	0.0	6.75	8.05	1.30	1.02	0.204
0.2	-0.2	4.97	8.57	3.60	1.06	0.001
0.0	-0.2	5.67	8.57	2.89	1.03	0.004
-0.2	-0.2	6.31	8.59	2.29	1.01	0.024
-0.4	-0.2	6.76	8.63	1.86	1.01	0.064
0.2[2]	-0.4[2]	4.97	8.97	4.01	1.07	<0.001
0.0	-0.4	5.68	8.96	3.28	1.03	0.002
-0.2	-0.4	6.33	8.98	2.64	1.01	0.009
-0.4	-0.4	6.78	9.01	2.22	1.01	0.027

1. Ψ_5 and Ψ_6 are the regression coefficients (placebo and drug, respectively) for the association between the current, possibly missing efficacy scores and the logit for probability of dropout

2. This combination of values is not plausible based on previous experience but is included for completeness of illustration.

Selection Model Results - *high dropout*

Input Values Ψ_5[1]	Ψ_6[1]	Week 8 LSMEANS Placebo	Drug	Endpoint Contrast	Standard Error	P value
0.2	**0.2**	**4.87**	**7.33**	**2.46**	**1.09**	**0.023**
0.0	0.2	5.60	7.38	1.78	1.05	0.091
-0.2	0.2	6.28	7.41	1.18	1.05	0.282
-0.4[2]	0.2[2]	6.76	7.42	0.66	1.06	0.527
0.2	0.0	4.94	7.97	3.03	1.07	0.005
0.0	**0.0**	**5.63**	**8.00**	**2.37**	**1.04**	**0.022**
-0.2	0.0	6.29	8.04	1.75	1.02	0.087
-0.4	0.0	6.75	8.05	1.30	1.02	0.204
0.2	-0.2	4.97	8.57	3.60	1.06	0.001
0.0	-0.2	5.67	8.57	2.89	1.03	0.004
-0.2	**-0.2**	**6.31**	**8.59**	**2.29**	**1.01**	**0.024**
-0.4	-0.2	6.76	8.63	1.86	1.01	0.064
0.2[2]	0.4[2]	4.97	8.97	4.01	1.07	<0.001
0.0	-0.4	5.68	8.96	3.28	1.03	0.002
-0.2	-0.4	6.33	8.98	2.64	1.01	0.009
-0.4	**-0.4**	**6.78**	**9.01**	**2.22**	**1.01**	**0.027**

- With **equal negative (positive) values** for Ψ_5 and Ψ_6 the within group mean changes were **greater (less)** than from MAR

- Little impact on endpoint contrasts with equal values for Ψ_5 and Ψ_6. But with more dropout on placebo, slight decrease in endpoint contrast with equal negative values for Ψ_5 and Ψ_6

Selection Model Results - *high dropout*

Input Values Ψ_5[1]	Ψ_6[1]	Week 8 LSMEANS Placebo	Drug	Endpoint Contrast	Standard Error	P value
0.2	0.2	4.87	7.33	2.46	1.09	0.023
0.0	0.2	5.60	7.38	1.78	1.05	0.091
-0.2	0.2	6.28	7.41	1.18	1.05	0.282
-0.4[2]	0.2[2]	6.76	7.42	0.66	1.06	0.527
0.2	**0.0**	**4.94**	**7.97**	**3.03**	**1.07**	**0.005**
0.0	**0.0**	**5.63**	**8.00**	**2.37**	**1.04**	**0.022**
-0.2	**0.0**	**6.29**	**8.04**	**1.75**	**1.02**	**0.087**
-0.4	**0.0**	**6.75**	**8.05**	**1.30**	**1.02**	**0.204**
0.2	-0.2	4.97	8.57	3.60	1.06	0.001
0.0	-0.2	5.67	8.57	2.89	1.03	0.004
-0.2	-0.2	6.31	8.59	2.29	1.01	0.024
-0.4	-0.2	6.76	8.63	1.86	1.01	0.064
0.2[2]	-0.4[2]	4.97	8.97	4.01	1.07	<0.001
0.0	-0.4	5.68	8.96	3.28	1.03	0.002
-0.2	-0.4	6.33	8.98	2.64	1.01	0.009
-0.4	-0.4	6.78	9.01	2.22	1.01	0.027

- Negative Ψ means subjects with negative residuals more likely to withdraw, leading to **decreased** observed mean. The selection model compensates by increasing the LSMEAN.

- Corresponding change to endpoint contrast

Selection Model Results - *high dropout*

Input Values Ψ_5[1]	Ψ_6[1]	Week 8 LSMEANS Placebo	Drug	Endpoint Contrast	Standard Error	P value
0.2	0.2	4.87	7.33	2.46	1.09	0.023
0.0	**0.2**	**5.60**	**7.38**	**1.78**	**1.05**	**0.091**
-0.2	0.2	6.28	7.41	1.18	1.05	0.282
-0.4[2]	0.2[2]	6.76	7.42	0.66	1.06	0.527
0.2	0.0	4.94	7.97	3.03	1.07	0.005
0.0	**0.0**	**5.63**	**8.00**	**2.37**	**1.04**	**0.022**
-0.2	0.0	6.29	8.04	1.75	1.02	0.087
-0.4	0.0	6.75	8.05	1.30	1.02	0.204
0.2	-0.2	4.97	8.57	3.60	1.06	0.001
0.0	**-0.2**	**5.67**	**8.57**	**2.89**	**1.03**	**0.004**
-0.2	-0.2	6.31	8.59	2.29	1.01	0.024
-0.4	-0.2	6.76	8.63	1.86	1.01	0.064
0.2[2]	-0.4[2]	4.97	8.97	4.01	1.07	<0.001
0.0	**-0.4**	**5.68**	**8.96**	**3.28**	**1.03**	**0.002**
-0.2	-0.4	6.33	8.98	2.64	1.01	0.009
-0.4	-0.4	6.78	9.01	2.22	1.01	0.027

- Negative Ψ means subjects with negative residuals more likely to withdraw, leading to an **decreased** observed mean. The selection model compensates by increasing the LSMEAN.

- Corresponding change to endpoint contrast

Pattern Mixture Model Results

High dropout

Identifying Restriction[1-3]	Endpoint Contrast	Standard Error	P value
ACMV	2.67	1.17	0.0224
CCMV	2.51	1.05	0.0166
NCMV	2.87	1.69	0.0895

1. ACMV = available case missing values
2. CCMV = complete case missing values
3. NCMV = neighboring case missing values

Low Dropout

Number of patients in patterns insufficient to estimate parameters

Reference-Based Imputation Results

	LSMEANS		LSMEAN	Std	
	Placebo	Drug	Difference[1]	Error	P value
High dropout					
MAR	-5.95	-8.24	2.29	1.00	0.024
J2R	-5.97	-7.57	1.60	0.99	0.110
CR	-5.96	-7.71	1.75	0.98	0.075
CIR	-5.95	-7.78	1.83	0.97	0.004
Low dropout					
MAR	-10.56	-12.40	1.84	0.70	0.009
J2R	-10.55	-12.26	1.71	0.70	0.016
CR	-10.55	-12.27	1.72	0.70	0.015
CIR	-10.55	-12.27	1.72	0.70	0.015

Difference MAR vs. J2R over 6x greater in high dropout

Delta-adjustment Results

	Low Dropout			*High dropout*		
Value of Delta	Endpoint Contrast	Std Error	Pvalue	Endpoint Contrast	Std Error	PValue
0	1.85	0.71	0.009	2.31	1.02	0.024
0.5	1.77	0.71	0.013	2.00	1.03	0.051
2.0	1.52	0.73	0.037			
2.5	1.44	0.74	0.051			

- Delta required to overturn significance of primary result 5x larger for high dropout data set

- Change in endpoint contrast per 1 unit change in Delta
  - Low dropout = 0.16
  - High dropout = 0.62

Discussion

- **Proper sensitivity analyses allow us to *assess* sensitivity**
- **Lower rates of dropout is the only way to *improve* sensitivity**
- **Controlled imputations are useful and intuitive**