

# 【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  
—統計手法・estimandと架空の事例に対する流れの整理—

## (8) 架空の事例の検討3 (主解析・感度分析)

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  
欠測のあるデータに対する解析方法論・SASプログラム検討チーム

マルホ株式会社 駒寄弘  
杏林製薬株式会社 縄田成毅  
大正製薬株式会社 横溝孝明

1. 事例の試験デザイン、主解析、感度分析
2. 感度分析
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈

1. 事例の試験デザイン、主解析、感度分析
2. 感度分析：
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈

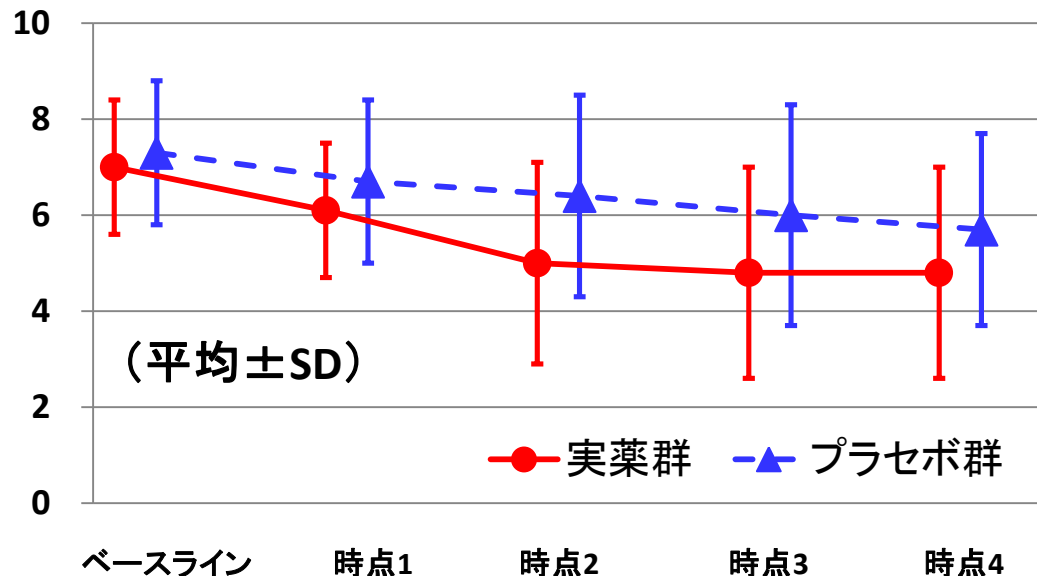
# 想定している試験デザイン

## 架空の事例の検討1より

|            |  |
|------------|--|
| 対象疾患       | 慢性疼痛                                       |
| デザイン       | プラセボ対照, ランダム化, 二重盲検, 並行群間比較, 優越性           |
| 開発段階       | 第三相  |
| 主要評価項目     | 投与16週時点における疼痛スコア(0~10)の週平均値のベースラインからの変化量   |
| 中止割合       | 10~50%                                     |
| 来院間隔       | 0,4,8,12,16週後(5時点)                         |
| 試験治療中止後の対応 | 試験治療中止後も可能な限りデータを収集<br>新規治療(標準治療を含む)の開始も可能 |

# 解析対象データ

## ◎慢性疼痛の第III相試験を想定したシミュレーションデータ



● 主要評価項目：疼痛スコア  
 → **スコア低下：改善**  
 (解析には変化量使用)

● 実薬群 vs プラセボ群  
 ・1群100例(ベースライン時)

◎単調な欠測のみ

|       | ベースライン |           | 時点1 |           | 時点2 |           | 時点3 |           | 時点4 |           |
|-------|--------|-----------|-----|-----------|-----|-----------|-----|-----------|-----|-----------|
|       | 例数     | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   |
| 実薬群   | 100    | 7.0 (1.4) | 80  | 6.1 (1.4) | 72  | 5.0 (2.1) | 69  | 4.8 (2.2) | 67  | 4.8 (2.2) |
| プラセボ群 | 100    | 7.3 (1.5) | 80  | 6.7 (1.7) | 67  | 6.4 (2.1) | 58  | 6.0 (2.3) | 52  | 5.7 (2.0) |

# 主解析の解析仕様

## 架空の事例の検討1より

|          |   |
|----------|---|
| Estimand | 計画している最終時点での群間差(全症例完了を想定)<br>Estimand 3(efficacy) |
| 欠測メカニズム  | MAR   |
| モデル      | MMRM  |
| 共変量      | ベースラインの疼痛スコア, 治験薬, 時点, 治験薬×時点                     |
| 分散共分散構造  | Unstructured                                      |
| 有意水準     | 両側5%  |

# 主解析の解析仕様と検討事項

---

## 主解析・感度分析の全体像

### ◎NRC (2010)の感度分析

- (1) 完全データの分布の検討
  - (2) 外れ値・外れた症例の探索
  - (3) 欠測メカニズムに対する感度
- (3)が最も重要. (3)に絞って解説.

### ◎Mallinckrodt (2013)の考え方

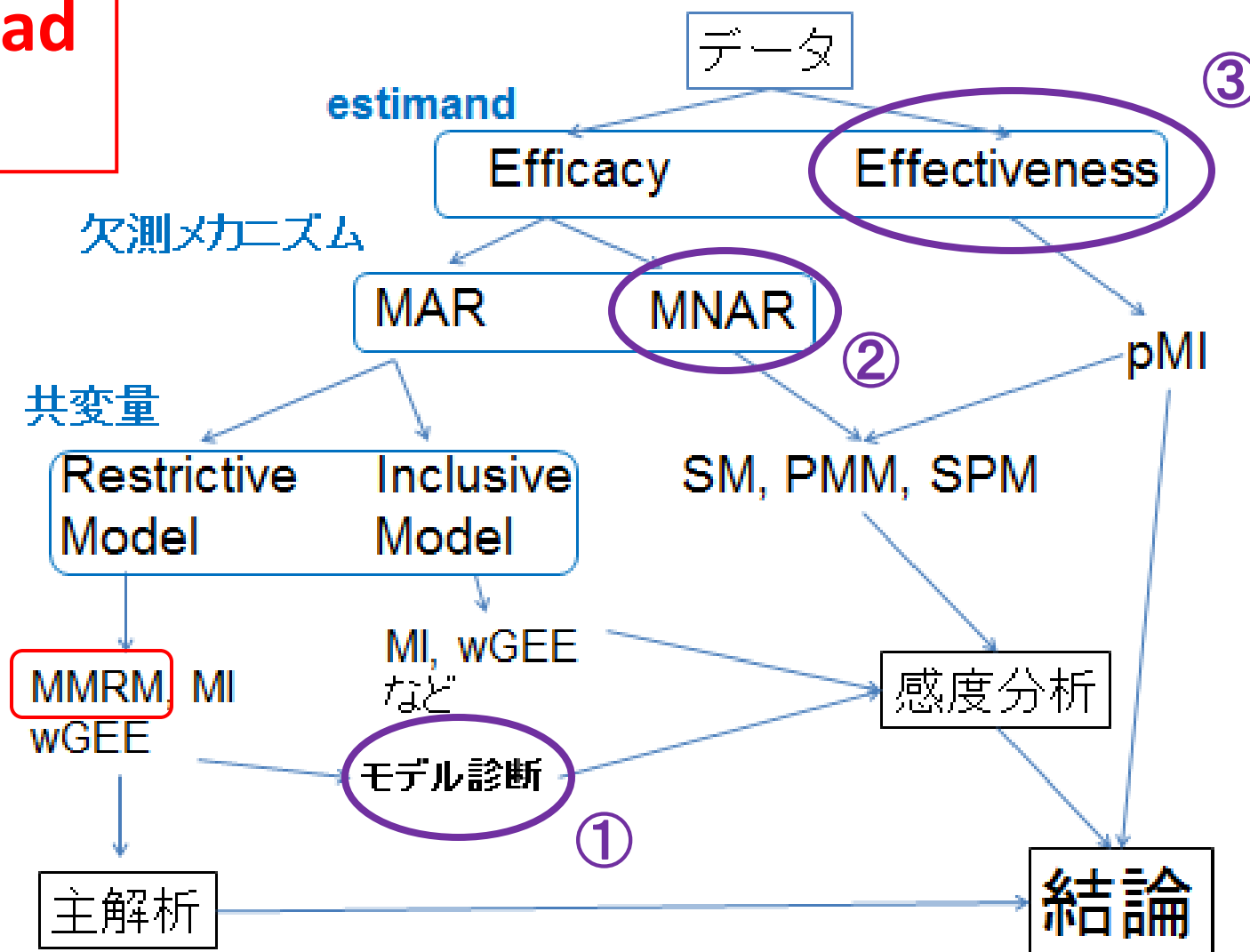
“Analytic Road Map” を提示し, NRCの(1)～(3)に対して, より広い枠組みで紹介(こちらでも(3)が最も重要)

**本発表ではAnalytic Road Mapに従って, 主解析, 感度分析を実際に行ってみる.**

# 主解析の解析仕様と検討事項

## Analytic Road Map

Mallinckrodt (2013)より  
一部編集





# 主解析・感度分析の一覧

|       | 仮説            | Estimand      | 欠測<br>メカニズム | 解析<br>手法         | その他             |
|-------|---------------|---------------|-------------|------------------|-----------------|
| 主解析   | Efficacy      | 3 (primary)   | MAR         | MMRM             |                 |
| 感度分析1 | Efficacy      |               | MAR         | MMRM             | 残差診断<br>・影響診断   |
| 感度分析2 | Efficacy      |               | MNAR        | SM               | 感度パラメータ<br>-1~1 |
|       |               |               | MNAR        | PMM<br>+<br>MMRM | 感度パラメータ<br>-1~1 |
| 感度分析3 | Effectiveness | 6 (secondary) | MNAR        | pMI<br>+<br>MMRM |                 |

◎3種類の感度分析を実行

1. 事例の試験デザイン、主解析、感度分析
2. 感度分析：
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈

# 完全データの分布に対する感度分析

---

- ① 共変量
- ② 残差診断
- ③ 影響診断
- ④ 分散共分散構造

# 完全データの分布に対する感度分析

## ①共変量

共変量を2種類に分ける

### •デザイン変数

- ✓ 投与群やベースライン時に得られる共変量.
- ✓ 全ての症例に対して観測され, 主解析に用いられる.

### •補助変数

- ✓ ベースライン以降に得られるデータも含める.
- ✓ 欠測したデータに対する推測に利用できる共変量.
- ✓ 主解析の共変量には用いられないが, 欠測確率や欠測データの分布のモデリングに役立つ.

→ ①MNARがMARへ近づく可能性  
②主解析のパラメータ推定量の  
精度向上

# 完全データの分布に対する感度分析

## ①共変量

共変量の感度分析は二つに分類される

- 主解析で用いるデザイン変数の選択

例) 応答変数のベースライン値の使用

- ✓ モデルに含める
- ✓ モデルに含めない
- ✓ ベースライン値 × 時点の交互作用として含める

- 補助変数を含めた解析

- ✓ 主解析(デザイン変数のみ): **Restrictive Model**
- ✓ 感度分析(デザイン変数 + 補助変数): **Inclusive Model**



ベースライン + 投与群 + 時点 + 投与群と時点の交互作用 + 前時点の応答変数

# 完全データの分布に対する感度分析

## ①共変量

### MI

欠測のあるデータセット



MIの実施

主解析 : **Restrictive model**  
感度分析 : **Inclusive model**



補完後のデータセット



主解析の実施

(主解析, 感度分析 : **Restrictive model**)



パラメータ推定

### wGEE

欠測のあるデータセット



欠測確率モデルの構築

主解析 : **Restrictive model**  
感度分析 : **Inclusive model**



wGEEの実施

(主解析, 感度分析 : **Restrictive model**)



パラメータ推定

# 完全データの分布に対する感度分析

## ②残差診断③影響診断



### ②残差診断

残差プロットから「全体の傾向」「外れ値」をみる。

### ③影響診断

パラメータ推定に影響の大きい症例(施設)をみる。

→ CookのD統計量など

**外れ値データ, 影響の大きい症例を除いたときの結果の頑健性を確認する**

(プログラム例)

```
PROC MIXED DATA=DAT;  
  CLASS TREAT TIME ID;  
  MODEL OUTCOME = BASELINE TREAT TIME TREAT*TIME  
                / INFLUENCE(EFFECT=ID ITER=3) RESIDUAL;  
  REPEATED TIME / TYPE = UN SUBJECT = ID;  
  LSMEANS TREAT*TIME / DIFF;  
RUN;
```

# 完全データの分布に対する感度分析

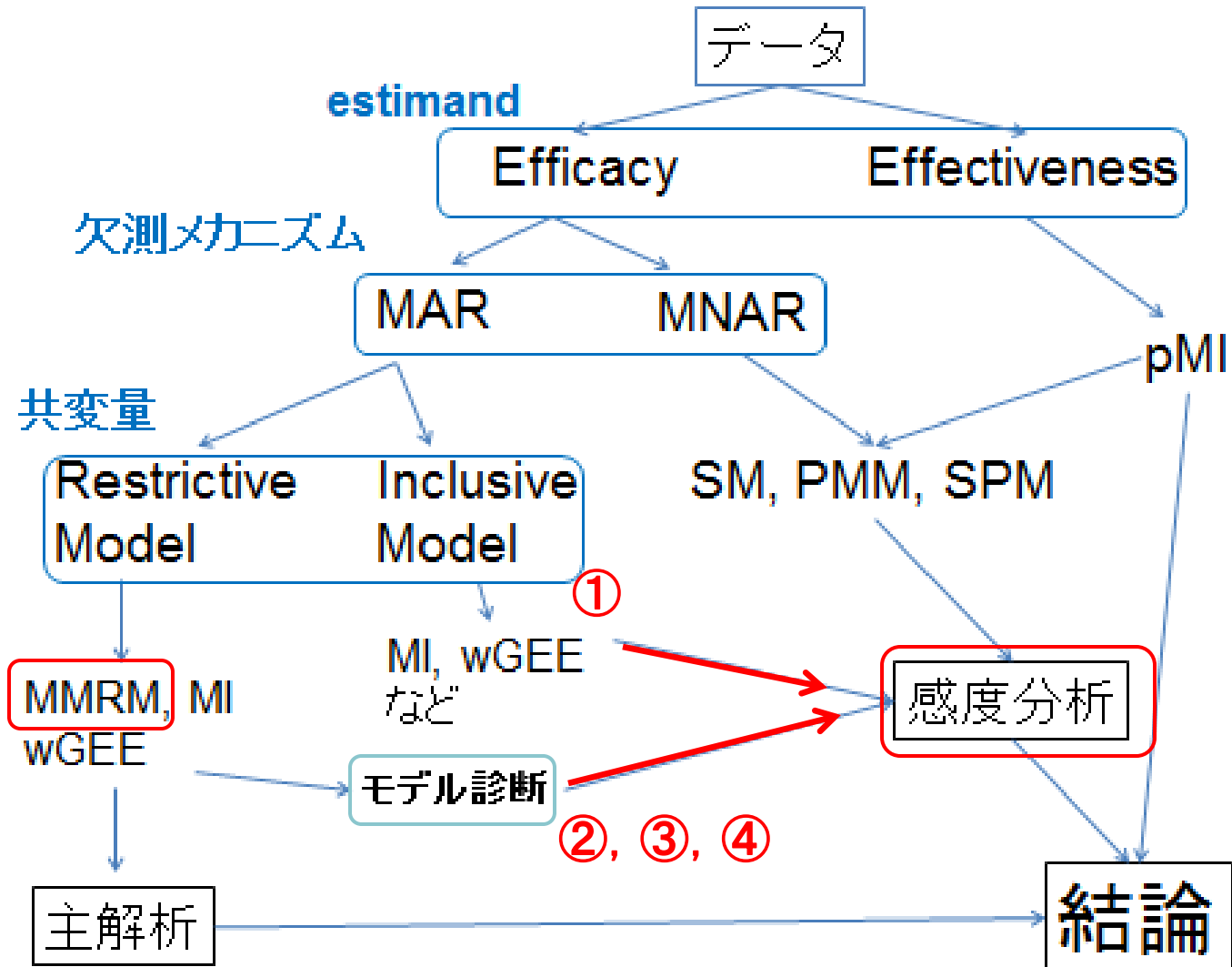
## ④分散共分散構造



- MMRMでは Unstructured (無構造) を指定することが多い。
  - 他の構造では？
  - 色々試して以下の点を確認
    1. 結果の安定性
    2. モデルの当てはまり度合い (AIC等)
    3. 時点間の相関



# 完全データの分布に対する感度分析 Analytic roadmapでおさらい



1. 事例の試験デザイン、主解析、感度分析
2. 感度分析：
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈

# 欠測メカニズムに対する感度分析

- MARかMNARかは検証不可能 (Type (i)) な仮定  
→ 主解析でMARを仮定した場合でも, MNARの可能性は否定できない.
- NRC (2010)とMallinckrodt (2013)で「感度分析の中で最も重要」と指摘.

**MARとMNARそれぞれの仮定の下での結果を比較し, 頑健性を確認することが近年推奨されている. 本発表ではNRC(2010)で紹介されている感度パラメータを用いた感度分析を紹介する.**

# SM及びPMMを用いたMNARを仮定した解析

## ▪ SM:

◎感度パラメータを設定 → 欠測確率のモデルを「色々試す」

$$f(\mathbf{Y}_i, \mathbf{R}_i | \theta, \psi) = f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \theta) \cdot f(\mathbf{R}_i | \mathbf{Y}_i^o, \mathbf{Y}_i^m, \psi)$$

## ▪ PMM:

◎感度パラメータを設定 → 欠測データの分布を「色々試す」

$$f(\mathbf{Y}_i, \mathbf{R}_i | \theta, \psi) = f(\mathbf{R}_i | \psi) \cdot f(\mathbf{Y}_i^o | \mathbf{R}_i, \theta) \cdot f(\mathbf{Y}_i^m | \mathbf{Y}_i^o, \mathbf{R}_i, \theta)$$

★SM, PMMで感度パラメータを入れる場所が異なる

# 感度パラメータを用いた 欠測メカニズムに対する感度分析 (SM)

$$\begin{aligned} f(\mathbf{Y}_i, \mathbf{R}_i | \boldsymbol{\theta}, \boldsymbol{\psi}) &= f(\mathbf{Y}_i | \boldsymbol{\theta}) \cdot f(\mathbf{R}_i | \mathbf{Y}_i, \boldsymbol{\psi}) \\ &= f(\mathbf{Y}_i^o, \mathbf{Y}_i^m | \boldsymbol{\theta}) \cdot \underbrace{f(\mathbf{R}_i | \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\psi})}_{\text{Type (i) の仮定}} \end{aligned}$$

Type (i) の仮定

$$\begin{aligned} \text{logit}\{Pr(R_{ij} = 0 | R_{i1} = 1, \dots, R_{i,j-1} = 1, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi})\} \\ = \psi_0 + \psi_1 Y_{i,j-1} + \psi_2 Y_{ij} \end{aligned}$$

**感度パラメータ:** 値を解析者が設定  
色々な値を代入して結果の頑健性をみる

# 感度パラメータを用いた 欠測メカニズムに対する感度分析(PMM)

$$f(\mathbf{Y}_i, \mathbf{R}_i | \theta, \psi) = f(\mathbf{R}_i | \psi) \cdot f(\mathbf{Y}_i^o | \mathbf{R}_i, \theta) \cdot \underbrace{f(\mathbf{Y}_i^m | \mathbf{Y}_i^o, \mathbf{R}_i, \theta)}_{\text{Type (i) の仮定}}$$

2時点の場合で考える

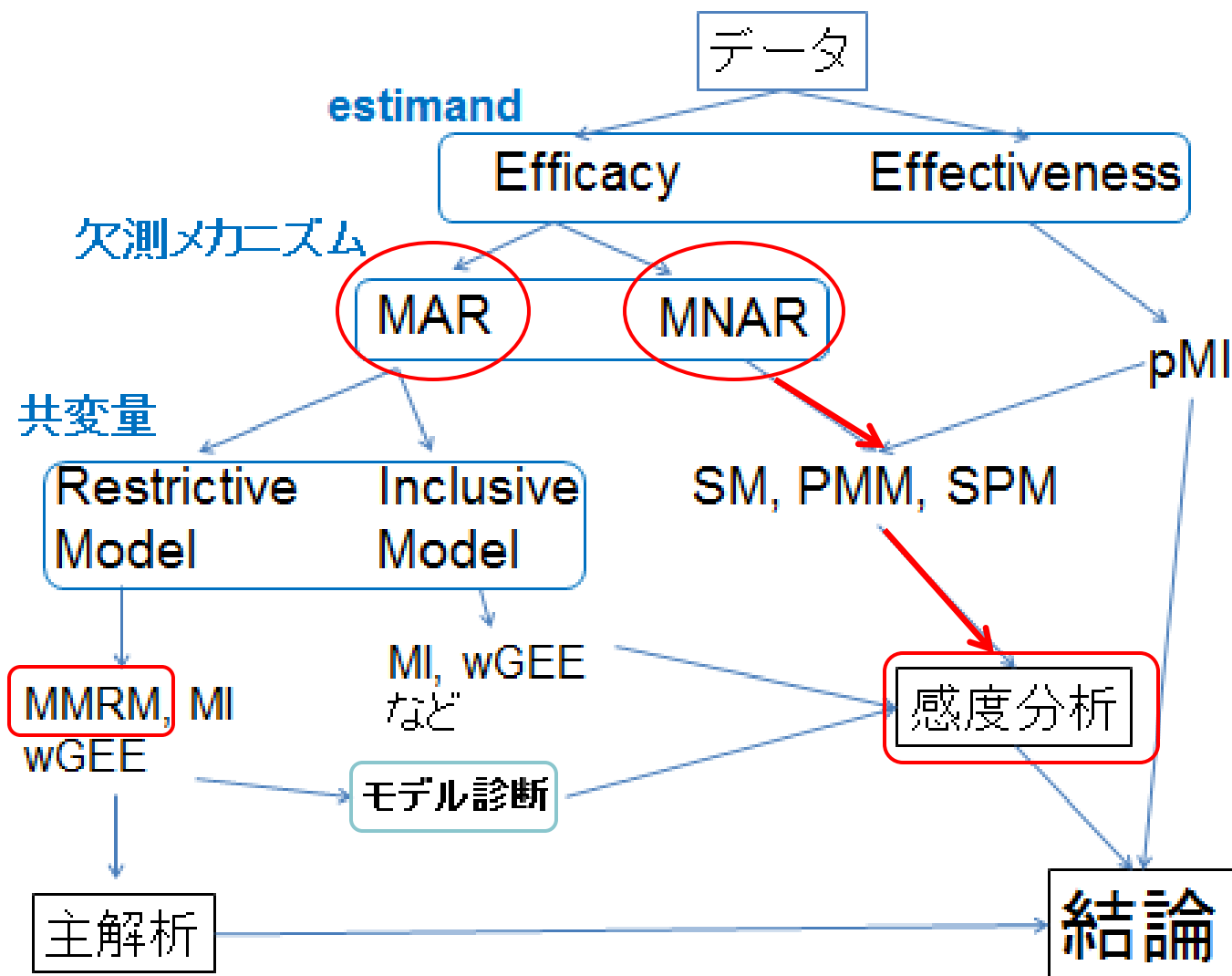
$Y_0$ : pre,  $Y_1$ : post

Type (i) の仮定

$$\underbrace{E(Y_1 | Y_0, R = 0)}_{\text{欠測データの平均}} = \underbrace{E(Y_1 | Y_0, R = 1)}_{\text{観測データの平均}} + \underbrace{\Delta}_{\text{感度パラメータ}}$$

- ✓ 欠測データはNFMV + ACMV ( $\Leftrightarrow$  ACMV  $\Leftrightarrow$  MAR)の制約条件で補完.
- ✓ さらにMARからのずれを $\Delta$ で表現する( $\Delta=0 \Leftrightarrow$  MAR)

# 欠測メカニズムに対する感度分析 Analytic roadmapでおさらい



1. 事例の試験デザイン、主解析、感度分析
2. 感度分析：
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈



# Estimandに対する感度分析

| Estimand | 仮説            | 推測の対象                | 被験者               | 評価時点        | Rescue Medication後のデータ |
|----------|---------------|----------------------|-------------------|-------------|------------------------|
| 1        | Effectiveness | 割り付け群<br>(治療方針)      | 全被験者              | 計画された<br>時点 | 主解析に<br>含める            |
| 2        | Efficacy      | 最初に<br>割り付けられた<br>治療 | 最初に耐えられ<br>た被験者のみ | 計画された<br>時点 | 主解析に<br>含めない           |
| 3        | Efficacy      | 最初に<br>割り付けられた<br>治療 | 全被験者              | 計画された<br>時点 | 主解析に<br>含めない           |
| 4        | Effectiveness | 最初に<br>割り付けられた<br>治療 | 全被験者              | 未定義         | 主解析に<br>含めない           |
| 5        | Effectiveness | 最初に<br>割り付けられた<br>治療 | 全被験者              | 未定義         | 主解析に<br>含めない           |
| 6        | Effectiveness | 最初に<br>割り付けられた<br>治療 | 全被験者              | 計画された<br>時点 | 補完することが<br>望ましい        |

# Estimandに対する感度分析

---

**Efficacy**: 計画通りに投与された薬剤の効果

→ per-protocol estimand (de jure)

**Effectiveness**: 実際に投与された薬剤の効果

→ ITT estimand (de facto)

## 疑問点

Estimand間の比較は感度分析？

→ 試験の頑健性を確認する上で有用な場合もある。

(Mallinckrodt 2013)

# Estimandに対する感度分析

✓ 主解析: Estimand **3** (Efficacy)

→16週投与した上での有効性を評価

(全例完了したことを想定)

✓ 感度分析: Estimand **6** (Effectiveness)

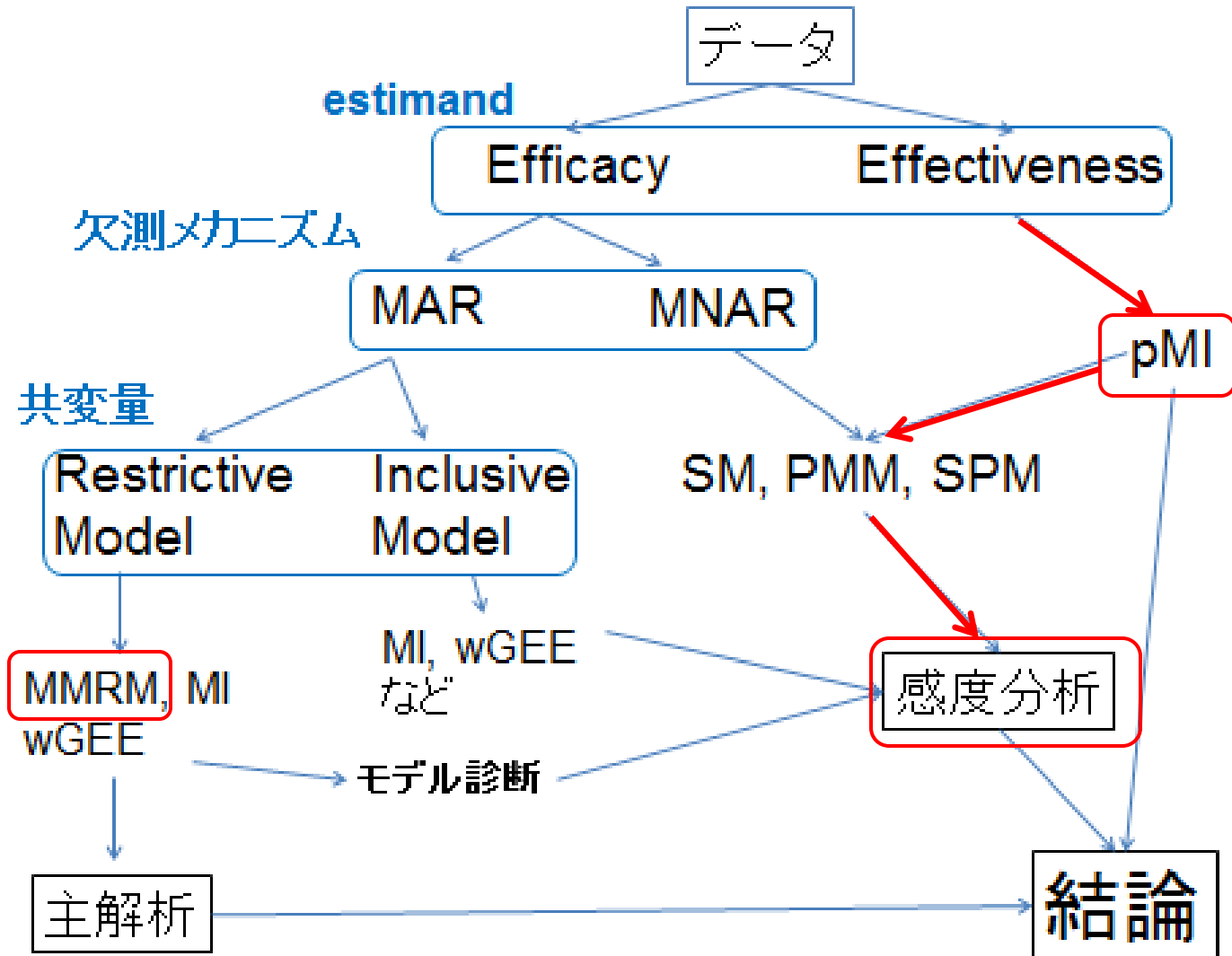
→**中止症例の影響**を考慮したもとの有効性を評価

→中止後のデータは**プラセボ群のデータを用いて多重補完**

**する** Placebo multiple imputation(pMI)を実行.その後

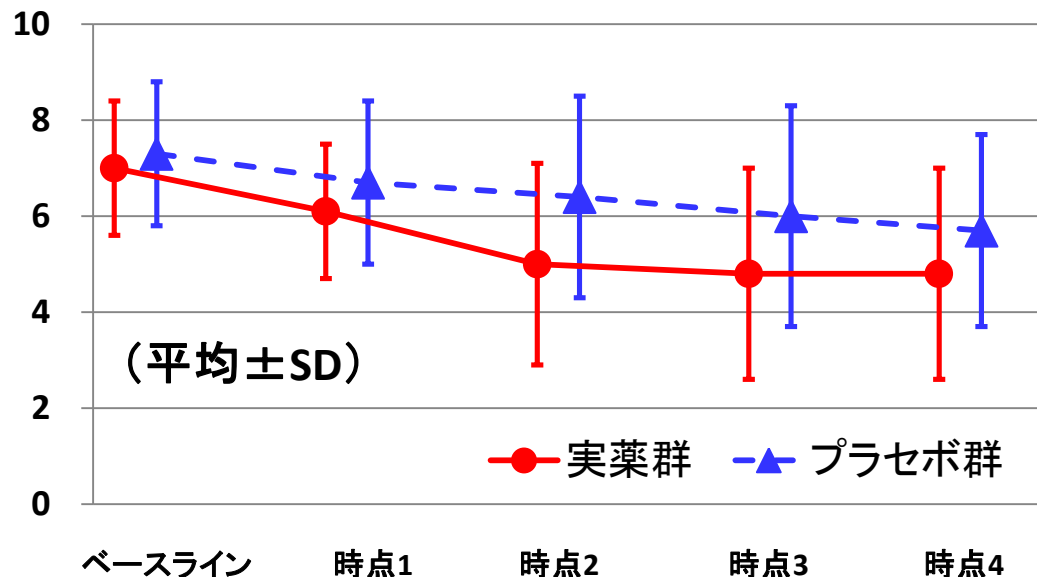
MMRMにより解析する.

# Estimandに対する感度分析 Analytic roadmapでおさらい



1. 事例の試験デザイン、主解析、感度分析
2. 感度分析：
  - Analytic road mapに記載されている感度分析—
  - A) 完全データの分布(モデル適合含む)
  - B) 欠測メカニズム
  - C) Estimand
3. 解析結果と解釈

## ◎慢性疼痛の第III相試験を想定したシミュレーションデータ



● 主要評価項目：疼痛スコア  
 → **スコア低下：改善**  
 (解析には変化量使用)

● 実薬群 vs プラセボ群  
 ・1群100例(ベースライン時)

◎単調な欠測のみ

|       | ベースライン |           | 時点1 |           | 時点2 |           | 時点3 |           | 時点4 |           |
|-------|--------|-----------|-----|-----------|-----|-----------|-----|-----------|-----|-----------|
|       | 例数     | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   | 例数  | 平均 (SD)   |
| 実薬群   | 100    | 7.0 (1.4) | 80  | 6.1 (1.4) | 72  | 5.0 (2.1) | 69  | 4.8 (2.2) | 67  | 4.8 (2.2) |
| プラセボ群 | 100    | 7.3 (1.5) | 80  | 6.7 (1.7) | 67  | 6.4 (2.1) | 58  | 6.0 (2.3) | 52  | 5.7 (2.0) |

# 主解析の詳細

|          |   |
|----------|---|
| Estimand | 計画している最終時点での群間差(全症例完了を想定)<br>Estimand 3(efficacy) |
| 欠測メカニズム  | MAR   |
| モデル      | MMRM  |
| 共変量      | ベースラインの疼痛スコア、治験薬、時点、治験薬×時点                        |
| 分散共分散構造  | Unstructured                                      |
| 有意水準     | 両側5%  |

# 主解析・感度分析の一覧

|       | 仮説            | Estimand      | 欠測<br>メカニズム | 解析<br>手法         | その他             |
|-------|---------------|---------------|-------------|------------------|-----------------|
| 主解析   | Efficacy      | 3 (primary)   | MAR         | MMRM             |                 |
| 感度分析1 | Efficacy      |               | MAR         | MMRM             | 残差診断<br>・影響診断   |
| 感度分析2 | Efficacy      |               | MNAR        | SM               | 感度パラメータ<br>-1~1 |
|       |               |               | MNAR        | PMM<br>+<br>MMRM | 感度パラメータ<br>-1~1 |
| 感度分析3 | Effectiveness | 6 (secondary) | MNAR        | pMI<br>+<br>MMRM |                 |

◎3種類の感度分析を実行



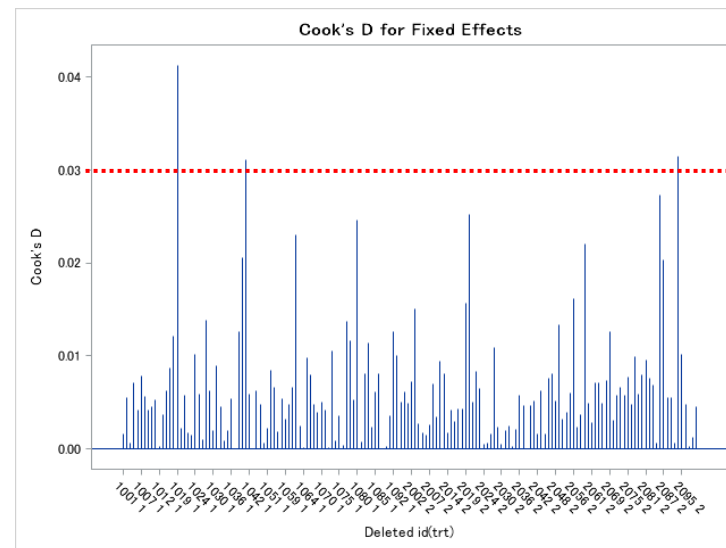
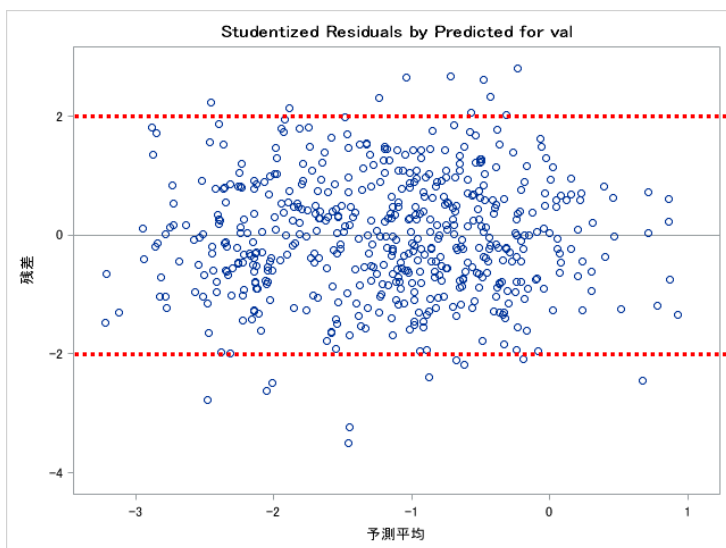
# 主解析の結果

| 解析手法 | 時点4における各群の<br>点推定値(SE) |             | 群間差   | 群間差<br>のSE | P値    |
|------|------------------------|-------------|-------|------------|-------|
|      | 実薬群                    | プラセボ群       |       |            |       |
| MMRM | -2.04(0.25)            | -1.03(0.27) | -1.02 | 0.37       | 0.006 |



この推定値の頑健性を確認する

# 感度分析1(残差診断・影響診断)の結果



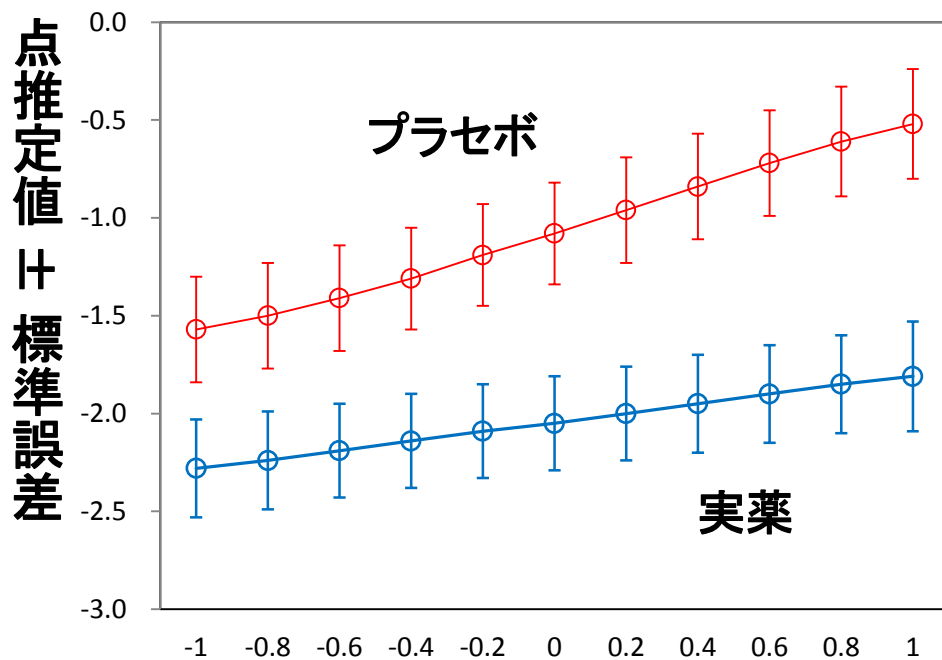
|                    | 除外判定基準               | 症例数 | データ数 | 群間差   | 群間差のSE |
|--------------------|----------------------|-----|------|-------|--------|
| 主解析                |                      | 160 | 545  | -1.02 | 0.37   |
| 感度分析1<br>(外れ値除外)   | Student化残差<br>絶対値2以上 | 159 | 525  | -0.94 | 0.34   |
| 感度分析1<br>(外れた症例除外) | Cook's D<br>0.03以上   | 157 | 533  | -0.95 | 0.35   |

# 感度分析2(欠測メカニズムの検討)の結果

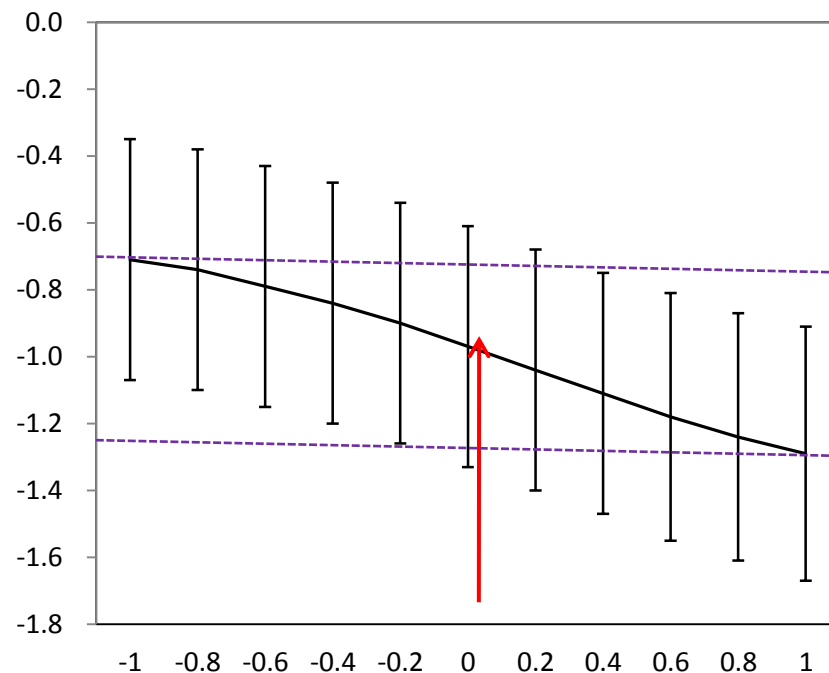
SM



## 各群



## 群間差



感度パラメータ  $\Delta$ : -1~1

$\Delta=0$ のとき、群間差の点推定値 = -0.97

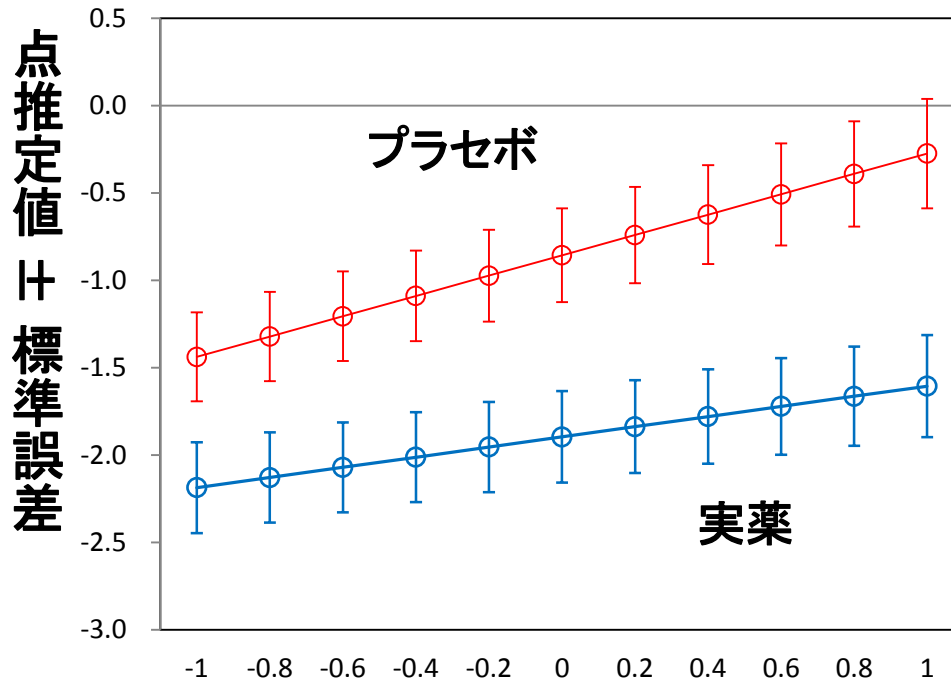
Min~Max: -0.71~-1.29

# 感度分析2(欠測メカニズムの検討)の結果

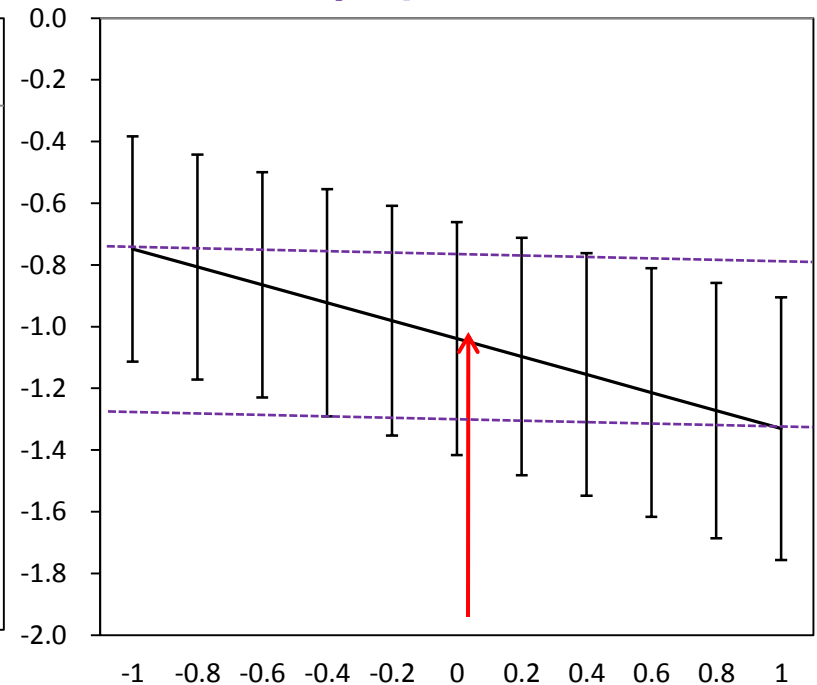
## PMM



### 各群



### 群間差



感度パラメータ  $\Delta$ : -1~1

$\Delta=0$ のとき、群間差の点推定値=-1.04

Min~Max: -0.74~-1.33

# 感度分析3 (estimandに対する感度分析)の結果

| 解析手法 | 時点4における各群の点推定値(SE) |             | 群間差   | 群間差のSE | P値    |
|------|--------------------|-------------|-------|--------|-------|
|      | 実薬群                | プラセボ群       |       |        |       |
| pMI  | -2.13(0.25)        | -1.16(0.27) | -0.97 | 0.36   | 0.008 |

pMIにより補完後, 完全データを用いてMMRMにより解析

# 主解析・感度分析の結果

|       | 解析手法 | その他           | 群間差   |       |       |
|-------|------|---------------|-------|-------|-------|
|       |      |               |       | 最小    | 最大    |
| 主解析   | MMRM |               | -1.02 |       |       |
| 感度分析1 | MMRM | 残差診断・<br>影響診断 |       | -0.94 | -0.95 |
| 感度分析2 | SM   | 感度パラメータ：-1~1  |       | -0.71 | -1.29 |
|       | PMM  | 感度パラメータ：-1~1  |       | -0.74 | -1.33 |
| 感度分析3 | pMI  |               | -0.97 |       |       |

→いずれの感度分析の結果についても主解析の結果と大きく乖離するものではないと考えられ、本薬剤の有効性の結果に対して**頑健性が認められる**と考えられた。

# 感度分析を行う際の留意点

---

- ・ 試験の目的と主解析で仮定している内容を明確化
- ・ 感度分析では何の仮定に対する感度を見ているか、意識
- ・ 感度分析の内容(何を確認すべきか)は疾患、試験デザイン、解析内容によって変わる
- ・ 感度パラメータ $\Delta$ の範囲は臨床的に意味のある範囲となるよう事前に検討

# 欠測メカニズムに対する感度分析での 考察方法

---

形式的な手順は確立していないが、以下の内容を考察することがNRC (2010)とMallinckrodt (2013)で提案されている。

- ・ 群間の試験中止割合や試験中止理由の確認
- ・ 個々の感度分析での結果の変動(点推定値及び信頼区間の変動が臨床的に意味のある範囲であるか)



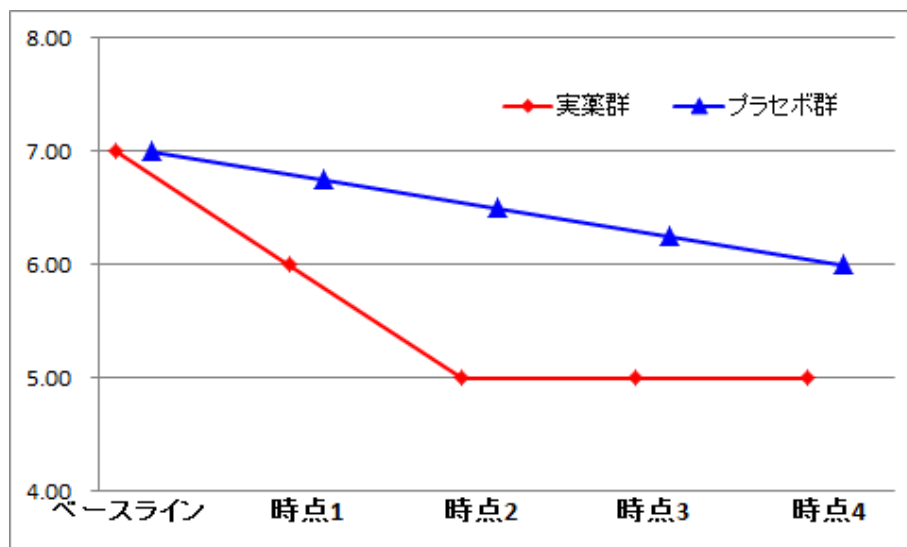
## 架空の事例の検討1

- **対象疾患**: 慢性疼痛
- **中止割合**: 10~50% (効果不十分, 有害事象)
- **主要評価項目**: 投与16週時点の疼痛スコアのベースラインからの変化量
- **Estimand**: Estimand 3 (Efficacy)
- **試験デザイン**: プラセボ対照ランダム化二重盲検並行群間試験
- **欠測メカニズムの仮定**: MAR
- **主要な解析**: MMRM

# 架空の事例の検討 まとめ

## 架空の事例の検討2

### - シミュレーションデータの提示



2通りの欠測メカニズムに基づく  
データを作成  
MCAR+MARの混合  
MCAR+MNARの混合

### 時点毎の累積欠測確率

| 群     | ベースライン | 時点1  | 時点2  | 時点3  | 時点4  |
|-------|--------|------|------|------|------|
| 実薬群   | 0 %    | 20 % | 26 % | 30 % | 34 % |
| プラセボ群 | 0 %    | 18 % | 30 % | 40 % | 48 % |

## 架空の事例の検討2

- 主解析の選択のシミュレーション (MMRM, MI, wGEE, LOCF ANCOVA, OC ANCOVA)
- $\alpha$  エラー, 検出力, 推定精度の結果より **MMRM** を選択
- シミュレーションの際, 考慮した因子:
  - 平均構造, 分散共分散構造, 欠測メカニズム,
  - 脱落確率の関数, 来院間隔, 来院回数

## 架空の事例の検討2

- Lu et al (2008) を用いてMMRMに基づく例数設計を実施
- 分散を設定する際, 時点ごとの残存率(または脱落割合)と時点間相関の構造を考慮.

# 架空の事例の検討 まとめ

## 架空の事例の検討2

### - シミュレーションデータの発生方法



必要最低限の  
時点

平均ベクトル  
分散共分散構造

欠測確率  
欠測メカニズム  
欠測確率の関数

# 架空の事例の検討 まとめ

## 架空の事例の検討3

- Analytic Road Mapに従って, 主解析, 感度分析を実行.

### Analytic Road Map

Mallinckrodt (2013)より  
一部編集

