#### 【企画セッション】欠測のあるデータにおける主解析の検討

### (2)多重補完(Multiple Imputation: MI)に基づく 連続量経時データの解析

大江基貴<sup>1)2)</sup>, 鵜飼裕之<sup>1)3)</sup>, 大浦智紀<sup>1)4)</sup>, 土居正明<sup>1)5)</sup>, 藤原正和<sup>1)6)</sup>, 横山雄一<sup>1)7)</sup> 1)日本製薬工業協会 医薬品評価委員会 データサイエンス部会 タスクフォース4 欠測のあるデー タの解析検討チーム

2)株式会社大塚製薬工場, 3)日本ベーリンガーインゲルハイム株式会社, 4)日本イーライリリー株式会社, 5)東レ株式会社, 6)塩野義製薬株式会社, 7)持田製薬株式会社

[Planning session] Study to the candidates for the primary analysis for the longitudinal data with missing observations.

# Continuous longitudinal data analysis by Multiple Imputation

Motoki Oe<sup>1)2)</sup>, Hiroyuki Ugai<sup>1)3)</sup>, Tomonori Oura<sup>1)4)</sup>, Masaaki Doi<sup>1)5)</sup>, Masakazu Fujiwara<sup>1)6)</sup>, Yuichi Yokoyama<sup>1)7)</sup>

1)The team for statistical analysis of data analysis with missing data, task force 4, data science expert committee, drug evaluation committee, Japan Pharmaceutical Manufacturers Association, 2)Otsuka Pharceutical Factory, Inc., 3)Nippon Boehringer Ingelheim, Co., Ltd., 4)Eli Lilly Japan K.K., 5)Toray Industries, Inc., 6)Shionogi & Co., Ltd., 7)Mochida Pharmaceutical Co., Ltd.

### 要旨:

本発表では、欠測のあるデータの解析方法の一つである多重補完 (MI)について、推測の概要を示すとともに、シミュレーション・データの解析を行い、補完回数と検出力の関係について示す.

キーワード: 多重補完, MI, 補完回数, パターン混合モデル, PMM

### 発表構成

- 多重補完(Multiple Imputation: MI)
  - 基本的な理論
- ・シミュレーション
  - うつ病データ
  - 慢性疼痛データ
- ・おわりに

### 多重補完(Multiple Imputation: MI)

## 多重補完(Multiple Imputation: MI)

- Rubin(1978,1987)によって提案された方法
- 欠測メカニズムがMARであれば、MARに基づくMIの解析(主解析)は妥当
- 複数回の補完を行うことで、欠測値の補完に対して不確実性を考慮
- MIIにおける解析モデルと補完モデルが尤度ベースのモデルと同じならば、MIの結果は尤度ベースの結果と類似.(Mallinckrodt, 2013)

補完モデル:欠測値を補完するための統計モデル

解析モデル:多重補完された完全データを用いて解析するための

統計モデル

- MIが有用な状況(Dmitrienko et al, 2005):
  - 共変量の欠測を補完
  - PMM(MNAR)の枠組みでも適用可能
  - 非単調な欠測を補完

### SASによるMIの実行

- 単調又は非単調な欠測パターンを想定
- Proc MIにおける補完方法 (Yuan, 2011)

Pattern of missingness	Type of imputed variable	Available methods
Monotone	Continuous	Monotone regression
		Monotone predicted mean matching
		Monotone propensity score
Monotone	Classification (ordinal)	Monotone logistic regression
Monotone	Classification (nominal)	Monotone discriminant function
Arbitrary	Continuous	MCMC full-data imputation MCMC monotone-data imputation

SAS/STAT 9.2. SAS/STAT 9.22. SAS/STAT 9.3

### MIの補完モデルにおける単調回帰の概略

• ある時点jにおける欠測データ $\underline{Y_j}$ を補完するため、 観測データ $Y_1,\ldots,Y_{j-1}$ を用いて事後予測分布を構築

$$Y_j \sim f(Y_j|Y_1,\ldots,Y_{j-1})$$

事後予測分布を構築する為の回帰モデル

$$Y_j = \beta_0 + \beta_1 Y_1 + \dots + \beta_{j-1} Y_{j-1}$$

- $oldsymbol{eta}$  回帰パラメータの推定値  $\hat{oldsymbol{eta}}=\left(\hat{eta}_0,\hat{eta}_1,\ldots,\hat{eta}_{j-1}
  ight)'$  観測された  $Y_j$
- 共分散行列  $\hat{\sigma}_j^2 \mathbf{V}_j$

ただし、
$$\mathbf{V}_j = \left(\mathbf{X}'\mathbf{X}
ight)^{-1}$$

 $(\mathbf{X}'\mathbf{X})$  は、 $Y_1,\ldots,Y_{j-1}$  で構成されるデザイン行列

観測された  $Y_j$  および  $Y_1, \ldots, Y_{j-1}$  により推定される

### MIIにおける単調回帰の概略

1. パラメータの事後分布から下記をサンプリング

$$\boldsymbol{\beta}_* = (\beta_{*0}, \beta_{*1}, \dots, \beta_{*j-1})', \ \sigma_{*j}^2$$

分散 :  $\sigma_{*j}^2 = \hat{\sigma}_j^2 (n_j - j - 2)/G$ ,  $G \sim \chi_{n_j - j - 2}^2$ 

 $n_j$  は観測されている  $Y_i$  のデータ数

回帰係数 :  $\boldsymbol{\beta}_* = \hat{\boldsymbol{\beta}} + \sigma_{*j} \mathbf{V}_{hj}' \mathbf{Z}$ 

 $\mathbf{V}_{hj}'$ はコレスキー分解  $\mathbf{V}_j = \mathbf{V}_{hj}' \mathbf{V}_{hj}$  により得られる上三角行列

 ${f Z}$  はj 個の互いに独立な標準正規確率変数のベクトル

2. 欠測データ  $Y_j$ を以下の式で生成される値で補完

$$\beta_{*0} + \beta_{*1}Y_1 + \cdots + \beta_{*j-1}Y_{j-1} + z_i\sigma_{*j}$$

 $z_i$ :標準正規分布からサンプリング

3. 1, 2 をM回(補完回数)繰り返す. i:被験者を表す識別変数

### MIの概略(Rubin, 1987)

- 関心のある k+1 次元のパラメータ・ベクトル  $\theta$
- 補完ごとに heta を推定 o m 回目の補完の推定値  $\hat{m{ heta}}^{(m)}$
- M個(補完回数)の完全データセットから推定されるパラメータ

$$\hat{\boldsymbol{\theta}}^* = \frac{1}{M} \sum_{i=1}^{M} \hat{\boldsymbol{\theta}}^{(m)}, \qquad \left(\hat{\boldsymbol{\theta}}^{(m)} - \boldsymbol{\theta}\right) \stackrel{\cdot}{\approx} \operatorname{N}\left(\mathbf{0}, \mathbf{U}^{(m)}\right)$$

個々のパラメータには, 漸近正規性を仮定

• このとき、 $\hat{\theta}^*$ は以下の分布に従う.

$$(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}) \stackrel{\cdot}{\approx} N(\mathbf{0}, \mathbf{V})$$

$$\mathbf{V} = \widehat{\mathbf{W}} + \left(\frac{M+1}{M}\right) \widehat{\mathbf{B}}$$



### 平均補完内分散

$$\widehat{\mathbf{B}} = \frac{1}{M-1} \sum_{m=1}^{M} \left( \widehat{\boldsymbol{\theta}}^{(m)} - \widehat{\boldsymbol{\theta}}^* \right) \left( \widehat{\boldsymbol{\theta}}^{(m)} - \widehat{\boldsymbol{\theta}}^* \right)'$$

#### 補完間分散

### MIの概略(Rubin, 1987)

•  $\hat{ heta}^*$  の漸近正規性から、以下の検定統計量を構成できる.

$$\left(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}\right)' \mathbf{V}^{-1} \left(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}\right) \stackrel{\cdot}{\approx} \chi_k^2$$

Li et al. (1991)によるF分布に基づく推測を利用

- 
$$H_0: \boldsymbol{\theta} = \boldsymbol{\theta}_0, \ H_1: \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$$
 に対するF検定

$$F = \frac{\left(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}_0\right)' \mathbf{V}^{-1} \left(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}_0\right)}{k(R+1)} \approx F_w^k$$

$$R=rac{1}{k}\left(1+rac{1}{M}
ight)\mathrm{tr}\left(\mathbf{B}\mathbf{W}^{-1}
ight)$$
  $au=k(M-1)$  p値は, $\Pr\left(F_w^k>F
ight)$ 

 $w = 4 + (\tau - 4) \left( 1 + \frac{1 - 2\tau^{-1}}{R} \right)^{2}$ 

・ 漸近理論に基づく推測は 標本サイズと補完回数 M に依存する.

### シミュレーション

### 補完回数に関する諸種の議論

- Rubin(1978, 1987, 1996)
  - 推定効率の観点で、3~5回でも十分としている.
- Graham et al. (2007)
  - 10~30%の欠測で,検出力の低下を1%未満に したい場合は,20回以上の補完を推奨.

$$\frac{V(\theta_{\infty})}{V(\theta_{M})} = \left(1 + \frac{\gamma}{M}\right)^{-1}$$

M:補完回数

 $\gamma$ :欠測割合(0~1)

- Siddiqui(2011)
  - 40%以下の欠測割合で、補完回数を10回としたシミュレーションを実施
- Royston & White (2011)
  - MCMCによる補完を行う場合は、100~1000回の補完が必要.
- White et al. (2011)
  - 100×欠測割合の補完回数を推奨. ただしデータ解析では, 100回/ 500回を利用している.

### これまでの検討

• うつ病データ(土居ら, 2014)

- 補完モデル : 投与群ごとの単調回帰モデル

— 補完回数 : 5回

- 解析方法 : 共分散分析

結果 : MAR, MNARともに、MMRMとほとんど同じ結果

• 慢性疼痛データ(横山ら, 2015)

- 補完モデル : 投与群ごとの単調回帰モデル

- 補完回数 : 100回

- 解析方法 : 共分散分析

- 結果 : MCAR+MAR, MCAR+MNARともに, MMRMより4~5%

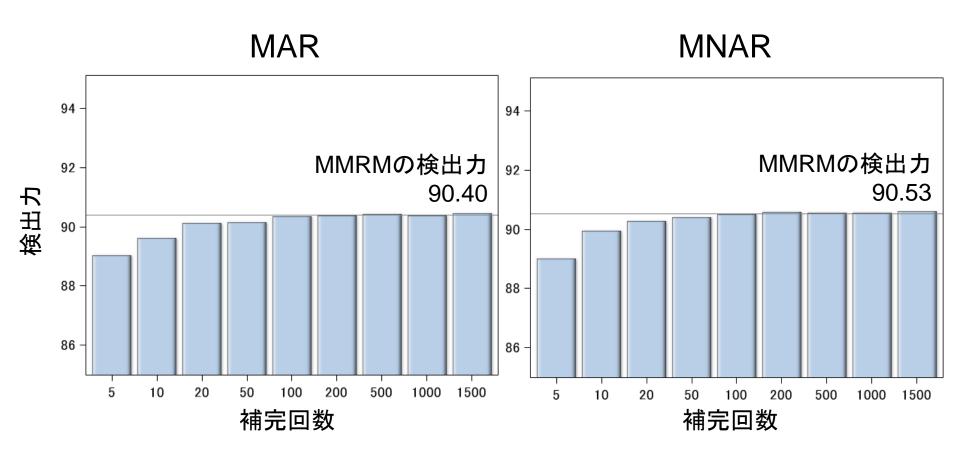
検出力が劣る結果

### シミュレーション

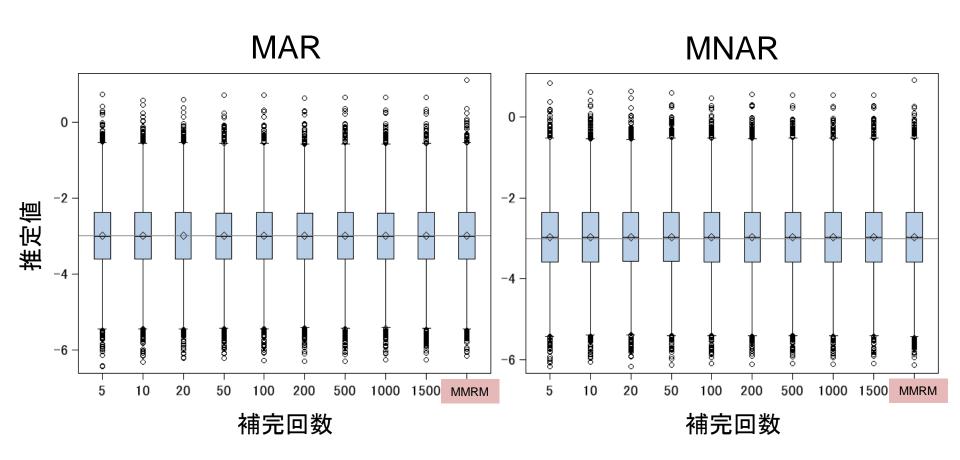
#### 補完回数を増やした場合のMIの推測性能の評価を行う

- 解析方法
  - 補完モデル
    - 群ごとの単調回帰モデル(REG)
    - 説明変数は、ベースライン値と各時点のベースライン値からの変化量
  - 解析モデル
    - 時点4を応答, 投与群を説明変数, ベースライン値を共変量としたANCOVA
    - 他の共変量はなし
  - 補完回数
    - 5, 10, 20, 50, 100, 200, 500, 1000, 1500の9水準
- 評価項目
  - 時点4の群間差に対する検出力,推定値,標準誤差,推定効率

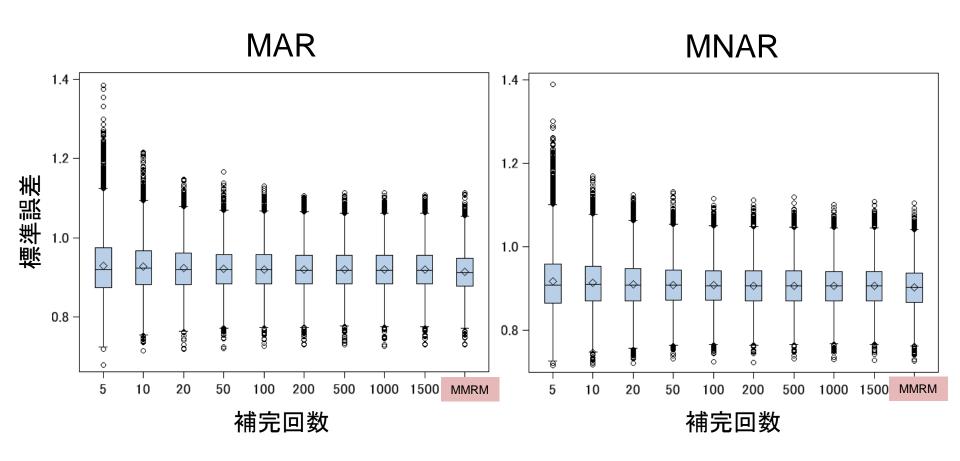
### シミュレーション結果(うつ病データ,検出力)



### シミュレーション結果(うつ病データ,推定値)



### シミュレーション結果(うつ病データ,標準誤差)



※各シミュレーションで計算されるSEのボックスプロット

### シミュレーション結果(うつ病データ,標準誤差2)

#### MAR

	MI:補完回数									MMRM
	5	10	20	50	100	200	500	1000	1500	
SEの平均値	0.930	0.927	0.923	0.922	0.921	0.920	0.920	0.920	0.920	0.914
SEの中央値	0.921	0.923	0.921	0.920	0.920	0.919	0.919	0.919	0.918	0.913
推定値のSD	0.917	0.912	0.906	0.905	0.904	0.904	0.904	0.904	0.904	0.908

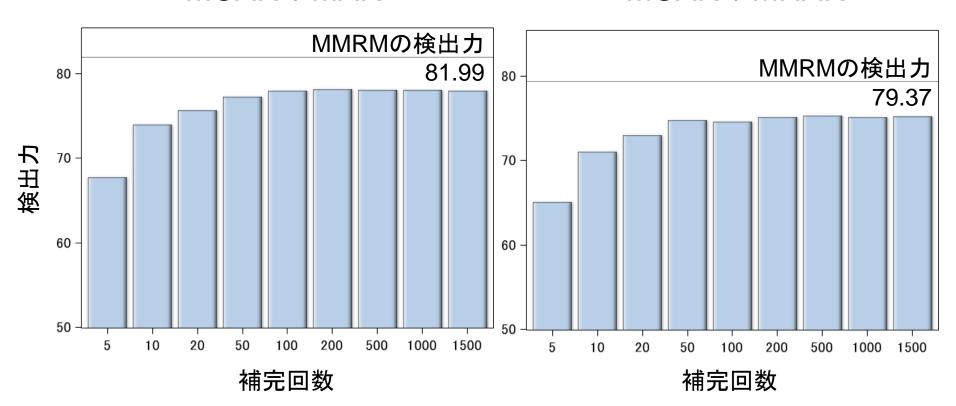
#### MNAR

	MI : 補完回数									MMRM
	5	10	20	50	100	200	500	1000	1500	
SEの平均値	0.917	0.914	0.909	0.908	0.908	0.908	0.907	0.907	0.907	0.903
SEの中央値	0.909	0.911	0.909	0.908	0.908	0.907	0.907	0.907	0.907	0.902
推定値のSD	0.904	0.898	0.894	0.893	0.893	0.892	0.892	0.892	0.892	0.897

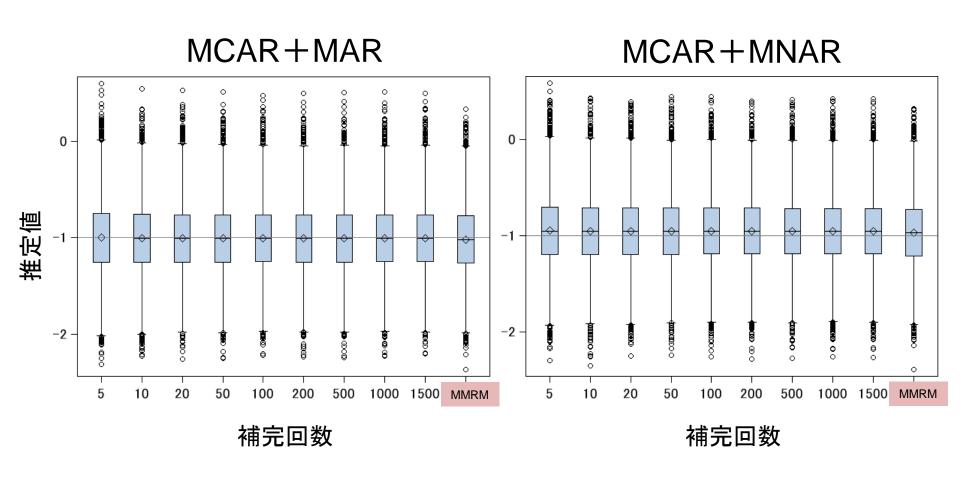
### シミュレーション結果(慢性疼痛データ,検出力)



#### MCAR+MNAR

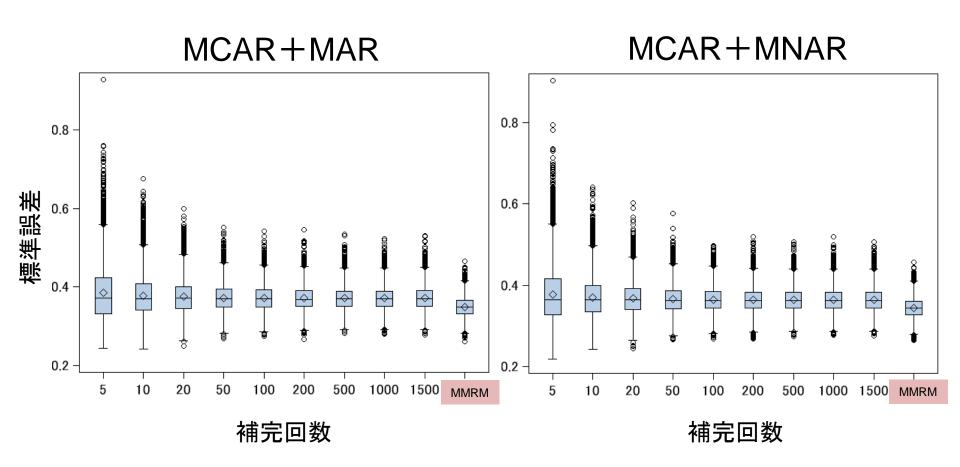


### シミュレーション結果(慢性疼痛データ,推定値)



### SASユーザー総会

### シミュレーション結果(慢性疼痛データ,標準誤差)



※各シミュレーションで計算されるSEのボックスプロット

### シミュレーション結果(慢性疼痛データ,標準誤差2)

#### MCAR + MAR

	MI:補完回数									MMRM
	5	10	20	50	100	200	500	1000	1500	
SEの平均値	0.385	0.377	0.375	0.373	0.372	0.371	0.371	0.371	0.371	0.350
SEの中央値	0.373	0.371	0.372	0.371	0.370	0.369	0.370	0.370	0.370	0.349
推定値のSD	0.375	0.365	0.362	0.360	0.359	0.358	0.358	0.358	0.358	0.353

#### MCAR + MNAR

	MI:補完回数									MMRM
	5	10	20	50	100	200	500	1000	1500	
SEの平均値	0.377	0.370	0.368	0.365	0.365	0.364	0.364	0.364	0.364	0.345
SEの中央値	0.365	0.364	0.365	0.363	0.363	0.363	0.363	0.362	0.362	0.344
推定値のSD	0.367	0.362	0.358	0.355	0.354	0.353	0.353	0.353	0.353	0.349

#### CALK, COCOAS SASI-1-1-18

### 推定効率の検討

• Rubin(1987)による漸近推定効率

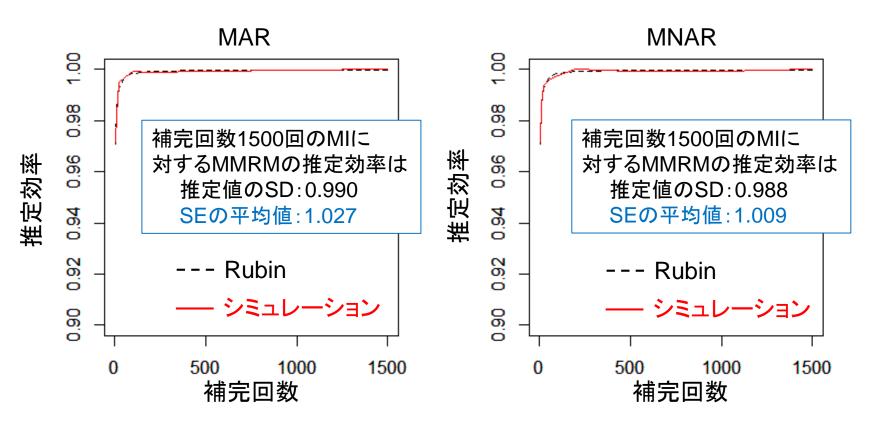
$$rac{\mathrm{V}( heta_\infty)}{\mathrm{V}( heta_M)} = \left(1 + rac{\gamma}{M}
ight)^{-1} \qquad M:$$
補完回数  $\gamma:$ 欠測割合(0~1)

- 補完回数が増えると、推定値のバラつき(分散)がどうなるかを表す。
- 推定効率の観点では、補完回数はM=10程度で十分?

補完回数と欠測割合の組み合わせにおける漸近推定効率×100(Siddiqui, 2011)

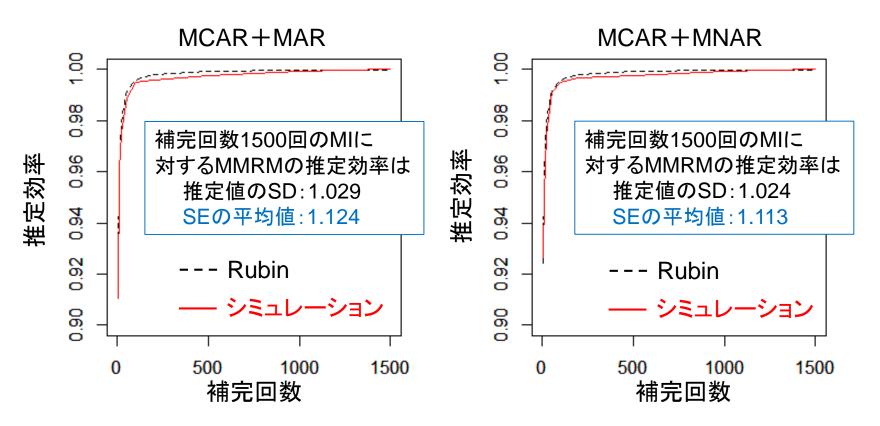
	欠測データの割合 $\gamma$											
M	0.1	0.3	0.5	0.7	0.9							
1	91	77	67	59	53							
3	97	91	86	81	77							
5	99	94	91	88	85							
10	99	97	95	93	92							
∞	100	100	100	100	100							

### シミュレーション1結果(うつ病データ,推定効率)



- ・ 補完回数1500回を∞と見做して、推定値のSDに基づき推定効率を計算した。
- Rubinの推定効率の算出に用いる欠測割合は、時点4での群平均 0.15 を利用した。

### シミュレーション1結果(慢性疼痛データ,推定効率)



- 補完回数1500回を∞と見做して、推定値のSDに基づき推定効率を計算した。
- Rubinの推定効率の算出に用いる欠測割合は、時点4での群平均 0.41 を利用した。

### 結果のまとめ

- 補完回数と推定効率
  - シミュレーションの計算値は、Rubin(1987)の計算式にほぼ合致した.
  - 推定効率の観点では、補完回数は10回程度で十分かもしれない。
- 補完回数と検出力
  - 少ない補完回数では、検出力が低下が示唆された.
    - うつ病データ(MAR)では、90.34(100回)に対して89.02(5回)
    - 慢性疼痛データ(MCAR+MAR)では、77.94(100回)に対して67.74(5回)
  - − いずれのシミュレーション・データでも、補完回数の増加に伴って検出力が増加した(SEの平均値が小さくなった)。
  - ただし、補完回数が100回を超えると、検出力が頭打ちするように見えた.
  - SEの分布も、補完回数100回のあたりからほぼ同様と考えられた.

### おわりに

- 欠測メカニズムがMARならば、MIは妥当な推測を与えることが知られている
  - 検討した範囲では、慢性疼痛データ(MCAR+MNAR)に対して若干のバイアスが認められたものの、MNARでも大きな問題はなさそうであった。
- 補完回数が十分であれば、MMRMと同様の結果を与える。
  - ・・・はずだが、今回の検討の結果、データによってMMRMより保守的となる場面があることが示唆された。
    - 補完回数を増やせば、ある程度は検出力が増加した.
    - しかし、慢性疼痛データでは、MMRMには追いつかなかった。
- SEを過大推定している可能性がある。
  - MMRMの結果に比べて、αエラーも名義水準から若干離れて小さい値
  - MIの保守的な結果を示した論文もある(Kim, 2004; Kim et al., 2006; Siddiqui, 2011).

### 参考文献

- Dmitrienko, A., Molenberghs, G., Chuang-Stein, C. & Offen, W. W. (2005). *Analysis of clinical trials using SAS: A practical guide*. SAS Institute.
- Graham, J. W., Olchowski, A. E. & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, **8**, 206-213.
- Kim, J. K. (2004). Finite sample properties of multiple imputation estimators. *The Annals of Statistics*. **32**(2), 766-783.
- Kim, J. K., Brick, J. M., Fuller, W. A. & Kalton, G. (2006). On the bias of the multiple-imputation variance estimator in survey sampling. J. R. Statist. Soc. B. 68(3), 509-521.
- Li, K. H., Raghunathan, T. E. & Rubin, D. B. (1991). Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution. *Journal of the American Statistical Association*. **86**(416), 1065-1073.
- Little, R. J. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*. **88**(421), 125-134.
- Little, R. J. (1994). A class of pattern-mixture models for normal incomplete data.
   Biometrika. 81(3), 471-483.
- Little, R. J. (1995). Modeling the drop-out mechanism in repeated-measures studies. Journal of the American Statistical Association. **90**(431), 1112-1121.

### 参考文献

- Mallinckrodt, C. H. (2013). Preventing and treating missing data in longitudinal clinical trials: a practical guide. Cambridge University Press.
- National Research Council (2010). *The Prevention and Treatment of Missing Data in Clinical Trials*. Washington, DC: The National Academies Press.
- Rubin, D. B. (1978). Multiple Imputation in sample surveys- A phenomenological Bayesian approach to nonresponse. Imputation and Editing of Faulty or Missing Survey Data. Washington, DC: *U.S. Department of Commerce*.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*. **91**(434), 473-489.
- Ratitch, B., O'Kelly, M. & Tosiello, R. (2013). Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical statistics*. **12**(6), 337-347.
- Royston, P. & White, I. R. (2011). Multiple imputation by chained equations (MICE): Implementation in Stata. *Journal of Statistical Software*. **45**(4), 1-19.
- Siddiqui, O. (2011). MMRM versus MI in dealing with missing data a comparison based on 25 NDA data sets. *Journal of Biopharmaceutical Statistics*. **21**, 423-436.

### 参考文献

- Thijs, H., Molenberghs, G., Michiels, B., Verbeke, G. & Curran, D. (2002). Strategies to fit pattern mixture models. *Biostatistics*. **3**(2), 245-265.
- Yuan, Y. (2011). Multiple imputation using SAS software. *Journal of Statistical Software*. **45**(6), 1-25.
- White, I. R., Royston, P. & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*. **30**(4), 377-399.
- 土居正明, 大浦智紀, 大江基貴, 駒嵜弘, 髙橋文博, 縄田成毅, 藤原正和, 横溝孝明, 横山雄一. (2014). 欠測のあるデータに対する総合的な感度分析と主解析の選択. SASユーザー総会論文集.
- ・ 横山雄一,横溝孝明,大浦智紀,大江基貴. (2015).【日本製薬工業協会シンポジウム】臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について 一統計手法・estimandと架空の事例に対する流れの整理 (7)架空の事例2(主解析の選択・例数設計・データの発生方法)

### ご清聴ありがとうございました

