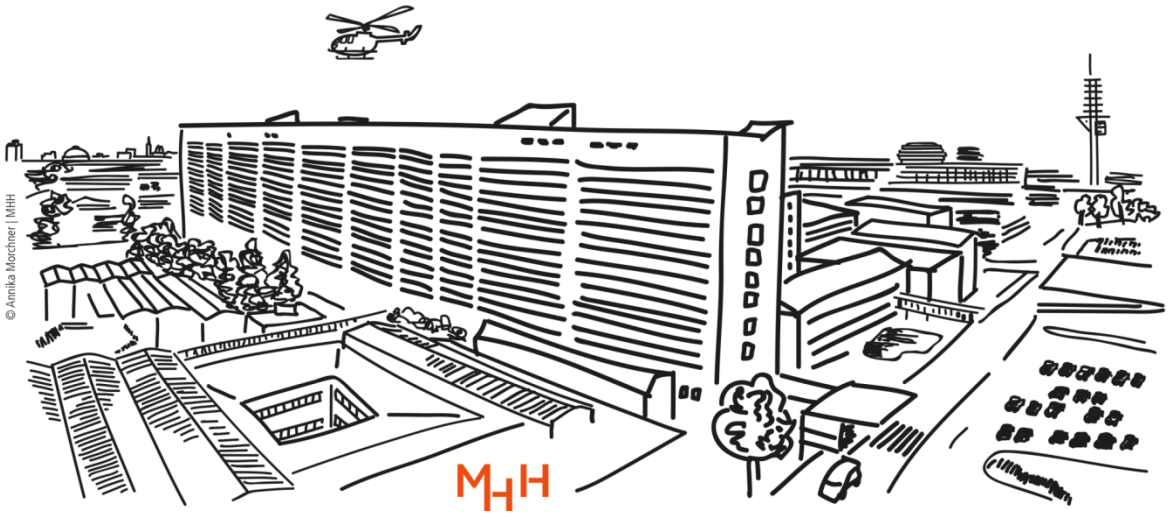
Empirical evaluation of the implementation of the EMA guideline on missing data in confirmatory clinical trials:

Specification of mixed models for longitudinal data in study protocols

Sebastian Häckl, Armin Koch and Florian Lasch



doi: 10.1002/pst.1964

Dr. Florian Lasch
Hannover Medical School
16.09.2020

Introduction

- Motivation: experience as a members of the Ethics Committee at Hannover Medical School

- Observation:
  - primary analysis not always fully specified (with or without reference to a SAP)
  - Strategies for handling missing values often not/vaguely defined
  - Multiplicity problem + type-I error rate inflation

→ Empirical study to investigate the quality of reporting of primary analysis models

- Focus on mixed models for longitudinal data because:
  (a) in longitudinal studies missing data are a common and relevant problem without a standard approach
  (b) mixed models are commonly used for handling missing data[1]
  (c) to allow a detailed assessment of model parameters

[1]: *Fletcher C, Tsuchiya S, Mehrotra DV. Current practices in choosing estimands and sensitivity analyses in clinical trials: results of the ICH E9 survey. Ther Innov Regul Sci. 2017;51(1):69-76.*

EMA Guideline on Missing Data in Confirmatory Clinical Trials

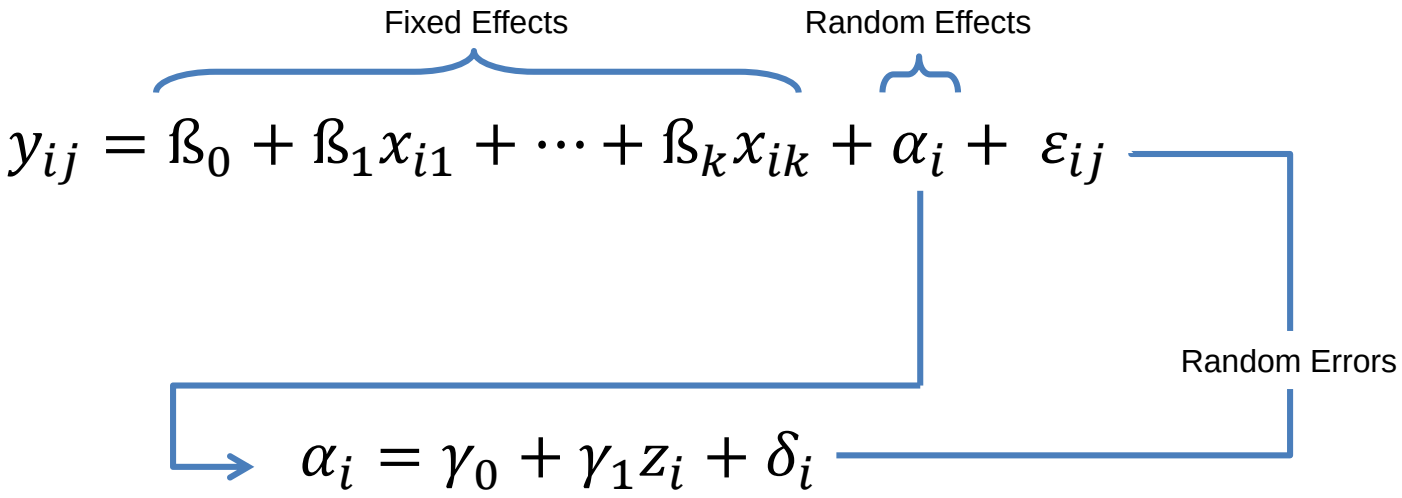
Published 2011:

"To avoid concerns over data-driven selection of methods, it is essential to pre-specify the selected methods in the statistical section of the study protocol or analysis plan (…)" (p. 6)

"Therefore, ***the precise option settings*** *must be fully justified and* ***predefined in advance in detail, so that the results could be replicated****, if required,* ***by an external data analyst*** *and so that it can be established that the choice has not been made post hoc."* (p. 10)

European Medicines Agency (EMA). Guideline on Missing Data in Confirmatory Clinical Trials. 2011:1-12

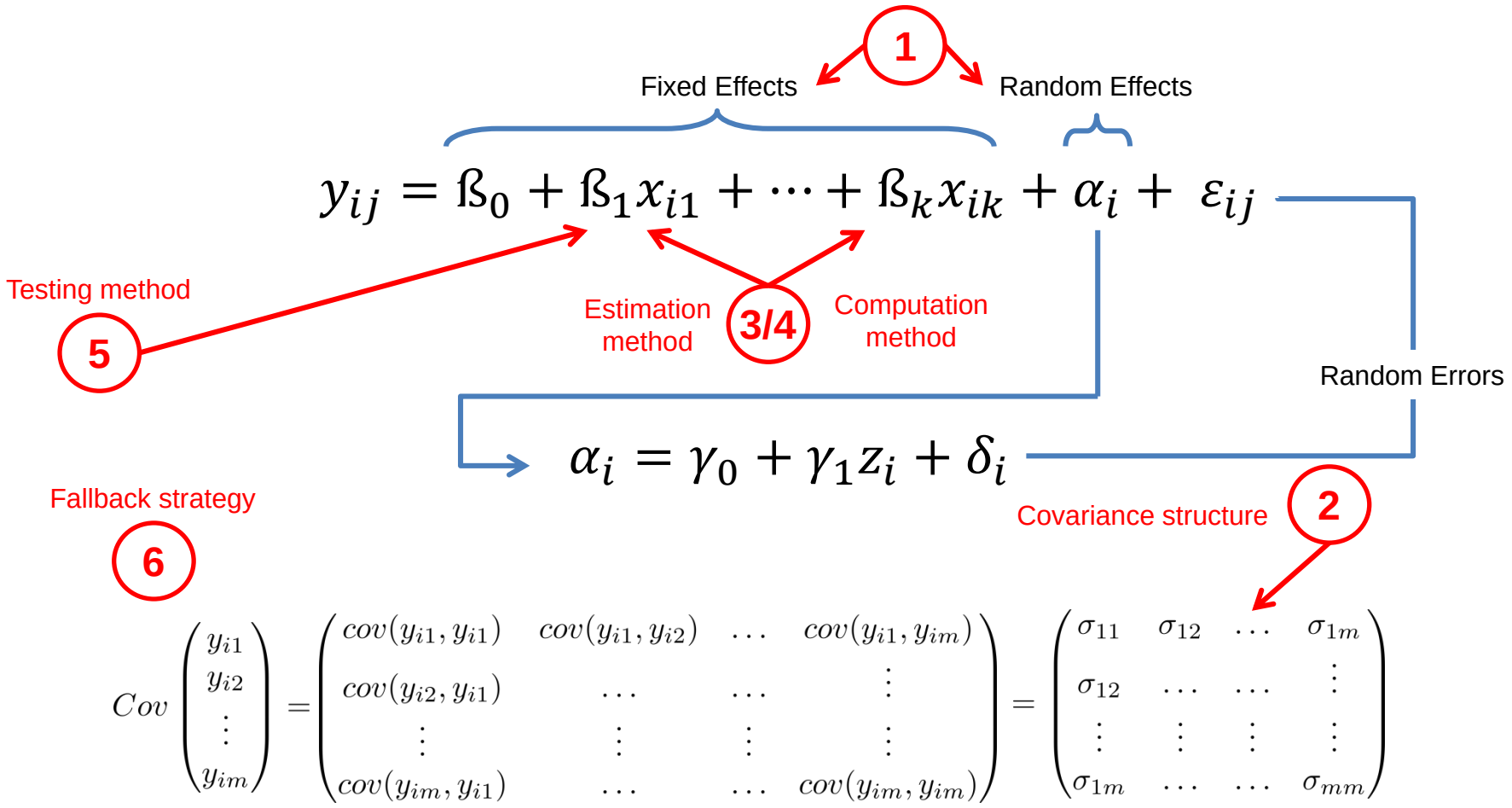
Mixed Model for Longitudinal Data



$$Cov\begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{im} \end{pmatrix} = \begin{pmatrix} cov(y_{i1}, y_{i1}) & cov(y_{i1}, y_{i2}) & \dots & cov(y_{i1}, y_{im}) \\ cov(y_{i2}, y_{i1}) & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ cov(y_{im}, y_{i1}) & \dots & \dots & cov(y_{im}, y_{im}) \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{12} & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{1m} & \dots & \dots & \sigma_{mm} \end{pmatrix}$$

Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. Verlag New York, LLC: Springer; 2009.

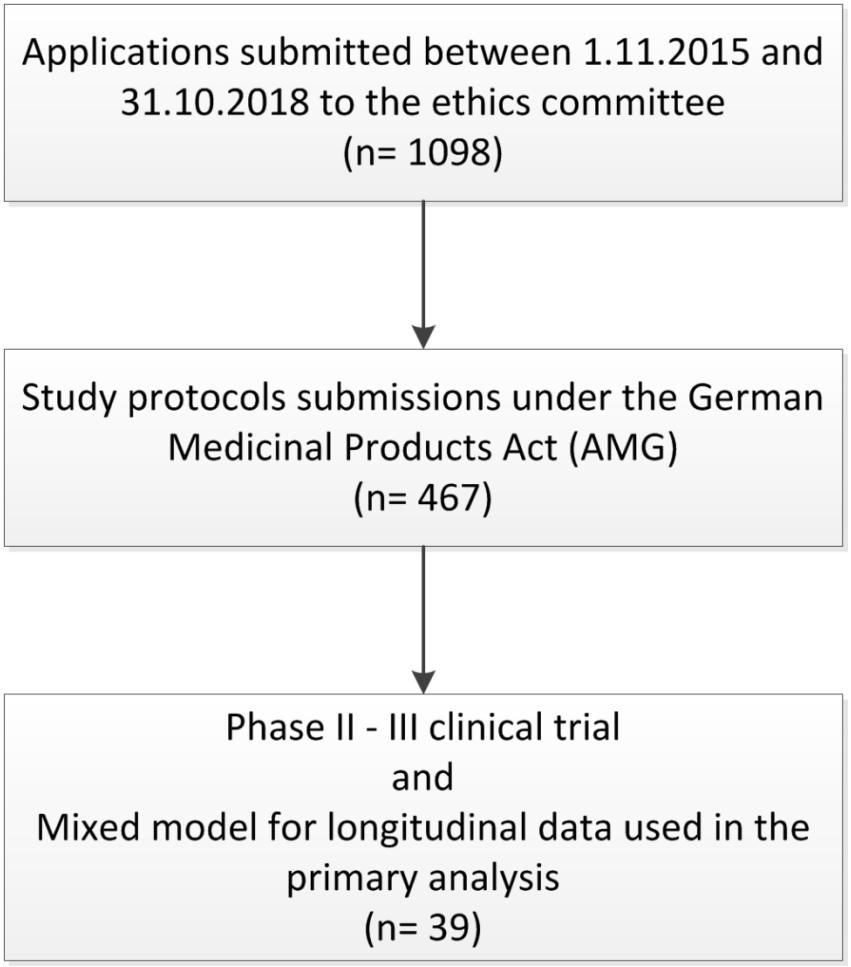
Mixed Model for Longitudinal Data



Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. Verlag New York, LLC: Springer; 2009.

Methods

- Access to study protocols granted by MHH ethics committee
- Study protocols evaluated independently
- Discordances resolved by consensus agreement



Trial characteristics

Table 1 Characteristics of included clinical trial protocols

Evaluation Item	Development phase II (n = 15)	Development phase III (n = 24)	All trials (n=39)
Sponsor			
Pharmaceutical company[a]	12 (80%)	23 (96%)	35 (90%)
Top 21 pharmaceutical company[b]	6 (40%)	11 (46%)	17 (44%)
Planned sample size			
mean (±standard deviation)	141 (±125)	651 (±992)	455 (±815)
median (minimum, maximum)	99 (30, 500)	232 (15, 4126)	180 (15, 4126)
Therapeutic area[c]			
Blood or blood-forming organs	0 (0.0%)	3 (12.5%)	3 (7.7%)
Endocrine, nutritional or metabolic diseases	0 (0.0%)	2 (8.3%)	2 (5.1%)
Mental, behavioural or neurodevelopmental disorders	0 (0.0%)	2 (8.3%)	2 (5.1%)
Nervous system	0 (0.0%)	3 (12.5%)	3 (7.7%)
Visual system	1 (6.7%)	2 (8.3%)	3 (7.7%)
Ear or mastoid process	1 (6.7%)	0 (0.0%)	1 (2.6%)
Circulatory / cardiovascular system	0 (0.0%)	2 (8.3%)	2 (5.1%)
Respiratory system	6 (40.0%)	6 (25.0%)	12 (30.8%)
Digestive System	3 (20.0%)	1 (4.2%)	4 (10.3%)
Skin	1 (6.7%)	1 (4.2%)	2 (5.1%)
Immune System	2 (13.3%)	1 (4.2%)	3 (7.7%)
Genitourinary system	1 (6.7%)	0 (0.0%)	1 (2.6%)
Developmental abnomalies	0 (0.0%)	1 (4.2%)	1 (2.6%)

[a] : Including the top 21 pharmaceutical companies

[b] : By global sales in 2017

[c] : Only ICD-11 [20] superior categories 01-20 are considered since categories 21-26 do not represent therapeutic areas

Results

Table 2 Reporting of primary mixed model analyses by clinical development phase

	Development phase			All Trials
Evaluation item	**II (n = 15)**	**III (n = 24)**	**P Value**	**(n = 39)**
Fixed and random effects				37/39 (94.9%)
Covariance structure				30/39 (76.9%)
Testing method				14/39 (35.9%)
Estimation method				11/39 (28.2%)
Computation method				1/39 (2.6%)
Fallback strategy				7/39 (17.9%)
SAP reference				12/39 (30.8%)
All items specified[+]				0/39 (0.0%)
Main items specified*				12/39 (30.8%)
Main items specified* or SAP reference				21/39 (53.8%)

[+] excluding reference to SAP

* Main items are fixed/random effects, covariance structure and testing method

[#]: p-value derived from Chi^2-test comparing proportions between study phases

[##]: p-value derived from Fisher's exact test comparing proportions between study phases

Results

Table 2 Reporting of primary mixed model analyses by clinical development phase

	Development phase			All Trials
Evaluation item	**II (n = 15)**	**III (n = 24)**	**P Value**	**(n = 39)**
Fixed and random effects				37/39 (94.9%)
Covariance structure				30/39 (76.9%)
Testing method				14/39 (35.9%)
Estimation method				11/39 (28.2%)
Computation method				1/39 (2.6%)
Fallback strategy				7/39 (17.9%)
SAP reference				12/39 (30.8%)
All items specified[+]				0/39 (0.0%)
Main items specified*				12/39 (30.8%)
Main items specified* or SAP reference				21/39 (53.8%)

[+] excluding reference to SAP

* Main items are fixed/random effects, covariance structure and testing method

[#]: p-value derived from Chi^2-test comparing proportions between study phases

[##]: p-value derived from Fisher's exact test comparing proportions between study phases

Results

Table 2 Reporting of primary mixed model analyses by clinical development phase

	Development phase			All Trials
Evaluation item	**II (n = 15)**	**III (n = 24)**	**P Value**	**(n = 39)**
Fixed and random effects	15/15 (100.0%)	22/24 (91.7%)	.51 ##	37/39 (94.9%)
Covariance structure	13/15 (86.7%)	17/24 (70.8%)	.44 ##	30/39 (76.9%)
Testing method	5/15 (33.3%)	9/24 (37.5%)	.79 #	14/39 (35.9%)
Estimation method	4/15 (26.7%)	7/24 (29.2%)	1.00 ##	11/39 (28.2%)
Computation method	0/15 (0.0%)	1/24 (4.2%)	1.00 ##	1/39 (2.6%)
Fallback strategy	1/15 (6.7%)	6/24 (25.0%)	.22 ##	7/39 (17.9%)
SAP reference	3/15 (20.0%)	9/24 (37.5%)	.31 ##	12/39 (30.8%)
All items specified+	0/15 (0.0%)	0/24 (0.0%)	-	0/39 (0.0%)
Main items specified*	5/15 (33.3%)	7/24 (29.2%)	.78 #	12/39 (30.8%)
Main items specified* or SAP reference	7/15 (46.7%)	14/24 (58.3%)	.48 #	21/39 (53.8%)

+ excluding reference to SAP

* Main items are fixed/random effects, covariance structure and testing method

#: p-value derived from Chi^2-test comparing proportions between study phases

##: p-value derived from Fisher's exact test comparing proportions between study phases

Results

Table 3 Reporting of primary mixed model analyses by sponsor type

Evaluation Item	Sponsor type		P Value*	All Trials
	Minor (n = 22)	Major (n = 17)		(n = 39)
Fixed and random effects	20/22 (90.9%)	17/17 (100.0%)	.50 **	37/39 (94.9%)
Covariance structure	15/22 (68.2%)	15/17 (88.2%)	.25 **	30/39 (76.9%)
Testing method	5/22 (22.7%)	9/17 (52.9%)	.05 *	14/39 (35.9%)
Estimation method	3/22 (13.6%)	8/17 (47.1%)	.03 **	11/39 (28.2%)
Computation method	0/22 (0.0%)	1/17 (5.9%)	.44 **	1/39 (2.6%)
Fallback strategy	2/22 (9.1%)	5/17 (29.4%)	.21 *	7/39 (17.9%)
SAP reference	9/22 (40.9%)	3/17 (17.6%)	.17 **	12/39 (30.8%)
All items specified[a]	0/22 (0.0%)	0/17 (0.0%)	-	0/39 (0.0%)
Main items specified[b]	3/22 (13.6%)	9/17 (52.9%)	.01 **	12/39 (30.8%)
Main items specified[b] or SAP reference	11/22 (50.0%)	10/17 (58.8%)	.58 *	21/39 (53.8%)

[a] Excluding reference to SAP

[b] Main items are fixed/random effects, covariance structure and testing method

*: P Value derived from Chi^2-test comparing proportions between sponsor types

**: P Value derived from Fisher's exact test comparing proportions between sponsor types

Results

Table 3 Reporting of primary mixed model analyses by sponsor type

Evaluation Item	Sponsor type		P Value*	All Trials
	Minor (n = 22)	Major (n = 17)		(n = 39)
Fixed and random effects	20/22 (90.9%)	17/17 (100.0%)	.50 **	37/39 (94.9%)
Covariance structure	15/22 (68.2%)	15/17 (88.2%)	.25 **	30/39 (76.9%)
Testing method	5/22 (22.7%)	9/17 (52.9%)	.05 *	14/39 (35.9%)
Estimation method	3/22 (13.6%)	8/17 (47.1%)	.03 **	11/39 (28.2%)
Computation method	0/22 (0.0%)	1/17 (5.9%)	.44 **	1/39 (2.6%)
Fallback strategy	2/22 (9.1%)	5/17 (29.4%)	.21 *	7/39 (17.9%)
SAP reference	9/22 (40.9%)	3/17 (17.6%)	.17 **	12/39 (30.8%)
All items specified[a]	0/22 (0.0%)	0/17 (0.0%)	-	0/39 (0.0%)
Main items specified[b]	3/22 (13.6%)	9/17 (52.9%)	.01 **	12/39 (30.8%)
Main items specified[b] or SAP reference	11/22 (50.0%)	10/17 (58.8%)	.58 *	21/39 (53.8%)

[a] Excluding reference to SAP

[b] Main items are fixed/random effects, covariance structure and testing method

*: P Value derived from Chi^2-test comparing proportions between sponsor types

**: P Value derived from Fisher's exact test comparing proportions between sponsor types

Discussion

- Not a single study protocol specified all items
- Specification of main model items generally poor
  → Control of type-I-error rate at intended level is not guaranteed
- Subgroup analyses:
  - No apparent difference between study phases
  - Major pharmaceutical companies show slightly better compliance
- Future research
  - Confirmatory replication (e.g. multicenter, international study, larger sample size)
  - Extension to different methodologies
  - Investigation of the magnitude of type-I-error inflation (publication in preparation)

Take home messages

- Clinical trial sponsors:
  - Model specifications in study protocols should be improved
  - Reference to a SAP and delay to blind data review should be avoided
  - Better option:
    (i) full specification in study protocol before study start
    (ii) If necessary, documented modifications of pre-specified model at blind data review

- Institutional review boards / ethics committees:
  - Specification of primary analysis models needs to be checked thoroughly
  - Request full specification in line with EMA guidelines

- Drug regulatory agencies:
  - Request and compare model specifications from marketing authorization applications and the final study protocol / SAP

References

Häckl, S, Koch, A, Lasch, F. Empirical evaluation of the implementation of the EMA guideline on missing data in confirmatory clinical trials: Specification of mixed models for longitudinal data in study protocols. *Pharmaceutical Statistics*. 2019; 18: 636– 644.

Fletcher C, Tsuchiya S, Mehrotra DV. Current practices in choosing estimands and sensitivity analyses in clinical trials: results of the ICH E9 survey. Ther Innov Regul Sci. 2017;51(1):69-76.

Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. Verlag New York, LLC: Springer; 2009.

European Medicines Agency (EMA). Guideline on Missing Data in Confirmatory Clinical Trials. 2011:1-12

Thank you for your attention!

Model Items

Item	Examples
Fixed and Random Effects	• Random intercept, random slopes
Covariance Structure	• Unstructured, Compound Symmetry, First-order Autoregressive, Toeplitz, Variance Components • homogeneous or heterogeneous variance
Estimation Method	• ML, REML, minimum variance quadratic unbiased estimation, empirical sandwich estimation
Computation Method	• Expectation-maximization-algorithm, Newton-Raphson algorithm, Fisher scoring algorithm
Testing Method	• Test: type III F-test, likelihood ratio test • Degrees of Freedom: Kenward-Rogers estimation, Satterthwaite approximation
Fallback Strategy	• Modification of covariance structure • Modification of random effects

Results

Table 4: Reporting of primary mixed model analyses by sponsor type (Phase II)

Evaluation item	Sponsor type		p-value[##]
	Minor (n=9)	Major (n=6)	
Fixed and random effects	9/9 (100.0%)	6/6 (100.0%)	-
Covariance structure	7/9 (77.8%)	6/6 (100.0%)	0.49
Testing method	1/9 (11.1%)	3/6 (50.0%)	0.09
Estimation method	1/9 (11.1%)	3/6 (50.0%)	0.24
Computation method	0/9 (0.0%)	0/6 (0.0%)	-
Fallback strategy	0/9 (0.0%)	1/6 (16.7%)	0.40
SAP reference	3/9 (33.3%)	0/6 (0.0%)	0.23
All items specified[+]	0/9 (0.0%)	0/6 (0.0%)	-
Main items specified*	1/9 (11.1%)	4/6 (66.7%)	0.09
Main items specified* or SAP reference	3/9 (33.3%)	4/6 (66.7%)	0.31

[+] excluding reference to SAP
* Main items are fixed/random effects, covariance structure and testing method
[##]: p-value derived from Fisher's exact test comparing proportions between study phases

Results

Table 5: Reporting of primary mixed model analyses by sponsor type (phase III)

Evaluation item	Sponsor type		p-value
	Minor (n=13)	Major (n=11)	
Fixed and random effects	11/13 (84.6%)	11/11 (100.0%)	0.48 ##
Covariance structure	8/13 (61.5%)	9/11 (81.8%)	0.39 ##
Testing method	4/13 (30.8%)	5/11 (45.5%)	0.68 ##
Estimation method	2/13 (15.4%)	5/11 (45.5%)	0.18 ##
Computation method	0/13 (0.0%)	1/11 (9.1%)	0.46 ##
Fallback strategy	2/13 (15.4%)	4/11 (36.4%)	0.36 ##
SAP reference	6/13 (46.2%)	3/11 (27.3%)	0.42 ##
All items specified+	0/13 (0.0%)	0/11 (0.0%)	-
Main items specified*	2/13 (15.4%)	5/11 (45.5%)	0.18 ##
Main items specified* or SAP reference	8/13 (61.5%)	6/11 (54.5%)	0.73 #

+ excluding reference to SAP
* Main items are fixed/random effects, covariance structure and testing method
#: p-value derived from Chi^2-test comparing proportions between study phases
##: p-value derived from Fisher's exact test comparing proportions between study phases