

第8回 データサイエンスラウンドテーブル会議

# 『アダプティブデザインにおけるバイアスの検討』 事前説明資料

上村 鋼平<sup>1</sup>, 加藤 凌輔<sup>2</sup>,

○中村 将俊<sup>3,6</sup>, 高津 正寛<sup>4,6</sup>, 田中 勇輔<sup>5,6</sup>,

---

<sup>1</sup>東京大学, <sup>2</sup>PMDA,

<sup>3</sup>ファイザーR&D合同会社, <sup>4</sup>持田製薬株式会社, <sup>5</sup>アステラス製薬株式会社,

<sup>6</sup>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

# はじめに

---

本発表内容及びディスカッションでの発言は、個人の見解に基づくものであり、当事者が所属する組織の公式見解ではございません。

# はじめに

---

アダプテーションに関連するバイアスの低減又は除去のための推定値の調整方法は提案されている一方で、実際に使用された事例は限られている。

FDAガイダンスでは

“ベネフィット・リスクの評価を円滑に進め、新薬の適切なラベリングのために根拠に基づく医療の実践を可能にする、十分に信頼できる治療効果の推定値を算出することが重要である”  
と推定値へのバイアスを考慮し解析を計画することの重要性を述べている。

本日は、**治療効果を示す推定値としてふさわしいものは何か**という観点から、推定値のバイアス調整の利点/欠点について議論したい。

# そもそもバイアスを調整することについての産官学目線の認識

---

実際の事例を用いたディスカッションの前に、データサイエンスラウンドテーブルのテーマでもある

## 産・官・学

のそれぞれの目線から、バイアスを調整することについての認識について、またそれぞれ異なる分野の方々へ、何を求めている(期待している)かを、ファシリテーターのメンバーから簡単にご紹介頂きます。

# アカデミアの目線から

## 規制当局へ

- 試験継続の可否判断のみの中間解析を含む試験におけるバイアス調整の現状との関係性はどうか考えるべきか
- 今後、アダプティブデザイン一般にバイアス補正が必要と考える場合、実務的にどのような議論が必要と考えるか

## 企業へ

- 実装の問題へのインパクトは規制当局の要求レベル次第か
- 主要解析結果として求められるのであれば、産官学で技術的な議論を行い、少なくともある程度標準的な手法の確立が必要ではないか
- バイアス調整した結果を提示したいというニーズはないのか

# 規制当局の目線から

## アカデミアへ

- (事前規定の範囲内であるが)試験開始前の想定とズレた検証をすることについて、現在提案されている推定値のバイアス調整方法は、結果解釈の手助けにもなり得るか。
- 調整方法の妥当性を説明する上で、シミュレーションの検討範囲等、合理的な説明をしていくことが可能か。
- 上記のようなコンセンサスの得られる基準等を決めた上で、事前規定することは可能か。

## 企業へ

- バイアスの調整をした推定値を算出することを、主解析として採用する目処や取り組み等はあるか。
- 実際に推定値のバイアスの調整を考慮する際に、事前規定できる内容の限界はあるか。
- 上記のような対応を検討していく過程で、企業内でハードルが高い事項はあるか。

# 企業の目線から

## アカデミアへ

- バイアスの調整方法の考究
- 解析ソフトへの実装
- 調整方法の性能評価も必要な部分ではありますが、プログラムを自作しなければならないところが、調整方法を利用する際のハードルになっています。

## 規制当局へ

- 相談の方法・ガイドラインの規定
- 添付文書などでの記載
- バイアスの未調整の推定値を薬剤の効果として評価すべきか、調整した推定値を薬剤の効果として評価すべきか、疾患領域・主要な解析・Estimandなどにより状況は様々だが一緒に検討して頂きたい。

# 本日のグループディスカッションの内容

---

実際に

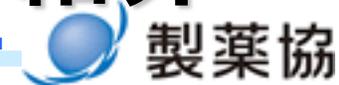
SSR (Sample size re-estimation : 症例数再設計)

TS (Treatment selection : 治療群選択)

を用いた2つのアダプティブデザインの事例を題材に、推定値のバイアス調整の利点/欠点について議論して頂きたいと思います。

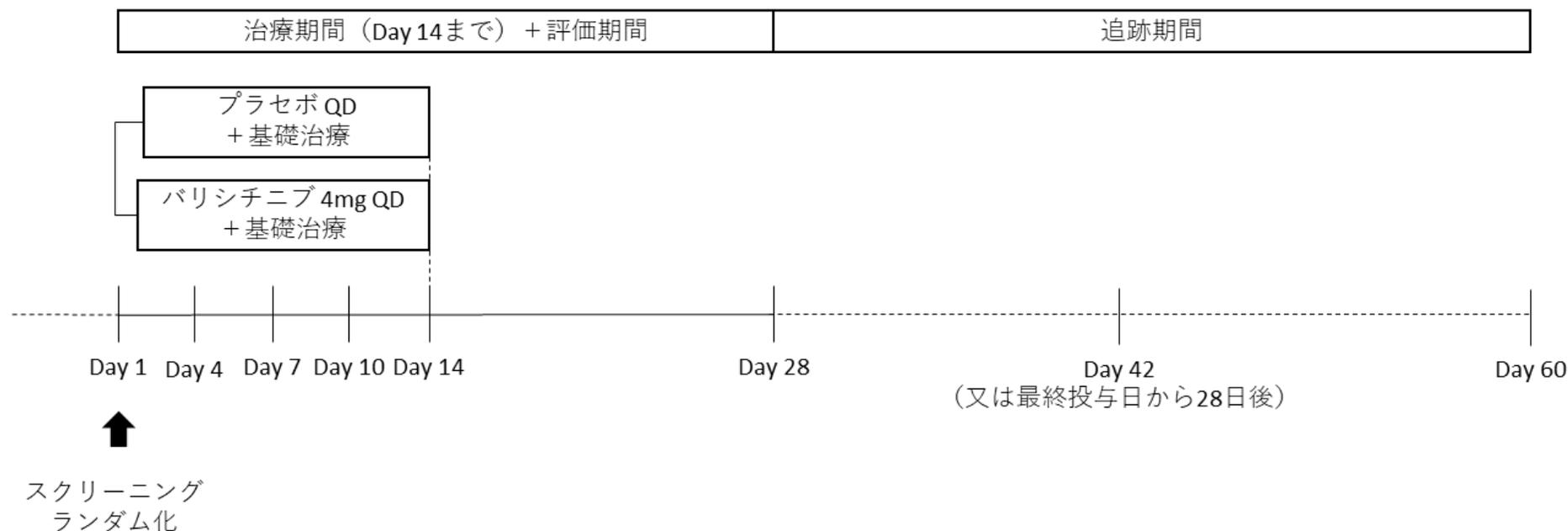
これから紹介する2つの臨床試験において、バイアスの調整をした推定値を試験結果の推定値として用いた場合の利点/欠点を整理頂き、各グループでディスカッションした内容について発表資料を作成・発表して頂きたいと思います。

# 症例数再推定のアダプティブデザイン： バイアスの調整方法とシミュレーション事例の紹介



# 参考事例：COV-BARRIER試験

試験表題： COVID-19に感染した患者を対象としたバリシチニブの二重盲検プラセボ対照ランダム化比較第III相試験



## 主要評価項目

Day28までに死亡，又は非侵襲的換気/高流量酸素療法もしくは侵襲的機械的換気[体外式膜型人工肺(ECMO)を含む]を必要とした被験者の割合

## 解析方法

主要評価項目に対して，各群で頻度及び割合を算出し，オッズ比とその信頼区間を算出する。また，治療群間の比較には層別因子(病期，年齢，地域，デキサメタゾン及び/又は全身性コルチコステロイドの使用の有無)で調整したロジスティック回帰モデルを用いて，両側有意水準5%で検定を行う。中間解析の結果として症例数が増加した場合，Cui, Hung and Wang(1999)の方法を適用し，第一種の過誤確率を制御する。

## SSRを用いた目的

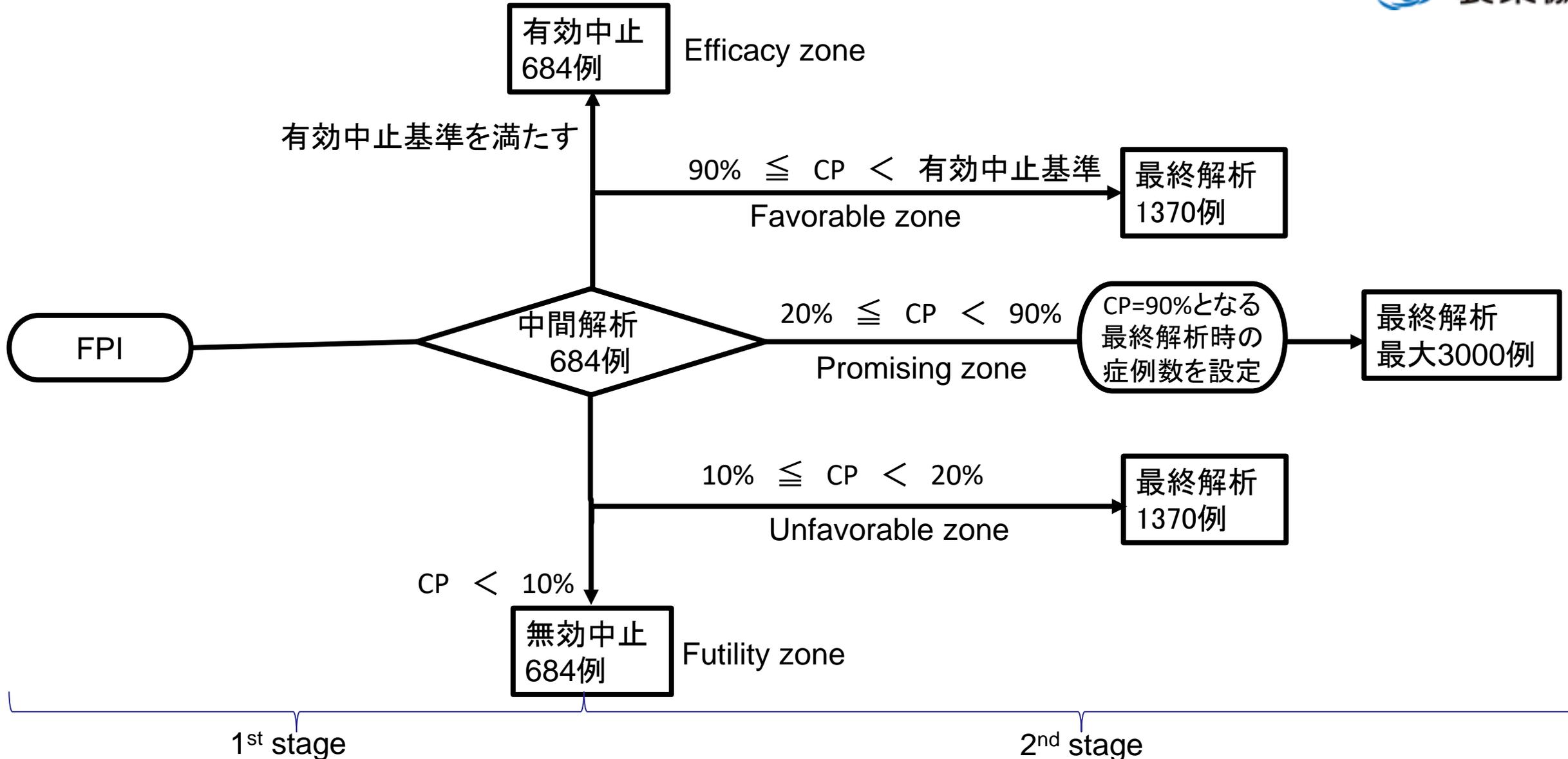
- ある併用薬を使用した患者の治療効果が利用可能なデータから正確に想定できないため、治療効果の大きさに応じた症例数再設計を可能にするデザインを検討したい。

主要評価項目の発現割合の差		統合した治療効果
ある併用薬を使用した患者 (全体の約6割):	ある併用薬を使用しなかった患者 (全体の約4割):	
7.5%	7.5%	7.5%
4.0% 	7.5%	5.4%

## 中間解析及び判定基準

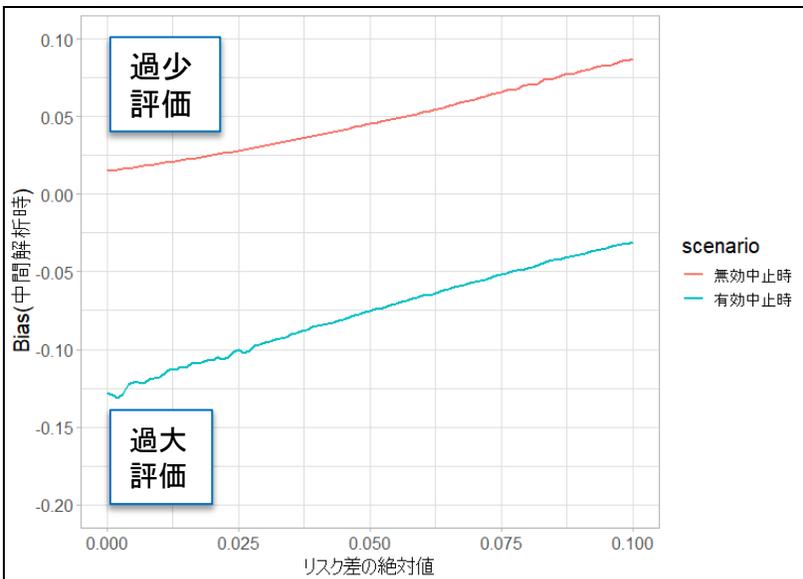
総症例数	1370例(治療効果の見積もり値として割合の差7.5%を設定し算出)
中間解析の回数及び目的	1回(有効及び無効中止)
中間解析の時期	全体症例の50%(684例)集積時
有効中止の基準	O'Brien-Fleming法に基づく棄却限界値
無効中止の基準	条件付き検出力(CP)が10%未満の場合に中止
症例数再推定の方法	Unfavorable zone : 無効中止基準 $\leq$ CP < 20% 症例数は変更しない Promising zone : 20% $\leq$ CP < 90% CP=90%となる最終解析時の症例数を設定(最大例数3000例) Favorable zone : 90% $\leq$ CP < 有効中止基準 症例数は変更しない

# シミュレーションの条件: Five Zone Design

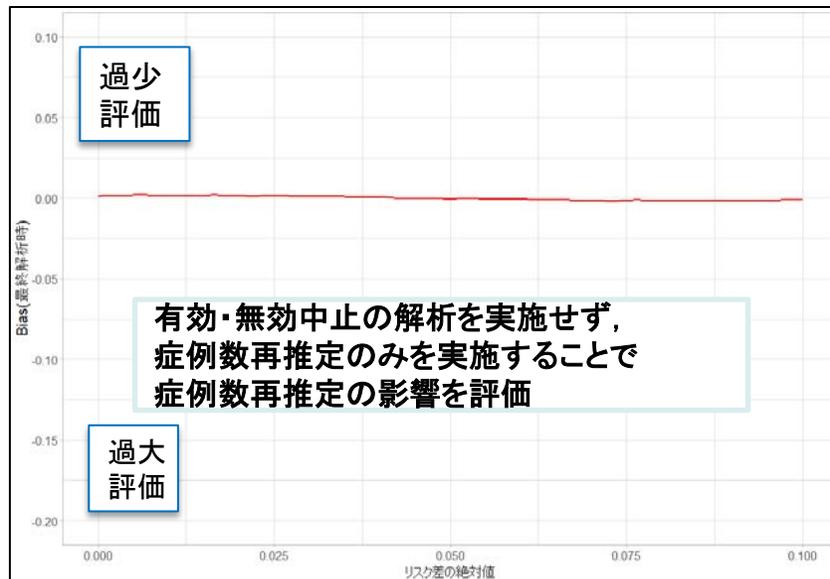


# バイアスが発生する原因

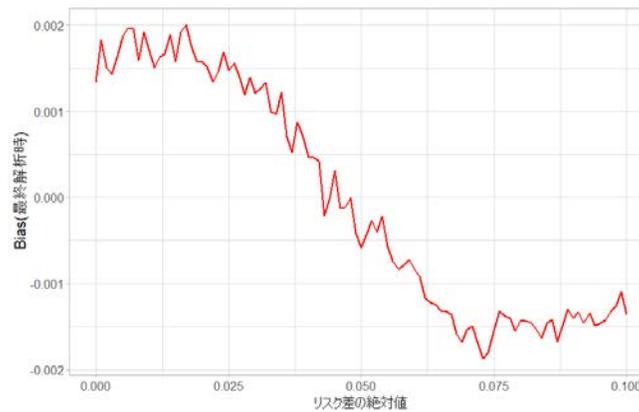
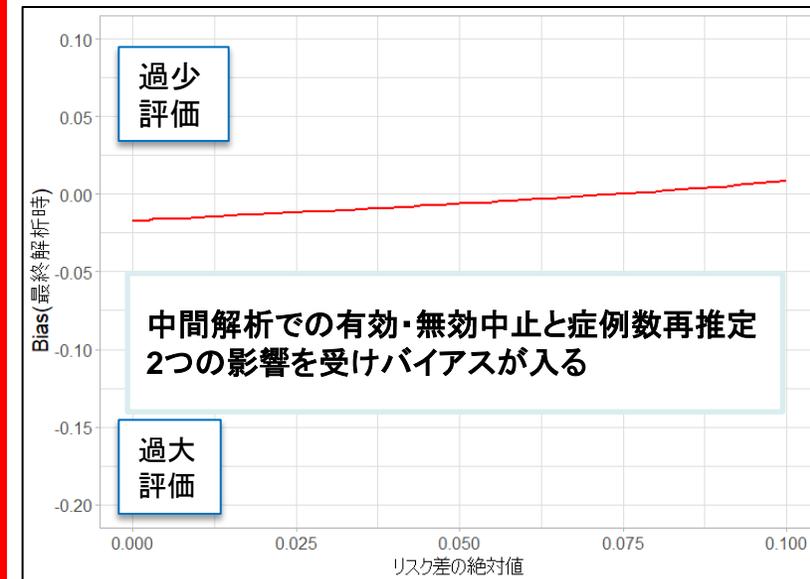
## 中間解析での有効・無効中止



## 症例数再推定



## 最終解析時



今回のシミュレーションのスコープ

# 最終解析時の推定値のバイアスを調整する方法

## ➤ Bias-corrected MLE (CBC-MLE)

$$\tilde{\theta}_c = \hat{\theta}_{obs} + \frac{\sqrt{I_1}}{I_2} \times \frac{\phi(e - \tilde{\theta}_c \sqrt{I_1}) - \phi(l - \tilde{\theta}_c \sqrt{I_1})}{\Phi(e - \tilde{\theta}_c \sqrt{I_1}) - \Phi(l - \tilde{\theta}_c \sqrt{I_1})} \text{の方程式の } \tilde{\theta}_c \text{ に関する解で与えられる}$$

## ➤ Uniformly minimum variance conditionally unbiased estimator (UMVCUE)

$$\text{UMCUE} = \hat{\theta}_{obs} - w_1 \frac{\phi(w_2(\hat{\theta}_{obs} - e/\sqrt{I_1})) - \phi(w_2(\hat{\theta}_{obs} - l/\sqrt{I_1}))}{\Phi(w_2(\hat{\theta}_{obs} - e/\sqrt{I_1})) - \Phi(w_2(\hat{\theta}_{obs} - l/\sqrt{I_1}))}$$
$$w_1 = 1/((I_2 - I_1)\sqrt{I_1^{-1}} + (I_2 - I_1)^{-1}), w_2 = I_1\sqrt{I_1^{-1}} + (I_2 - I_1)^{-1}$$

## ➤ Conditional median unbiased estimator (CMUE)

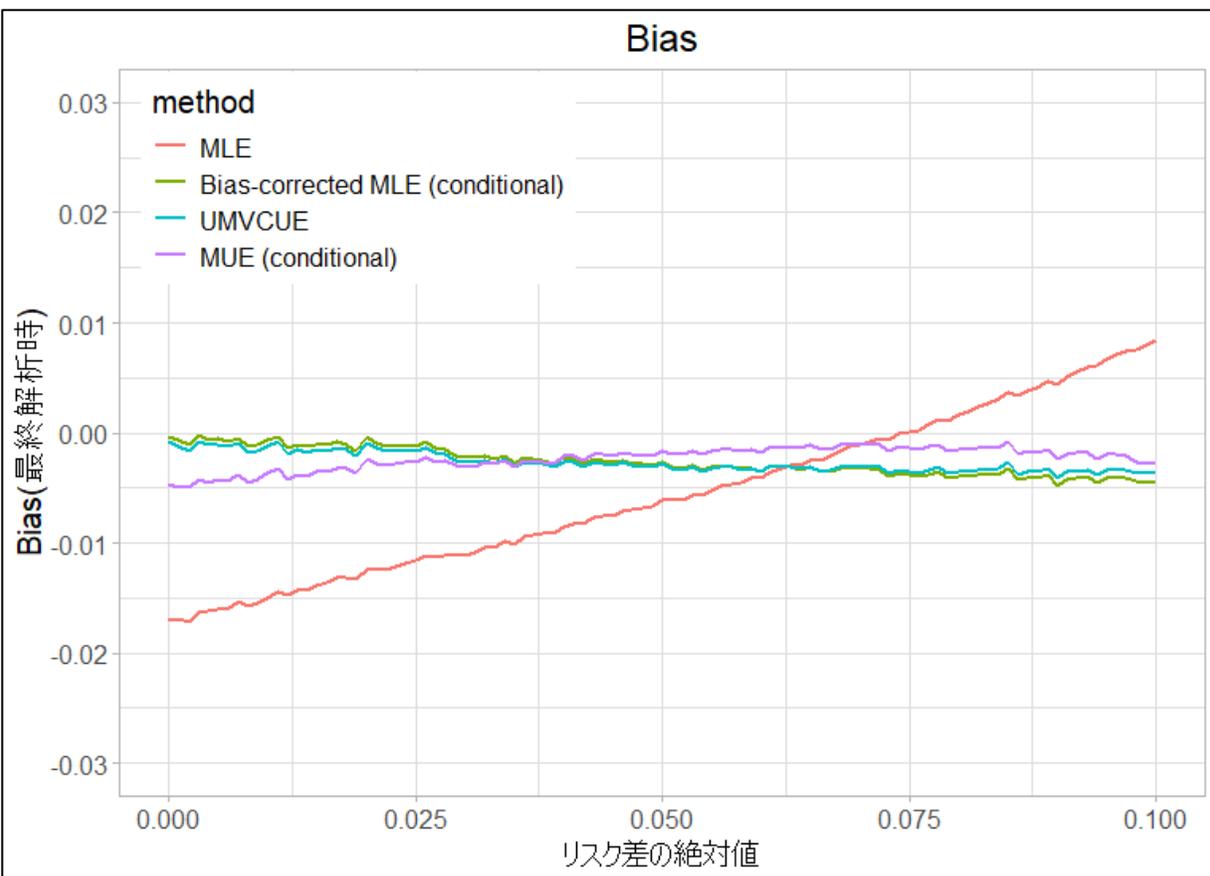
■ 以下の方程式の $\theta$ に関する解で与えられる

$$0.5 = \int_{-\infty}^{\hat{\theta}_{obs}} \frac{\Phi\left(\frac{\hat{\theta} - e/\sqrt{I_1}}{1/I_1 - 1/I_2}\right) - \Phi\left(\frac{\hat{\theta} - l/\sqrt{I_1}}{1/I_1 - 1/I_2}\right)}{\Phi(\theta\sqrt{I_2} - e) - \Phi(\theta\sqrt{I_2} - l)} \times \frac{\exp(-I_2(\hat{\theta} - \theta)^2/2)}{\sqrt{2\pi/I_2}} d\hat{\theta}$$

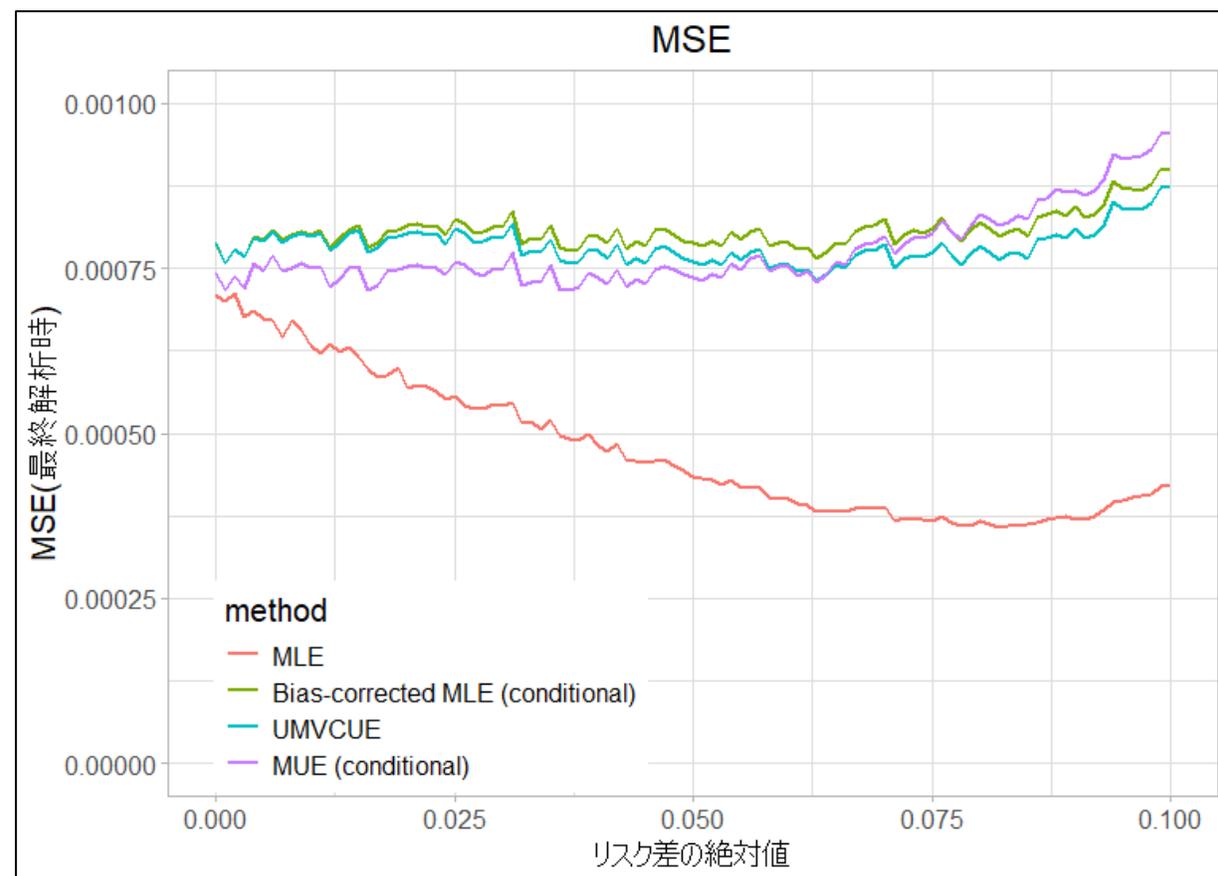
$I_1$ : 中間解析時の情報量,  $I_2$ : 最終解析時の情報量  
 $e$ : 有効中止の境界,  $l$ : 無効中止の境界  
 $\hat{\theta}_{obs}$ : ナイーブな推定量 (MLE)

# 最終解析時の治療効果の推定(バイアス/MSEの評価)

### Bias



### MSE



- 有効中止・無効中止時は大きなバイアスが発生する
- 最終解析時のバイアスを調整する方法を用いると
  - バイアス: どの補正法もMLEと比較して**減少**の傾向にある
  - MSE: どの補正法もMLEと比較して**増大**の傾向にある

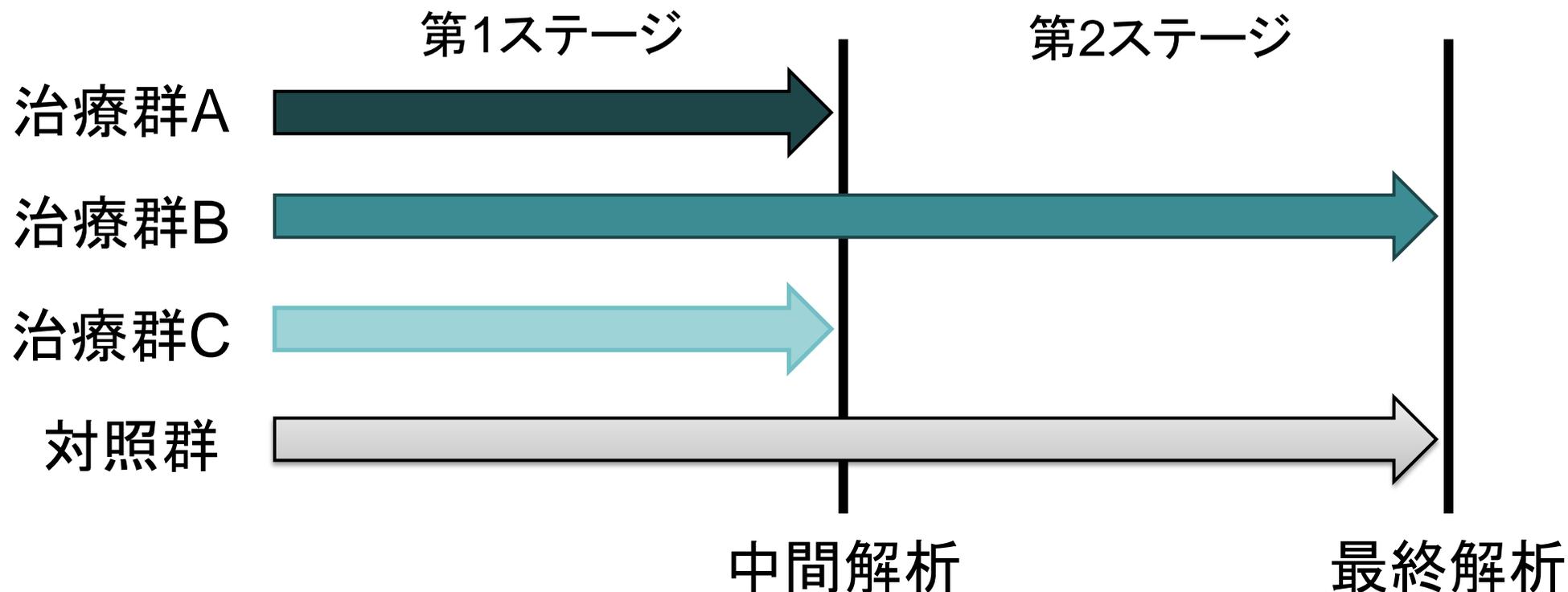
# 治療群選択のアダプティブデザイン： バイアスの調整方法とシミュレーション事例の紹介



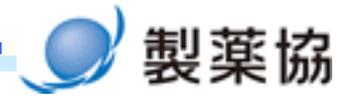
持田製薬株式会社 高津 正寛

# 治療群選択のアダプティブデザイン(TSデザイン)

- 複数の治療群を設定して試験を開始
- 中間解析にて、事前に計画した基準に基づいて治療群を選択
- 最終解析にて、選択された治療群と対照群を比較して効果を検証する



# TSデザインにおけるバイアスの定義 (Bauer et al. (2010))



## ➤ 選択バイアス (Selection bias)

### ■ 選択された条件付きの治療効果の推定量に生じるバイアス

- 第1ステージの効果が高いために選択される→過大評価のバイアス
- 第2ステージの推定値にはバイアスが含まれないが、分散が大きい

## ➤ 報告バイアス (Reporting bias)

### ■ 選択されたか否かに関わらず、治療効果の推定量に生じるバイアス

- 選択群 : 選択バイアスと同様
- 非選択群 : 第1ステージの効果が低いために選択されない→過小評価のバイアス

※以降では、選択された条件付きの治療効果の推定について取り扱う

# バイアスを調整する方法 (TSデザイン)

## ➤ Cohen & Sackrowitz (1989)

- 2ステージデザイン、正規分布のエンドポイント
- 第1ステージで最も優れた治療群を1つ選択する場合におけるUMVCUE
  - 一様最小分散条件付き不偏推定量;  
Uniformly Minimum Variance Conditionally Unbiased Estimator
- ステージ間で被験者数均等などの制約

## ➤ Bowden & Glimm (2008)

- 第1ステージでJ番目に優れた治療群のUMVCUEを提案
  - 安全性などを含む任意の選択ルールを利用可能
- 分散既知を想定; ステージ間での被験者数不均等も可

## ➤ 第1ステージで最も優れた群のMLE (ナイーブな推定量)

$$\hat{\mu}_{(1)} = \frac{n_1 X_{(1)} + n_2 Y}{n_1 + n_2}$$

$X_{(1)}$  : 第1ステージにおける最も優れた群 (= 選択された群) の推定値  
 $Y$  : 選択された群の第2ステージにおける推定値  
 $n_1, n_2$  : 各ステージの例数/群

## ➤ 第1ステージで最も優れた群のUMVCUE (Cohen & Sackrowitz (1989))

$$\tilde{\mu}_{(1)} = \frac{X_{(1)} + Y}{2} - \frac{1}{\sqrt{2}} \frac{\phi(W)}{\Phi(W)}$$

MLE推定量 - バイアス補正項

ただし、

$$W = \sqrt{2} \left( (X_{(1)} + Y) / 2 - X_{(2)} \right)$$

$\tilde{\mu}_{(1)}$  の推定に  $X_{(1)}$  と  $X_{(2)}$  を利用

$\phi, \Phi$  : 標準正規分布の確率密度関数および累積分布関数  
 $X_{(2)}$  : 第1ステージで2番目に優れた群の推定値

## ➤ 第1ステージでJ番目に優れた治療群のUMVCUE (Bowden & Glimm (2008))

$$\tilde{\mu}_{(j)} = \frac{\sigma_{2,j}^2 X_{(j)} + \sigma_{1,(j)}^2 Y_j}{\sigma_{1,(j)}^2 + \sigma_{2,j}^2} - \frac{\sigma_{2,j}^2}{\sqrt{\sigma_{1,(j)}^2 + \sigma_{2,j}^2}} \frac{\{\phi(W_{j,j+1}) - \phi(W_{j,j-1})\}}{\{\Phi(W_{j,j+1}) - \Phi(W_{j,j-1})\}}$$

ただし、

$$W_{s,t} = \frac{1}{\sigma_{1,(s)}^2} \left( \frac{\sigma_{2,s}^2 X_{(s)} + \sigma_{1,(s)}^2 Y_s}{\sqrt{\sigma_{1,(s)}^2 + \sigma_{2,s}^2}} - X_{(t)} \sqrt{\sigma_{1,(s)}^2 + \sigma_{2,s}^2} \right) \quad (s = 1, \dots, k, t = 0, \dots, k+1)$$

$$X_{(0)} := \infty > X_{(1)} \geq \dots \geq X_{(k)} > X_{(k+1)} := -\infty$$

$\tilde{\mu}_{(j)}$  の推定に  $X_{(j)}, X_{(j-1)}, X_{(j+1)}$  を利用

$X_i$  : 互いに独立に  $N(\mu_i, \sigma_{1,i}^2)$  に従う群  $i$  の第1ステージにおける推定値 ( $i = 1, \dots, k$ )  
 $Y_i$  :  $N(\mu_{(i)}, \sigma_{2,i}^2)$  に従う  $i$  番目に優れた群の第2ステージにおける推定値  
 $\sigma_{1,(s)}^2$  :  $s$  番目に優れた群の推定値の分散

# シミュレーションによる推定値の性能評価

➤ INHANCE (慢性閉塞性肺疾患患者対象シームレスP2/3)を下地に設定

パラメータ	設定
有意水準 ( $\alpha$ )	0.025 (片側)
群	第1ステージ: 6群 (試験治療 4用量, プラセボ群, 陽性対照群) 第2ステージ: 3-4群 (選択された試験治療 1-2用量, プラセボ群, 陽性対照群)
第1, 2ステージの例数/群	110, 230
中間解析の目的	最終解析において検証する試験治療群の選択 (※有効中止なし)
用量選択ルール	選択条件: 12週時トラフFEV <sub>1</sub> のプラセボとの差 (エフェクトサイズ) が0.4よりも大きいかつ 陽性対照との差が0より大きい  選択ルール: <u>選択条件を満たす最低用量①と, ①より1つ高い用量②の2群</u>  ※ ②は選択条件を満たしていなくても選択可能, ②がなければ1群のみ選択 ※ 条件を満たす群が1つもなければ無効中止
最終解析の目的	12週時トラフFEV <sub>1</sub> に基づく, 選択用量のプラセボに対する優越性の検証

# 設定した臨床シナリオ

## ➤ 臨床シナリオ

←低用量                      高用量→

No.	用量1	用量2	用量3	用量4	陽性対照	備考
1	0.4	0.433	0.467	0.5	0.4	直線的な用量反応
2	0.4	0.433	0.467	0.467	0.4	用量3で頭打ち
3	0.4	0.433	0.433	0.433	0.4	用量2で頭打ち
4	0.4	0.4	0.4	0.4	0.4	全用量同じ

シミュレーションの繰り返し回数: 100,000回

## ➤ 使用する推定量

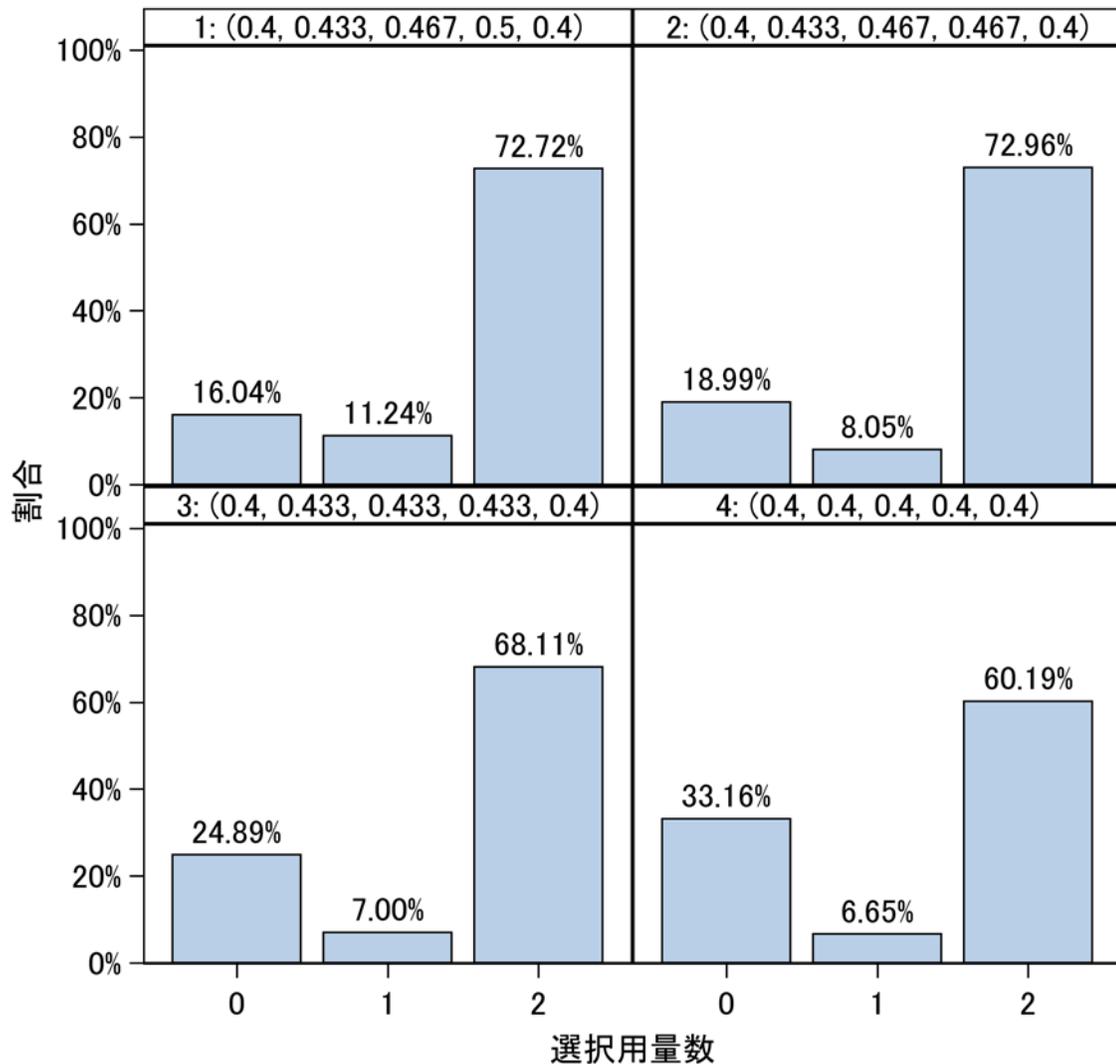
- MLEおよびBowden & Glimm (2008)のUMVCUE

## ➤ 評価指標

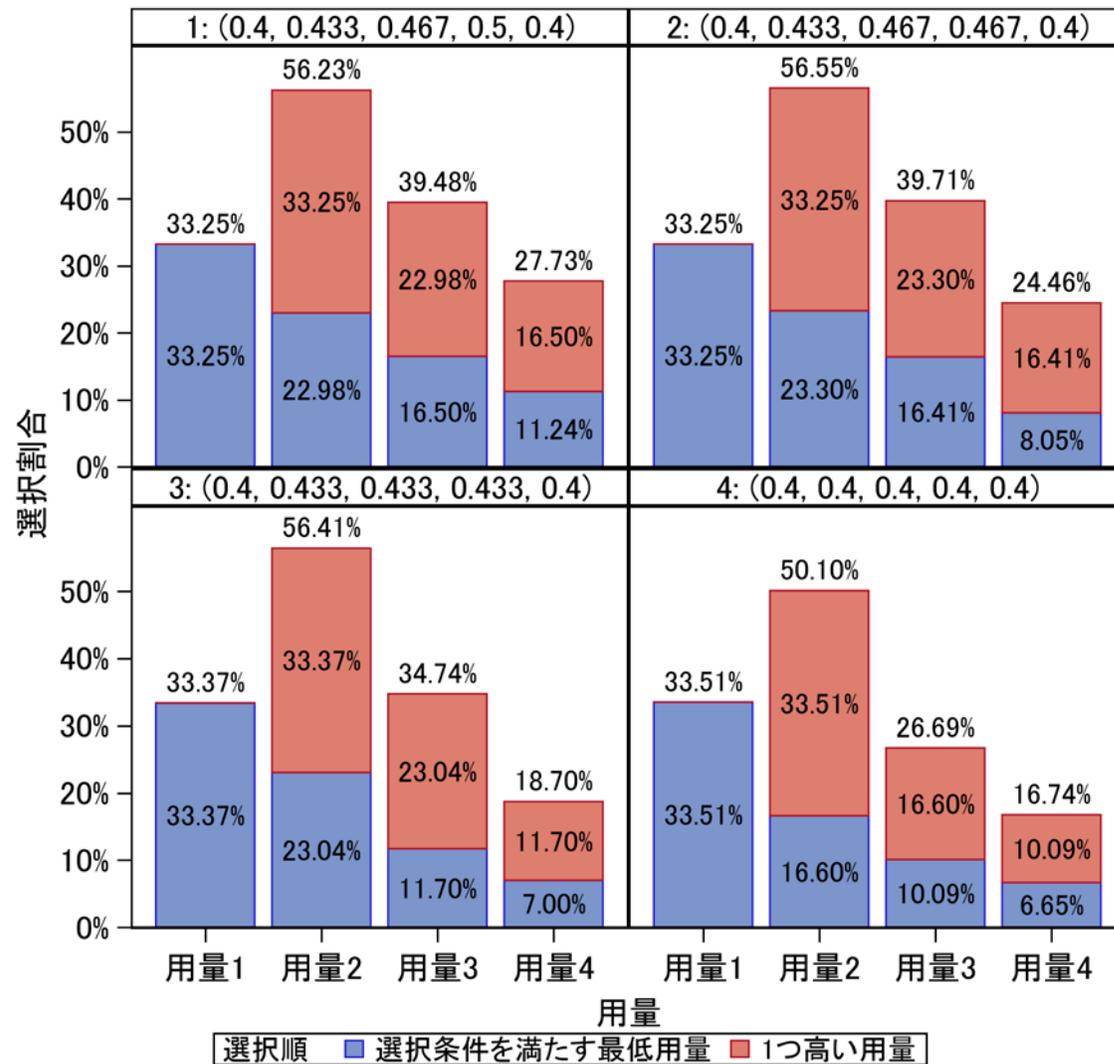
- 選択用量数および各用量の選択割合
- 推定値のバイアス・MSE・信頼区間の被覆確率

# 結果：選択用量数および各用量の選択割合

### 選択用量数

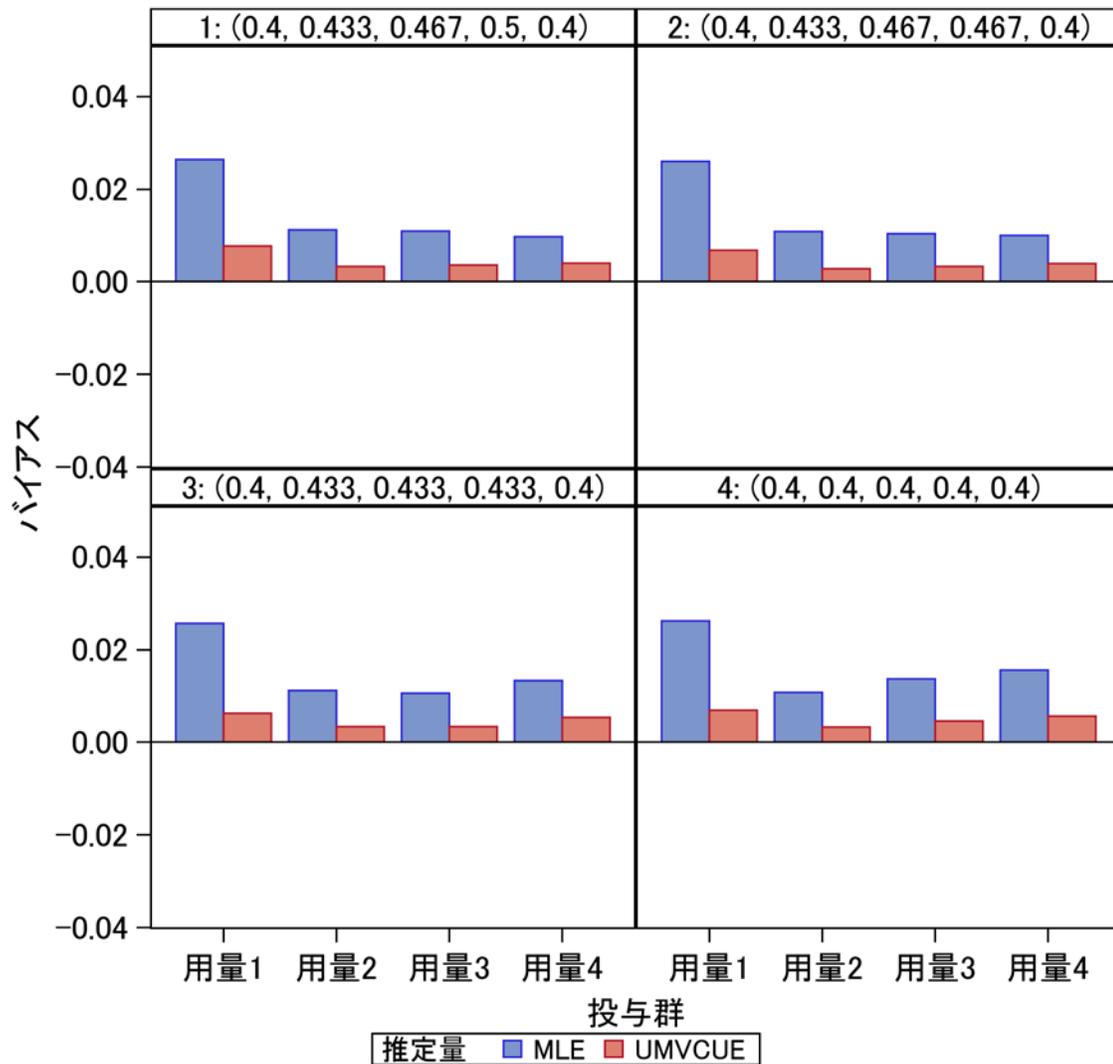


### 各用量の選択割合

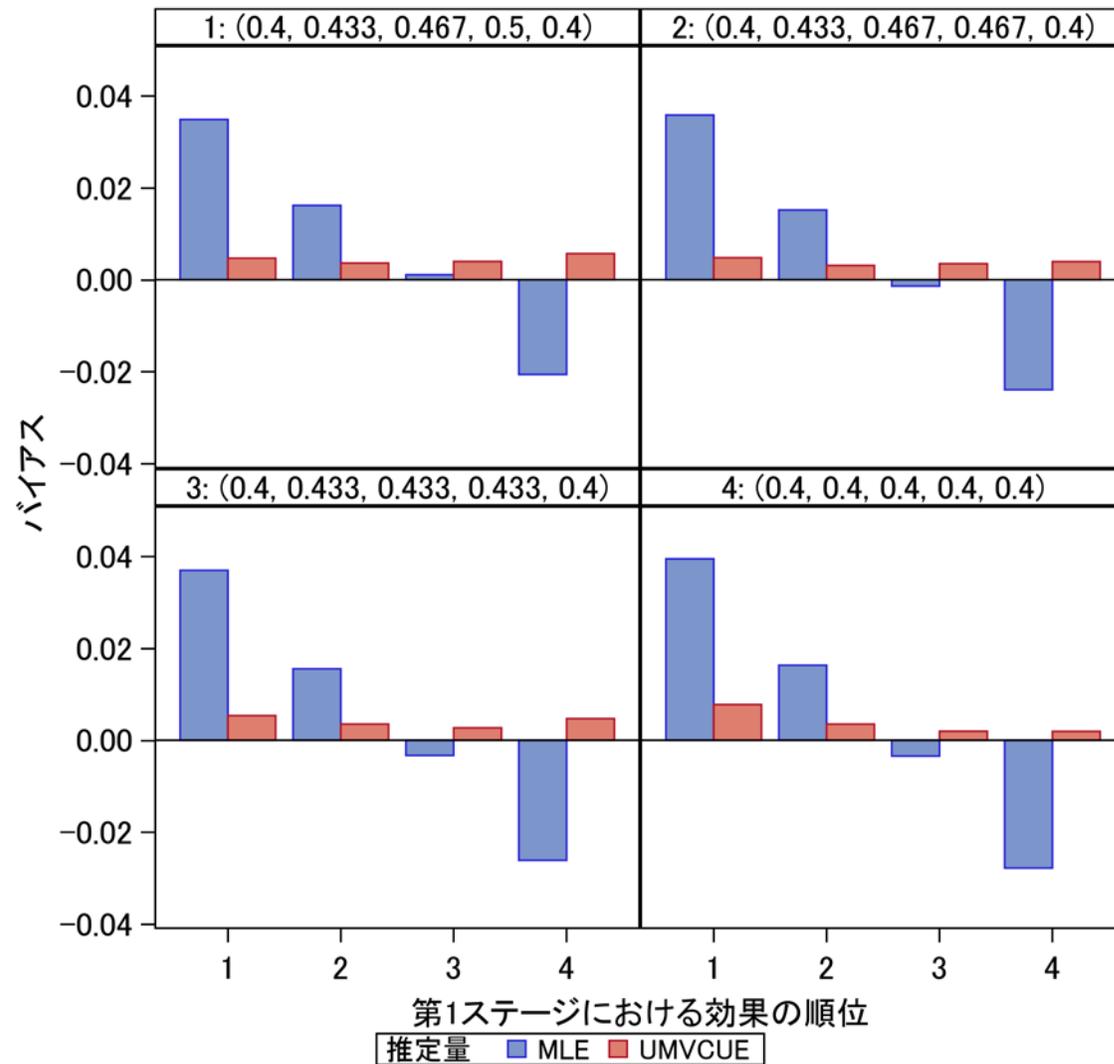


# 結果：各群の治療効果の推定(バイアスの評価)

バイアス(用量ごと)

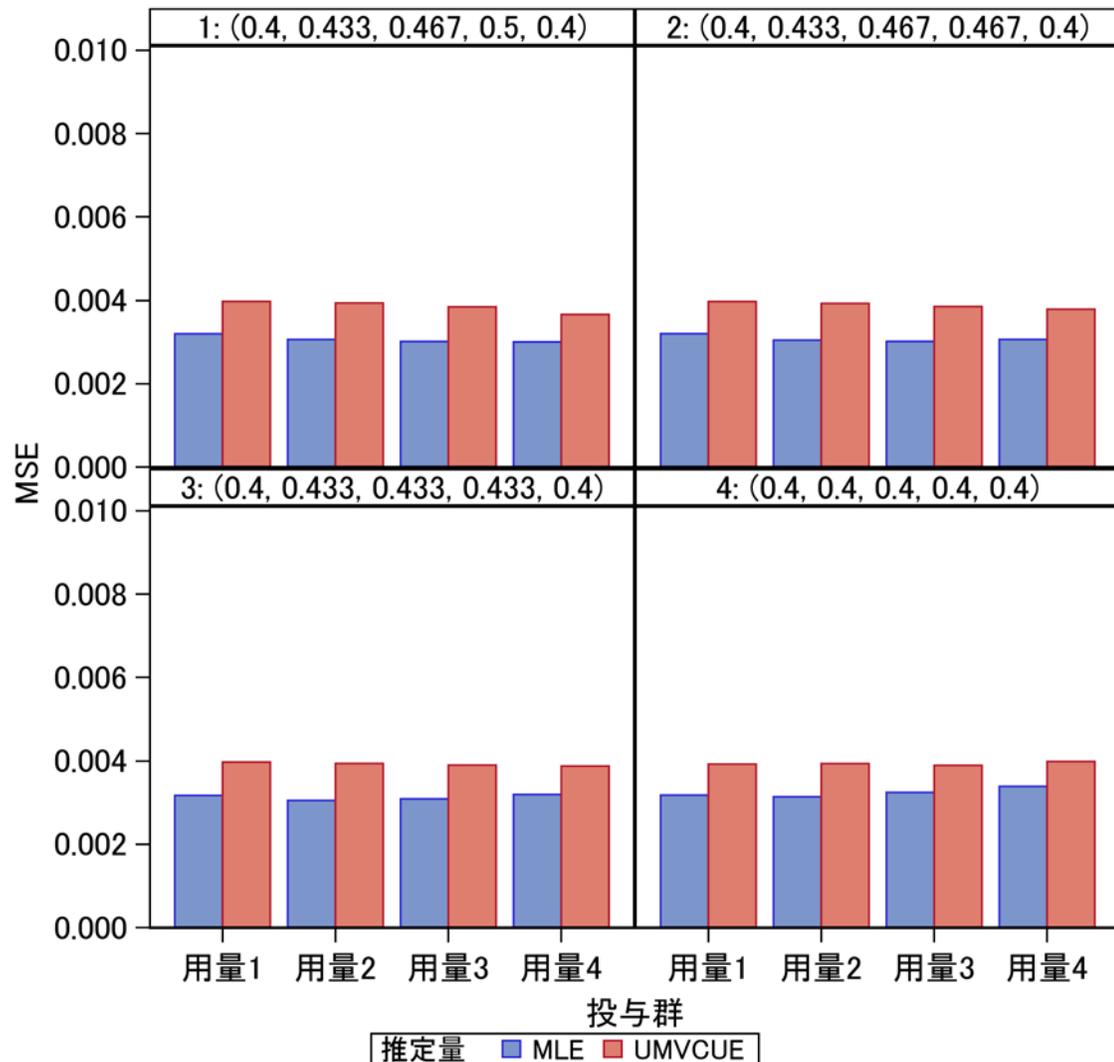


バイアス(第1ステージにおける効果の順位ごと)

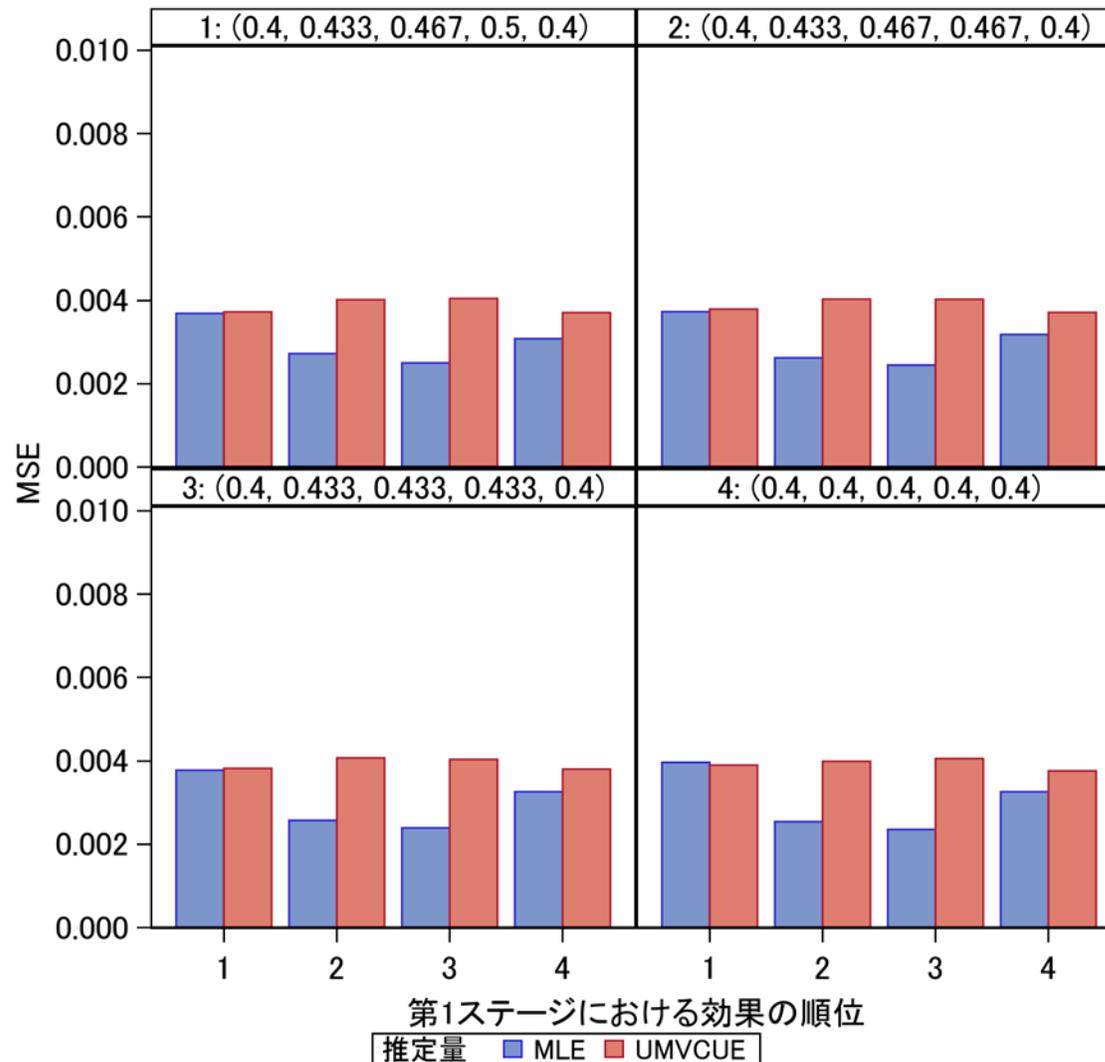


# 結果：各群の治療効果の推定(MSEの評価)

MSE(用量ごと)

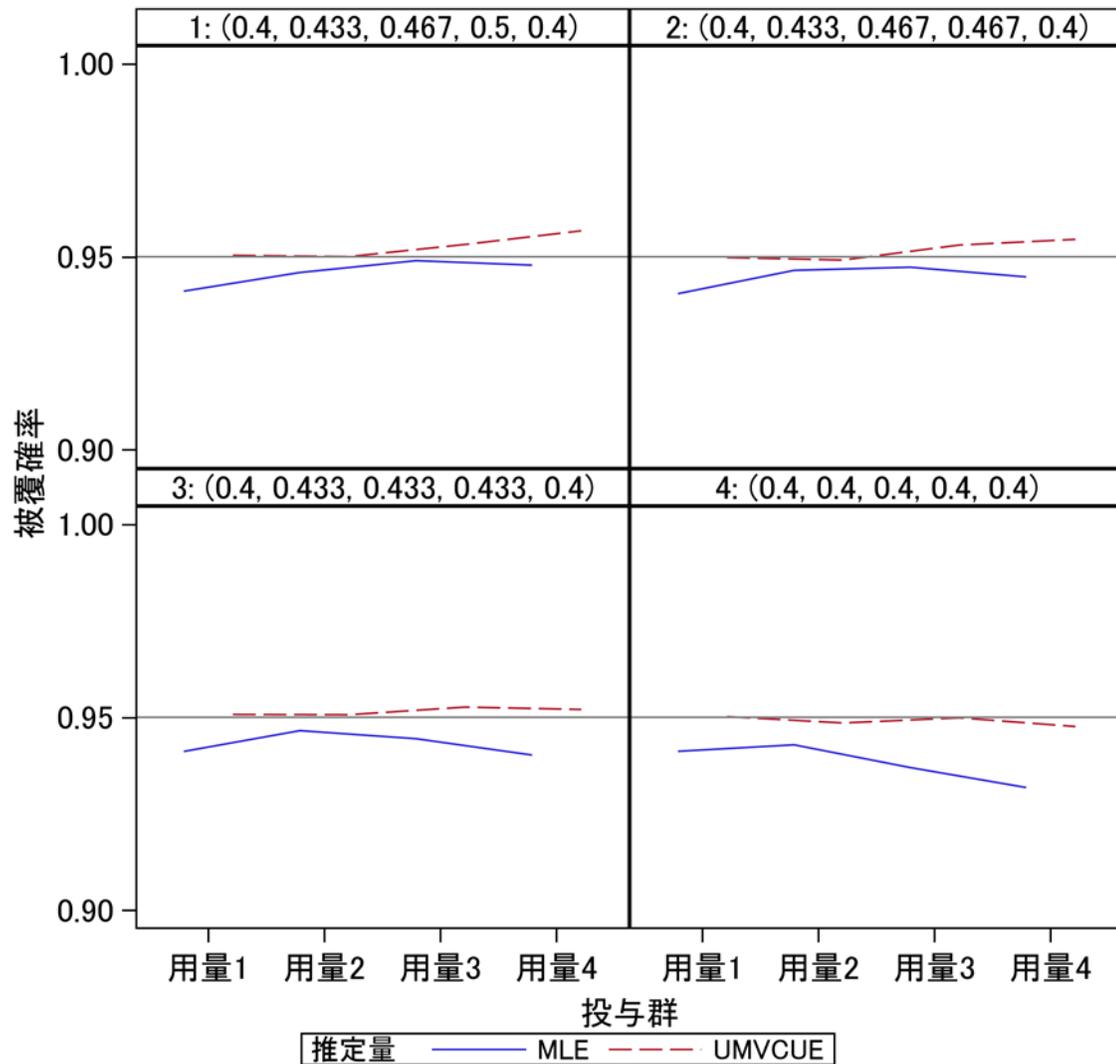


MSE(第1ステージにおける効果の順位ごと)

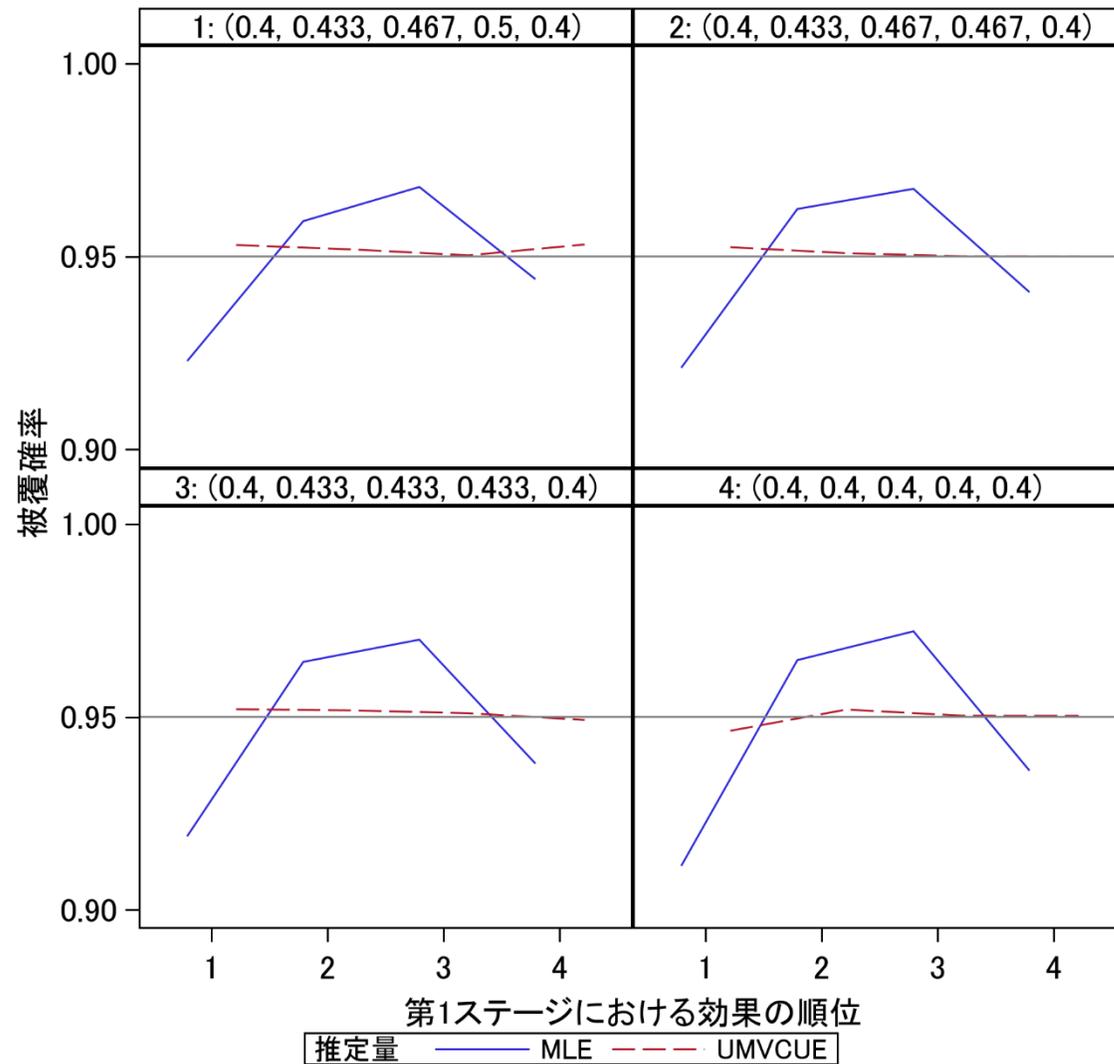


# 結果：各群の治療効果の推定(信頼区間の被覆確率の評価)

信頼区間の被覆確率(用量ごと)



信頼区間の被覆確率(第1ステージにおける効果の順位ごと)



# まとめ: MLEとUMVCUEの性質

## ➤ バイアス

- MLEと比べてUMVCUEはバイアスが補正された

## ➤ MSE

- ほとんどのケースでMLEと比べてUMVCUEの方がMSEが大きかった

## ➤ 被覆確率

- MLE: 一部で名義水準から大きく外れていた
- UMVCUE: おおむね名義水準付近を保っていた

## ➤ 参考: 中間解析時点の影響(※本シミュレーションの評価対象外)

- MLE: MSEは時点の影響を受けない
- UMVCUE: 中間解析時点が遅い(=第1ステージの情報量が多い)ほどMSE増大

# おわりに: 議論の参考として

- 推定値が持つべき性質 (Robertson et al. (2023a))
  - 使用するアダプティブデザインを適切に反映している
  - バイアスが無い、もしくは小さい
  - MSEが小さい(好ましいバイアス-バリエーション・トレードオフを反映している)
  - 簡易的に計算できる
  - 対応する信頼区間の算出方法が
    - 正しい被覆確率を持つ
    - (有効中止を含む) 仮説検定との一貫性/互換性がある
  
- バイアス-バリエーション・トレードオフについて (Robertson et al. (2023b))
  - どちらを重要視するかは試験の目的などによって異なる
    - 第II相: 後に続く検証治験のため精度の高い推定値が重要かもしれない
    - 第III相: 実社会での意思決定のためバイアスの無い推定値が重要かもしれない

- Food and Drug Administration. Adaptive Design Clinical Trials for Drugs and Biologics Guidance for Industry. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry> [accessed 2024/6/18]
- 日本製薬工業協会, アダプティブデザインに関するFDAガイダンスの邦訳, 2021年8月. [https://www.jpma.or.jp/information/evaluation/results/allotment/adaptive\\_design.html](https://www.jpma.or.jp/information/evaluation/results/allotment/adaptive_design.html) [accessed 2024/6/18]
- 日本製薬工業協会. アダプティブデザインの統計的推測に関する検討. 2024年4月. [https://www.jpma.or.jp/information/evaluation/results/allotment/DS\\_202302\\_adaptive.html](https://www.jpma.or.jp/information/evaluation/results/allotment/DS_202302_adaptive.html) [accessed 2024/6/18]
- 日本製薬工業協会. アダプティブデザインを計画するためのシミュレーションの検討. 2024年4月. [https://www.jpma.or.jp/information/evaluation/results/allotment/DS\\_202404\\_AD.html](https://www.jpma.or.jp/information/evaluation/results/allotment/DS_202404_AD.html) [accessed 2024/6/18]
- 高津正寛、他. 治療群の選択を伴うアダプティブデザインの動作特性の検討—事例に基づくシミュレーションの実践—. 2023年SASユーザー総会.
- Mayer C, Perevozskaya I, Leonov S, Dragalin V, Pritchett Y, Bedding A, Hartford A, Fardipour P, Cicconetti G. Simulation Practices for Adaptive Trial Designs in Drug and Device Development. *Statistics in Biopharmaceutical Research* 2019;11(4):325-335.
- Marconi, V. C., Ramanan, A. V., de Bono, S., Kartman, C. E., Krishnan, V., Liao, R., Piruzeli, M. L. B., Goldman, J. D., Alatorre-Alexander, J., de Cassia Pellegrini, R., Estrada, V., Som, M., Cardoso, A., Chakladar, S., Crowe, B., Reis, P., Zhang, X., Adams, D. H., Ely, E. W., & COV-BARRIER Study Group (2021). Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomised, double-blind, parallel-group, placebo-controlled phase 3 trial. *The Lancet. Respiratory medicine*, 9(12), 1407–1418.
- Cui L, Hung, HMJ, Wang SJ. Modification of sample-size in group sequential trials. *Biometrics* 1999;55:853-857.
- Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Statistics in Medicine* 2011;30(28):3267-3284
- O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;35:549-612
- Bauer P, Koenig F, Brannath W, Posch M. Selection and bias—two hostile brothers. *Statistics in Medicine* 2010;29(1):1-13.

- Bowden J, Glimm E. Unbiased estimation of selected treatment means in two-stage trials. *Biometrical Journal* 2008;50(4):515-527.
- Cohen A, Sackrowitz HB. Two stage conditionally unbiased estimators of the selected mean. *Statistics & Probability Letters* 1989;8(3):273-278.
- Lawrence D, Bretz F, Pocock S. INHANCE: An adaptive confirmatory study with dose selection at interim. In *Indacaterol - The First Once-Daily Long-Acting Beta2 Agonist for COPD*, Trifilieff A (ed.) Springer: Basel, 2014; 77–93.
- Robertson DS, Choodari-Oskoei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs I: A methodological review. *Stat Med.* 2023a;42(2):122-145.
- Robertson DS, Choodari-Oskoei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs II: Practical considerations and guidance. *Stat Med.* 2023b;42(14):2496-2520.

# 当日議論結果

SSR/TS	バイアス調整された推定値を用いた場合	バイアス未調整の推定値を用いた場合
<p>利点 Pros</p>	<ul style="list-style-type: none"> <li>- 真値に近い値を患者さんへ伝えられる</li> <li>- バイアスの許容範囲を事前に検討できる、正当化できる事によって頑健な結果を与えることができる、意味のある最小差を超えない事を確認できる</li> </ul>	<ul style="list-style-type: none"> <li>- 事例のようなバイアスが小さい場合には絶対にやらなければいけないという事ではない？</li> <li>- 費用、リソースを抑えられる(追加の検討をしなくてよい)</li> <li>- 解釈がしやすい(検討した上で調整しないという結論づけたうえで)</li> </ul>
<p>欠点 Cons</p>	<ul style="list-style-type: none"> <li>- 計算してみたらうまくいかないケースもありえる(リスクとして考える): どういったときにうまくいくかという点が不明確</li> <li>- どこまでシミュレーションをして規制当局と合意を取るのかが不明確</li> <li>- 検定と推定の齟齬についての事前規定、合意が難しい、解釈が難しい⇒決められれば、理想的 ※感度解析として実施した場合は齟齬は発生しない</li> </ul>	<ul style="list-style-type: none"> <li>- 事前検討していなかった場合、照会事項対応などで聞かれたときに審査の時間がかかり、承認時期に影響を与える可能性がある</li> <li>- 科学的妥当性を求められたときに答えられない</li> <li>- 参考情報として利用する際に間違ったバイアスのかかった情報によって新たな試験が失敗してしまう</li> <li>- メインの結果がだけが出回ってしまってしまう</li> </ul>

## アダプティブデザインを用いた試験結果に対するバイアス調整についての 利点/欠点を踏まえてのグループAの提言

- バイアスの調整をすべきかどうかは結論は出せない（ケースバイケース）
  - ハードルが高い事は認識しておきたい
- シミュレーションで検討して決定する
  - シミュレーション（手法、シナリオ、パラメータ）を行う範囲を決める事が現時点では困難
  - バイアス調整できる手法から入る事は妥当ではない⇒全部できて、バイアス調整もできる事が理想
  - まずはADを選択できるかどうかを検討した上でバイアス調整まで検討したい
- やらなくても社内でバイアスの程度を確認しておくべき
- バイアスの評価について提言することはあっても良い
  - 検討結果によって感度解析を検討し、合意することができる
- 場合によっては感度解析をやるかどうかを決めておく
- 指標の優先順位
  - 選択バイアスなども含めたアルファエラーの調整となるのでアルファエラーを優先すべきか？
  - 検証試験の準備のために実施するのであれば、バイアス調整を優先？⇒目的に応じて検討
  - 探索的な層でADを選択することで検証試験に役に立つ可能性がある
  - MSEとバイアスをどちらを優先するかは要検討
- 経験値は積んでいきたいが、コストやリソースなども考えるとまだハードルが高い
  - 今後の動きに応じてニーズが発生する可能性がある

# グループB

SSR(症例数再設計)/TS(治療群選択)についてのディスカッション  
Pivotal試験での想定

SSR/TS	バイアス調整された推定値を用いた場合	バイアス未調整の推定値を用いた場合
利点 Pros	①バイアスが少なくなる	①中間解析に影響されない ②解釈がしやすい ③MSEが小さくなる
欠点 Cons	①中間解析に影響される ②計算や事前準備が複雑 事前準備 シミュレーションでの検討内容 手法を選択する基準 ③検定と推定で結果が異なった場合, 直感的な一貫性がとれず説明が難しい	①真の評価が大きいのに過小評価してしまう可能性がある(企業にとってのリスク) ②バイアスが大きくなる

## 提言

事前にバイアス調整済みかバイアス未調整かを, 主解析にするか補足的解析にするかをシミュレーションを用いて検討して決めておく

# グループC



SSR/TS	バイアス調整された推定値を用いた場合	バイアス未調整の推定値を用いた場合
利点 Pros	<ul style="list-style-type: none"><li>バイアスが補正されてリスクベネフィットの評価に使える</li><li>次の試験に対して良い推定値を与えることができる</li></ul>	<ul style="list-style-type: none"><li>実装, 解釈が容易</li></ul>
欠点 Cons	<ul style="list-style-type: none"><li>実装, 検討に時間を要する</li><li>P値とCIは整合するのか? 整合しない場合に主解析として用いる事が難しくなるのでは?</li><li>一般的には未調整の推定値を示しているケースが多いので, 調整した推定値を示すことはデメリットになるのでは?</li><li>デザインによってはバイアス補正の方法がない場合がある。</li><li>場合によってはコンサバな結果となる可能性がある</li><li>バラつきが大きくなる</li></ul>	<ul style="list-style-type: none"><li>バイアスが含まれているので, 添文への記載, リスクベネフィットの評価が難しい</li></ul>

## アダプティブデザインを用いた試験結果に対するバイアス調整についての 利点/欠点を踏まえてのグループCの提言

- 立案時にはどちらを主解析に用いるかをPMDAと合意しておく必要があるのではないか？
- バイアスの検討は社内ディスカッションではあまり議論にならない
- 領域によって許容できるバイアスの程度を事前に検討しておくことが重要、ただし、基準を示すのは難しい。各シナリオが発生する確率も考慮する必要がある（非常に低い確率で生じるバイアスを調整する必要性があるか？）
- 使う推定量の解釈を事前に考えておかないといけない
- 例えば今回のシミュレーション結果であれば、調整しない推定値が主解析でも良いのではないか？
- 調整する／しない、PMDAとの議論を想定し、事前に両方検討しておいた方が良い