



Title	臨床試験の解析において重み付き推定を用いて複数の外部情報源を統合するための情報借用法に関する研究
Author(s)	高橋, 圭太
Description	配架番号 : 2993
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第16900号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16900
Doc URL	https://hdl.handle.net/2115/99745
Type	doctoral thesis
File Information	TAKAHASHI_Keita.pdf, 全文



学位論文

臨床試験の解析において重み付き推定を用いて複数の外部
情報源を統合するための情報借用法に関する研究
(Studies on an adaptive constrained weighted estimation for
incorporating multiple external information sources
in clinical trials)

2026年3月

北海道大学

高橋 圭太

Keita Takahashi

学位論文

臨床試験の解析において重み付き推定を用いて複数の外部
情報源を統合するための情報借用法に関する研究
(Studies on an adaptive constrained weighted estimation for
incorporating multiple external information sources
in clinical trials)

2026年3月

北海道大学

高橋 圭太

Keita Takahashi

目次

発表論文目録および学会発表目録.....	1 頁
要旨	2 頁
略語表.....	5 頁
緒言	6 頁
方法	13 頁
結果	25 頁
考察	40 頁
結論	43 頁
謝辞	45 頁
利益相反.....	46 頁
引用文献.....	47 頁
付録	52 頁

発表論文目録および学会発表目録

本研究の一部は以下の論文に投稿した.

1. Keita Takahashi, Kazufumi Okada, Shiro Tanaka, Isao Yokota
Adaptive Constrained Weighted Estimation for Incorporating Multiple External
Information Sources
Pharmaceutical Statistics, 2026, 25(1):e70059.

要旨

【背景と目的】

希少な疾患を対象とする臨床試験では、十分な数の対象者を登録することが難しいという観点から、ヒストリカルデータなどの外部情報を利用して情報量不足を補う（情報を借用する）選択肢が考慮される場合がある。特に近年は、外部情報を同時対照群の一部として用いるハイブリッド対照デザインが注目されており、また同デザインを採用した試験事例も報告されている。適切な外部情報を利用できれば、同時対照群に割り付ける対象者数を削減することで、試験の実施可能性および効率性を高められる可能性がある。

同時対照群と外部情報の同質性を判断するためには、原則としては情報源間で治療内容、対象集団、評価方法、施設、時期などが一致しているかを定性的に評価することが必要である。一方、現実的な応用では同質性の有無を二元論的に評価することは容易ではなく、また同質性があることを仮定できたとしてもそれが誤りである可能性は否定できない。

これらのことから、同時対照と外部情報源の間の同質性をデータによって評価し、その程度に基づいて借用量を適応させる定量的アプローチが注目されてきた。このアプローチは、情報借用の強みを活かしながらリスクを抑えるという観点から自然な発想に基づいているが、第一種の過誤確率および検出力を明示的に制御することができない。ほとんどの既存法では、手法固有のモデルパラメータをシミュレーションに基づいて調整することで動作特性を制御できるが、試験計画の観点では、重要な動作特性は解析手法の設計によって制御できることが望ましい。

また、情報借用に関する既存法の多くはベイズ法に基づいているが、第一種の過誤確率や検出力の明示的な制御という観点では頻度論的手法の方が整合的である。臨床試験において複数の外部情報源を適応的に利用することを念頭に置いて開発された頻度論的手法は、筆者の知る限り報告されていない。

以上を踏まえ、本研究では、第一種の過誤確率および検出力の明示的な制御と適応的借用を両立する方法を開発することを目的とする。特に、既存法では実現できていない、複数の外部情報源の頻度論的借用に焦点を当てる。

【方法】

本研究では、前述の課題に対処するために、対照群の有効サンプルサイズを一定に保つ制約条件を重み付き最尤推定に導入し、適応的な情報借用を頻度論に基づく制約付き最適化問題として設計する。具体的には、各外部情報源に個別の

重みパラメータを設定し、対照群の有効サンプルサイズに関する制約のもとで重み付き対数尤度を最大化することで重みを推定する。制約付ける有効サンプルサイズは標準的なサンプルサイズ設計に基づく目標値、すなわち対照群に対する所望のサンプルサイズとする。これによって、すべての外部情報源が同時対照と同質である場合には第一種の過誤確率と検出力が名目水準に制御され、そうでない場合には適応的な重み付けを通じて第一種の過誤確率の増加を抑制しつつ、検出力を上昇させることを意図している。

数値例によって提案法の各外部情報源への重み付けの性質を探索する。モンテカルロシミュレーションによって第一種の過誤確率や検出力などの動作特性を評価する。実際の臨床試験データへの適用によって、提案法の現実的な挙動を確認する。上記のすべての検討は代表的な既存法と比較させたうえで行い、提案法の統計的性質を様々な角度から検討する。

【結果】

数値例によって、提案法は同時対照と外部情報源の乖離に応じて重みを適応的に与える性質が確認された。乖離が著しく大きい外部情報源に対する重みは0となり、すべての外部情報源の標本平均が同値である場合には等しい重みを与える性質も確認された。同時対照と外部情報源の乖離だけでなく、サンプルサイズも重みにわずかに影響することがわかった。これは制約条件のもとで、サイズが小さい外部情報源により大きな重みを与える方が尤度の観点で有利であるためである。ただし、重み付きサンプルサイズのスケールではその影響は非常に小さいことも確認された。

モンテカルロシミュレーションによって、すべての外部情報源が同時対照と同質である場合には、提案法は第一種の過誤確率と検出力を名目水準に制御することが確認された。この性質は外部情報源の数およびサンプルサイズに依存しなかった。異質な外部情報源の存在下では、同時対照と類似した情報源により大きな重みを与える傾向が確認された。同時対照と外部情報源の乖離が大きいとも小さいとも言えない曖昧な状況では、提案法は既存法よりも検出力が高く、第一種の過誤確率が低い傾向があった。一部の外部情報源のみ同時対照と整合するシナリオでは、適切な情報源の特定に関して、提案法は一部の既存法に劣る場合があった。これは均一な重み付けが重み付き対数尤度の上昇に寄与する一つの要素だからである。

臨床試験データへの適用では、外部情報源の多くが同じ方向に偏っている状況では、重み付き推定量がその方向へ縮小しやすく、反対方向に位置する少数の外部情報源への重み付けが抑制される傾向がみられた。また、提案法の重み付けの挙動が既存法よりも安定していることも確認された。

【考察】

既存の適応的情報借用法は、対照情報源同士の同質性に基づいて借用量を適応させるアプローチを採用しているため、第一種の過誤確率と検出力を明示的に制御することができない。これは試験計画の観点では望ましくないため、これらの重要な計画事項をデザインによって保証しながら、適応的借用の利点（すなわち情報源間の同質性の解析への反映）を引き出すためにはどうしたらよいか、という問いが本研究の起点となった。本研究では、有効サンプルサイズを一定水準に制約付けるという、既存法とは逆の発想によってこの問いに応えた。

有効サンプルサイズを制約するという設計は、すべての外部情報源が同時対照と異なる場合にも借用することを意味し、これは望ましい性質ではない。一方で、同時対照と整合しない外部情報源に対する重みが 0 になるとは限らないという問題は既存法でも見られ、適応的な方法でこれを完全に防ぐことは難しいと考えられる。例えば既存法では、同時対照との乖離が著しく大きい外部情報源にも大きな重みが与えられることがあった。現実的に重要なことは、各手法が様々な種類の異質性のもとでどのような挙動を示すかを事前に理解していることである。本研究では主に 2 つの外部情報源を考慮したが、より多くの異質な外部情報源のもとでの提案法の性能評価は今後の研究課題である。

【結論】

本研究では、臨床試験において複数の外部情報源を利用する新たな適応的情報借用法を開発した。この方法は、適応的な重み付き最尤推定に有効サンプルサイズを一定に保つ制約条件を導入することで、頻度論の枠組みのもとで、すべての外部情報源が同時対照と同質である場合の第一種の過誤確率および検出力の名目水準への制御と適応的な借用を両立する。提案法は標準的な臨床試験の計画事項（有意水準、検出力、サンプルサイズ）に基づいて設計したものであるため、ハイブリッド対照試験を標準的な臨床試験の自然な拡張として位置付けることに寄与するかもしれない。これによって、情報借用の適用対象となりやすい希少な疾患領域を対象とした研究の発展につながることが期待される。

略語表

本文中および図中で使用した略語は以下のとおりである.

ACWE	adaptive constrained weighted estimation
CDAI	Crohn's disease activity index
CI	confidence / credible interval
CPP	conditional power prior
EBPP	empirical Bayes power prior
ESS	effective sample size
FDA	Food and Drug Administration
ICH	International Council for Harmonisation
JPP	joint power prior
MAC	meta-analytic combined
MAP	meta-analytic predictive
MCID	minimal clinically important difference
MPP	modified power prior
RCT	randomized controlled trial
R-MAP	robust meta-analytic predictive
RMSE	root mean squared error
std	standardized difference

緒言

1. ランダム化比較試験の対照群としての外部情報の利用

ランダム化比較試験 (randomized controlled trial; 以下, RCT) は, 一次研究としては最もエビデンスレベルの高い研究デザインであるとされる (Burns et al., 2011; Guyatt et al., 2008). 医薬品規制調和国際会議 (ICH; International Council for Harmonisation 〈旧称: International Conference on Harmonisation〉) によるガイドラインでは, 治療効果推定のバイアスを回避するための最も重要な要素のひとつとしてランダム化に言及されており, 医薬品の製造販売承認を目的とした臨床試験において RCT を採用することが推奨されている (ICH, 1998).

一方, 希少疾患領域においては, 臨床試験に参加可能な対象者数が限られていることから, 一つの試験内で対照群を設定する RCT の実施が困難となることがある (日本製薬工業協会, 2022). また製薬企業の立場からみると, 市場規模が小さい希少疾患領域は収益性の観点から開発が躊躇されやすいという問題も存在する (日本製薬工業協会, 2022). 仮に RCT を実施できた場合でも対象者登録が難しいという問題は変わらないため, 試験の長期化によって有効な治療薬の上市が遅れることも懸念される. さらに有望な標準治療が存在しない疾患領域では, 一部の患者を対照群に割り付けることが個人倫理の観点から批判されることもある (Yin, 2014).

ICH と米国食品医薬品局 (FDA; Food and Drug Administration) は, 希少疾患を対象とした臨床試験における選択肢として外部対照試験, すなわち試験外の情報を対照群として用いるデザインについて言及している (FDA, 2023; ICH, 2000). 外部対照として用い得る情報源としては, 関連する過去試験のデータや患者レジストリデータ, また文献で公表された要約統計量などが挙げられる (Ghadessi et al., 2020; 野村他, 2022). 外部対照試験は, 試験における必要対象者数の減少が図れる一方, ランダム化を実施しないため, 治療群間の系統的な差異を排除できないことが問題となり得る (Burger et al., 2021; FDA, 2023).

近年, RCT と外部対照試験の利点を併せ持つ第 3 の選択肢として, 同時対照群の一部を外部情報で補うハイブリッド対照試験が注目されている (Burger et al., 2021; Pang and Zhu, 2024; Pocock, 1976; Viele et al., 2014; 日本製薬工業協会, 2022; 野村他, 2022). このデザインは外部対照試験と同様に, 必要対象者数の減少による効率化が図れ, また対照群への割付数の削減によって前述した個人倫理の問題に部分的ではあるが対処することができる. さらに緒言 4 節で述べるとおり, 同時対照と外部対照の同質性をデータに基づいて評価することで, 必要に応じて外部対照の解析への寄与度を制御できることも利点であるといえる.

2. ハイブリッド対照試験の事例

同時対照群を外部情報で補強することを計画したいくつかの臨床試験が報告されている。Hueber et al. (2012)は、中等度から重度のクローン病患者に対して secukinumab とプラセボを比較する RCT を実施した。有効性の主要評価項目はベースラインから 6 週間後のクローン病活動性指数 (Crohn's disease activity index ; 以下, CDAI) の変化量 (Δ CDAI) であった。試験計画では、各群で 40 人と 20 人を登録し、さらに 6 つの過去試験 (Hanauer et al., 2006; Sandborn et al., 2001a; Sandborn et al., 2001b; Sandborn et al., 2007a; Sandborn et al., 2007b; Winter et al., 2004) のデータから 20 人分の情報を外部対照として用いることが意図されていた。当該試験の結果としては、secukinumab はプラセボよりも CDAI を減少させず、クローン病に対する有効性は示されなかった。

Baeten et al. (2013)は、強直性脊椎炎患者に対して secukinumab とプラセボを比較する RCT を実施した。有効性の主要評価項目は 6 週目の奏効割合であった。試験計画では、各群で 24 人と 6 人を登録し、さらに 8 つの過去試験のデータから 43 人分の情報を外部対照として用いることが意図されていた。6 週目奏効割合の事後平均は secukinumab 群で 59.2%, プラセボ群で 24.5%と推定され、リスク差が 0%以上となる事後確率が 99.8%となった。この事後確率が事前に定めた閾値 95%を上回ったことに基づき、secukinumab の有効性が示されたと結論づけられている。

以上の 2 つの試験はいずれも複数の外部情報源を利用しており、また解析手法には後述するベイズ流メタアナリシス的アプローチが用いられたという点で共通している。

ハイブリッド対照デザインを採用した試験は他にも報告されている (Bodey et al., 1990; Holmes et al., 2006; Salloway et al., 2021; Yamaue et al., 2017)。これらの事例では、外部情報の利用によって治療の有効性を証明できたとは限らないが、実施が困難であった状況下で試験を行えたこと自体に一定の意義がある。

3. 情報借用の正当性

前節まで、臨床試験の対照群の一部として外部情報を利用するモチベーションとその事例について述べた。本節では外部情報を用いることのリスクについて述べ、またその正当性を評価する方法を説明する。

情報借用の最大のリスクは、同時対照と外部情報の母集団の同質性が担保されない場合に、治療効果の推定にバイアスが生じることである。標準的な RCT では、すべての対象者が試験内でランダム化されるため、比較群間で患者背景は同質であることが期待でき、したがってアウトカムの群間差を治療効果の不偏

推定量とみなすことができる。情報借用の文脈では、真のパラメータが同時対照と外部対照で異なる場合、対照群の推定量にバイアスが生じ、したがって治療効果の推定量にもバイアスが生じる。これに伴い、検定仮説の方向によって第一種の過誤確率の上昇または検出力の低下も生じる。このため、対照情報源の間で真のパラメータが一致していることが必要である。しかし実際には真のパラメータを知ることはできないので、各情報源における患者背景や施設差などを総合的に考慮したうえで借用の正当性を定性的に仮定することが通常である。

この正当性を定性的に評価するための古典的な指針として、Pocock (1976) による基準 (Pocock 基準) が存在する。これは元々 RCT に過去試験データを統合するデザインを念頭に置いた基準であったが、現在では必ずしも過去試験データに限らず、外部情報全般の利用の正当性を評価する基準としてしばしば引用される (Jahanshahi et al., 2021; Pang and Zhu, 2024)。Pocock 基準は次の通りである。

- | |
|--|
| <ol style="list-style-type: none">1. 詳細に定義された、新規試験の対照治療群と同じ治療を受けている2. 新規試験と同じ適格基準を用いた最近実施された臨床試験である3. 治療の評価方法が新規試験と同じである4. 重要な患者背景の分布が新規試験と同様である5. 新規試験とほぼ同じ組織で実施されている6. 新規試験との結果の違いを生むと予想される他の要因が存在しない |
|--|

以上の 6 つの項目が提唱されているが、結局のところいずれも、「その外部情報を対照群として用いることで、治療効果の推定にバイアスは生じないのか」という問いに基づいている。情報借用の文脈に関わらず、標準的な臨床試験においても、治療効果をバイアスなく推定できることを事前に主張するためには種々の定性的な仮定が必要である。例えば多施設共同試験では、施設間の系統的な差異がかなり小さいことを事前に仮定できている必要がある。Pocock 基準はこのような仮定を外部対照を用いる状況に一般化したものと解釈できる。

項目 1 と 3 は、同時対照群と外部対照群の間で同じアウトカムを同じ方法で測定するという意味で、臨床試験に限らず科学として最も基本的な項目であるといえる。項目 1 と 3 以外は各対照情報源における母集団の同質性に関連する項目である。項目 2 と 5 は外部情報の測定時期や場所に関するものであり、例えば過去試験データや海外の試験データを外部情報として用いるような場合には厳密には成立しないかもしれない。

ここでは前節で紹介した Hueber 試験を例に挙げ、同試験における情報借用の妥当性を Pocock 基準に照らし合わせて検討する。まず、6 つの過去試験の対照

群ではすべて同時対照と同様にプラセボが用いられたため、項目 1 は満たされている。また、いずれの過去試験でも CDAI のベースラインからの変化量が測定されていたものの、その評価時期は試験ごとに多少異なっていた（4 週目または 8 週目）ことから、項目 3 は部分的にしか満たされていない。また、いずれの過去試験も CDAI が 220 から 450 の範囲にある中等度から重度のクローン病患者を対象としたため、重要な適格基準は同一であった。一方、年齢やクローン病の病歴などの適格基準は完全に同一ではないことに加え、試験の実施時期にも数年程度の差異があるため、項目 2 も部分的にしか満たされていない。また実施施設は過去試験間で異なるため項目 5 は満たされておらず、項目 4 と 6（重要な患者背景やその他の要因）については、前述のような試験の実施時期や施設、また適格/除外基準の微妙な差異を考慮すると、完全に満たされていると仮定することは難しい。

Pocock 基準は、情報借用の正当性を定性的に検討するうえで重要な視点を提供するものであるが、現実的にはすべての基準を満たした外部情報を利用できる状況は少なく、むしろ Hueber 試験のように、「部分的には満たされている」という状況がほとんどであると考えられる。また、ハイブリッド対照試験を念頭に置いた場合、Pocock 基準を「厳しい」と評価する研究者もいる（Li et al., 2023; Lim et al., 2020）。

以上のように、Pocock 基準のような定性的基準を満たしているかそうでないかの二元論的な評価は容易ではなく、また基準を満たすことが仮定できたとしてもそれが誤りであるリスクが存在する。このことから、対照情報源の間の同質性をデータによって評価し、その程度に基づいて借用量を適応させる定量的アプローチが注目されている。

4. 適応的情報借用

適応的情報借用法は、情報源の間の同質性をデータに基づいて評価し、それを借用量に反映させるアプローチである。この「同質性の評価」と「借用量への反映」をどのように実現するかに関して様々なアプローチが提案されている。

実際の臨床試験でも利用実績がある手法として、メタアナリシス的アプローチが挙げられる（Baeten et al., 2013; Hueber et al., 2012; Neuenschwander et al., 2010; Schmidli et al., 2014; Viele et al., 2014; 武田他, 2015）。このアプローチは階層ベイズモデルに基づいており、各対照群の推定対象パラメータが同じ分布に由来することを仮定し、その分散パラメータをデータから推定することを通じて、外部情報の解析への寄与が適応的に調整される。この考えのもとで、さらに meta-analytic predictive (MAP) 法および meta-analytic combined (MAC) 法と呼ばれるアプローチに大別することができる。MAP 法は、試験の計画段階に外部情報源

から導いたメタアナリシ스의予測事前分布（MAP 事前分布）を、試験終了後にベイズの定理によって同時対照データと組み合わせて推論する二段階アプローチである。MAC 法は、すべての対照情報源を同時に用いてパラメータを推定する一段階アプローチである。両者はモデル式のもとでは等価であるが、二段階アプローチである MAP 法は、外部情報から導いた事前分布を試験の実施前に示すことができるため計画時に利用しやすいという特長がある（野村他, 2022）。また、対照情報源の間の異質性にさらに対処するため、MAP 事前分布を曖昧な分布との混合によってロバスト化する robust-MAP（R-MAP）法が提案された（Schmidli et al., 2014）。この方法では、興味のあるパラメータの推定に用いる事前分布を、外部情報に基づく MAP 成分と外部情報に基づかないロバスト成分の両方によって構成することができる。これらのメタアナリシス的アプローチでは、推定対象パラメータの分散が推定精度に直結するため、その分散の事前分布のパラメータが借用の程度を調整し得るチューニングパラメータとなる。また、これらのアプローチは外部情報源の数に関わらず利用可能であるが、メタアナリシスの考えに基づいていることから、複数の外部情報源が利用可能な状況が想定されることが多い（武田他, 2015）。

また、0 から 1 の値を取り得る power parameter を用いて外部対照の尤度を割り引く power prior 法とその派生手法が提案されている。Ibrahim and Chen (2000) は、power parameter を定数として対数尤度関数を割り引く conditional power prior (CPP) と、power parameter を確率変数として興味のあるパラメータと同時に推測する方法である joint power prior (JPP) を提案した。その後、JPP に対して、外部情報に基づく尤度関数を定数倍すると power prior が異なる分布になるという問題、すなわち尤度原理への違反が指摘された（Duan et al., 2006; Neuenschwander et al., 2009）。同著者らによって、この問題を解決するために正規化定数が補正された手法である modified power prior (MPP) が提案された。Gravestock and Held (2017; 2019) は、power parameter の周辺尤度の最大化に基づく empirical Bayes power prior (EBPP) を提案した。同手法は JPP や MPP のように power parameter を確率変数として扱うのではなく、未知定数としてデータから最適化するアプローチである。そのため同手法では power parameter に事前分布を指定しない。

Power prior ベースの各手法は、メタアナリシス的アプローチと比較すると単一の外部情報源の利用が想定されることが多い（武田他, 2015）ものの、複数の外部情報源を利用することも可能である。実際、JPP と MPP は原著論文において外部情報源が複数ある状況も考慮されており、MPP については複数の外部情報源の利用により焦点を当てた派生手法も提案されている（Banbeta et al., 2019; 2022）。同様に EBPP も複数の外部情報源の利用を想定して拡張されている

(Gravestock and Held, 2019). Power prior ベースの手法では、各外部情報源に対する power parameter の事前分布のパラメータが借用の程度を制御する役割をもつ。ただし、EBPP は power parameter に事前分布を指定しないため借用性に関するチューニングパラメータも有さない。

以上の手法はベイズ法に基づいているが、頻度論に基づく手法も提案されている。例えば、仮説検定に基づいて同時対照と外部情報源の類似性を評価し、その結果に応じて外部情報の利用有無を決定する test-then-pool 法が挙げられる (Li et al., 2020; Okada et al., 2024; Viele et al., 2014)。この手法は非専門家の理解を得やすいと考えられるが、外部情報をすべて利用するか否かの二元論的アプローチであり、外部情報の寄与割合を細かく調整することはできない。他には、adaptive lasso に基づく方法、傾向スコア重み付けに基づく方法なども提案されている (Chen et al., 2020; Li et al., 2023)。これらの頻度流の手法は単一の外部情報源を利用する設計になっており、複数の外部情報源の利用は想定されていない。

5. 既存法の問題点

緒言 3 節で述べたように、情報借用の正当性を定性的に判断することは容易ではなく、またその判断には誤りが生じ得ることを考慮する必要がある。適応的な方法を用いることで、検出力の低下や第一種の過誤確率の上昇といった情報借用のリスクを抑制しながら、情報借用の主要な目的である検出力の上昇を達成できる可能性がある。

一方、既存の適応的借用法は、情報源同士の同質性に基づいて借用量を適応させるアプローチを採用しているため、第一種の過誤確率と検出力を明示的に制御することができない。RCT では、臨床的に意味のある最小の差 (minimal clinically important difference ; 以下, MCID) と事前指定する有意水準に基づき、所望の検出力を達成するためのサンプルサイズを設計することが標準的である (ICH, 1998)。ハイブリッド対照試験はこのサンプルサイズの達成が困難な状況に検討し得る選択肢であるが、外部情報をむやみに借用すればいいというわけではなく、所望の検出力と第一種の過誤確率に基づいた設計とすることが規制の観点からも望ましい。FDA は、外部情報を解析に用いる文脈で、最大・最小サンプルサイズを事前に規定することを推奨している (FDA, 2010)。ほとんどの既存法は、解析手法固有のモデルパラメータをシミュレーションに基づいて調整することで動作特性を制御できるが、試験計画の観点では、第一種の過誤確率と検出力は数値的試行錯誤ではなく解析手法の設計によって明示的に制御できることが望ましい。

また、情報借用に関する既存法の多くはベイズ法に基づいているものの、臨床試験の計画と解析においては頻度論に基づく手法が主流になっている (ICH,

1998). 本研究では解析の枠組みの優劣を論じることは意図していないものの、第一種の過誤確率や検出力の明示的な制御という観点ではベイズ法よりも頻度論的手法の方が整合的である。前節で述べたとおり、頻度論に基づく情報借用法も提案されているが、それらはすべて単一の外部情報源を利用するための方法であり、臨床試験において複数の外部情報源を適応的に利用することを念頭に置いて開発された頻度論的手法は筆者の知る限り存在しない。

6. 本研究の目的

ここまで、臨床試験の対照群の一部として外部情報を利用するモチベーションを示し、その中でも適応的方法が有用であることを事例を挙げながら説明した。また適応的情報借用に関する既存法を取り上げ、その問題点を指摘した。特に、検出力と第一種の過誤確率を明示的に制御できないという問題は臨床試験の計画において望ましくない。

情報借用の文脈では、検出力と第一種の過誤確率は用いる外部情報に依存するため、これらを一様に制御することはできない。例えば、帰無仮説の棄却の観点で有利な外部情報が用いられる場合は検出力と第一種の過誤確率の両方が上昇し、帰無仮説の棄却に不利な外部情報が用いられる場合は検出力と第一種の過誤確率の両方が低下することが通常である。また、第一種の過誤確率を制御したうえで検出力を上昇させることは一般的に不可能であることも示されている (Kopp-Schneider et al., 2020)。

以上の内容を踏まえ、本研究では、第一種の過誤確率と検出力の名目水準への制御と適応的な情報借用を両立する手法を開発することを目的とする。特に、既存法では実現できていない、頻度論の枠組みのもとで複数の外部情報源を利用することに焦点を当てる。具体的には、適応的な重み付き最尤推定に有効サンプルサイズ (effective sample size; 以下, ESS) を一定に保つ制約条件を導入した adaptive constrained weighted estimation (ACWE) を提案する。これによって、外部情報源が同時対照と同質な場合は第一種の過誤確率と検出力を明示的に制御し、そうでない場合は適応的な借用を実現できると考えた。

本研究では連続アウトカムを対象に提案法を定式化し、数値例によって提案法の各外部情報源への重み付けの性質を探索する。モンテカルロシミュレーションに基づいて動作特性を評価し、さらに実際の臨床試験データへの適用によって手法の挙動を確認する。上記のすべての検討は既存法と比較させたうえでを行い、提案法の統計的性質を様々な角度から検討する。

方法

1. 定式化

本節では提案法の基本設計を確立する．比較のために，複数の外部情報源を利用する文脈における代表的な既存法である MAP 法と power prior 法の数理も示す．Power prior は緒言で取り上げた MPP と EBPP を対象とする．

1.1. 提案法

本研究では，連続アウトカムを観測する RCT に基づき，次の仮説を検定することを主要な関心とする．

$$H_0: \delta \leq 0 \text{ versus } H_1: \delta > 0$$

ここで $\delta = \theta_{trt} - \theta^{(0)}$ であり， θ_{trt} と $\theta^{(0)}$ はそれぞれ試験治療群と同時対照群の母平均である．緒言 5 節で述べたとおり，RCT では事前指定する有意水準のもとで所望の検出力を達成するサンプルサイズを試験のデザイン段階で決定することが通常である．これに基づいた各群の必要サンプルサイズをそれぞれ n_{trt} と n^* とする．このとき，既知の分散 σ^2 のもとで，上式の仮説に対する検定統計量は

$$Z = \frac{\hat{\delta}}{\sqrt{\frac{\sigma^2}{n_{trt}} + \frac{\sigma^2}{n^*}}} \sim N(0, 1^2)$$

と表される．ただし $\hat{\delta} = \hat{\theta}_{trt} - \hat{\theta}^{(0)}$ であり， $\hat{\theta}_{trt}$ と $\hat{\theta}^{(0)}$ はそれぞれ試験治療群と対照群の標本平均である．

ここまでは標準的な RCT の状況設定を述べた．本研究では臨床試験における情報借用をテーマとするため，次の状況を想定する．

- 試験で実際に登録可能な対象者数は $(n_{trt} + n_0)$ であり，なおかつ $n_0 < n^*$ であるとする
- 同時対照群と関連する複数の外部情報源が利用可能である

このとき，緒言で説明したハイブリッド対照デザインに整合する形で，試験治療群と同時対照群にそれぞれ n_{trt} 人と n_0 人を割り当て，同時対照群で不足する $(n^* - n_0)$ 人に相当する情報を複数の外部情報源から補うことを考える．

対照群として用いる情報源を $\mathbf{X}^{(s)}$, $s = 0, \dots, S$ と表記する．ただし $s = 0$ は同時対照の番号を， $s \geq 1$ は外部情報源の番号を表す．後述するように，外部情報源はデータセットである必要はなく，要約統計量でも構わない．各情報源のサンプルサイズを n_s とし，アウトカムは正規分布に従い，

$$X_i^{(s)} \stackrel{i.i.d.}{\sim} N(\theta^{(s)}, \sigma^2), i = 1, \dots, n_s$$

を仮定する．また本研究を通じて、各情報源において共変量分布はそろっていると仮定する．以上の設定のもと、各情報源に重み $w^{(s)} \in [0,1]$ を与えた重み付き最尤推定に基づく情報借用を定式化する．ただし、同時対照群の情報はすべて用いたいため $w^{(0)} = 1$ とする．この場合、対照群全体における重み付き平均 θ^* の目的関数は

$$\begin{aligned} \log L(\theta^* | \mathbf{w}, \sigma^2, \mathbf{X}) &= \sum_{s=0}^S \sum_{i=1}^{n_s} w^{(s)} \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (X_i^{(s)} - \theta^*)^2 \right\} \\ &= \sum_{s=0}^S w^{(s)} \left[-\frac{n_s}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \{ (n_s - 1) \hat{\sigma}^{2(s)} + n_s (\bar{X}^{(s)} - \theta^*)^2 \} \right] \end{aligned} \quad (1)$$

と表せる．ここで、 $\mathbf{w} = (w^{(0)}, \dots, w^{(S)})$, $\mathbf{X} = (\mathbf{X}^{(0)}, \dots, \mathbf{X}^{(S)})$ である．順を追って説明するためこの時点では $w^{(s)}$ を固定値としていることに注意されたい．この場合、各情報源の標本平均 $\bar{X}^{(s)}$ が得られれば θ^* を推定できるが、後述するように $w^{(s)}$ もデータから動的に推定する場合は標本不偏分散 $\hat{\sigma}^{2(s)}$ も必要となる． θ^* の推定量は単純な重み付き平均の形で

$$\hat{\theta}^* = \frac{\sum_{s=0}^S n_s w^{(s)} \bar{X}^{(s)}}{\sum_{s=0}^S n_s w^{(s)}} \quad (2)$$

と表せる．また、 $\hat{\theta}^*$ の分散は

$$V(\hat{\theta}^*) = \frac{\sigma^2 \sum_{s=0}^S n_s w^{(s)2}}{(\sum_{s=0}^S n_s w^{(s)})^2} \quad (3)$$

と表せる．

ここで、推定対象パラメータに関する ESS (Kish, 1965; Pennello and Thompson, 2008) を導入する．外部情報を取り入れたうでの重み付き平均の ESS は

$$ESS = n_0 \frac{V(\hat{\theta}^{(0)})}{V(\hat{\theta}^*)} \quad (4)$$

と表せる．ここで、 $V(\hat{\theta}^*)$ と $V(\hat{\theta}^{(0)})$ はそれぞれ情報借用をした場合としない場合の推定量の分散である．ESS は分散比に基づいており、情報借用後の実質的なサンプルサイズを、推定精度の上昇に基づいて表した量と解釈できる．提案法は、対照群において不足する $(n^* - n_0)$ に相当する情報を外部情報源から借用することを目的としている．そのため、ESS が n^* に一致するように外部情報源に対する重み $(w^{(1)}, \dots, w^{(S)})$ を設定することでこの目的を達成することができると考えた．すなわち、(3)式および(4)式から、

$$\frac{(\sum_{s=0}^S n_s w^{(s)})^2}{\sum_{s=0}^S n_s w^{(s)2}} = n^* \quad (5)$$

を満たす $(w^{(1)}, \dots, w^{(S)})$ を設定する． $w^{(s)}$ を固定値として扱う場合、(5)式の制約

を満たす範囲で $(w^{(1)}, \dots, w^{(S)})$ を事前指定することになる．一方，緒言 3 節で述べたとおり，情報借用の正当性を定性的に判断することは容易ではないため，外部情報源に対する重みをデータから推定できることが望ましい．そこで，(5)式の制約のもとで，(1)式の対数尤度関数を最大化する重みを用いればよいと考えた．これは等式制約付きの最適化問題であり， $(w^{(1)}, \dots, w^{(S)})$ の最適解は

$$\begin{aligned} & (\hat{w}^{(1)}, \dots, \hat{w}^{(S)}) \\ &= \arg \max_{(w^{(1)}, \dots, w^{(S)})} \left[\sum_{s=0}^S w^{(s)} \left[-\frac{n_s}{2} \log(2\pi\sigma^2) \right. \right. \\ & \quad \left. \left. - \frac{1}{2\sigma^2} \{ (n_s - 1)\hat{\sigma}^{2(s)} + n_s(\bar{X}^{(s)} - \theta^*)^2 \} \right] \right], \\ & \text{subject to } \frac{(\sum_{s=0}^S n_s w^{(s)})^2}{\sum_{s=0}^S n_s w^{(s)^2}} = n^* \end{aligned}$$

によって得られる．この $(\hat{w}^{(1)}, \dots, \hat{w}^{(S)})$ を(2)式にプラグインする形で，重み付き平均の推定量が

$$\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)} = \frac{n_0 \bar{X}^{(0)} + \sum_{s=1}^S n_s \hat{w}^{(s)} \bar{X}^{(s)}}{n_0 + \sum_{s=1}^S n_s \hat{w}^{(s)}} \quad (6)$$

と得られる．一方，ここまでの説明に基づけば，(3)式から(5)式において， $(w^{(1)}, \dots, w^{(S)})$ の推定誤差を考慮していないため， $\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}$ の分散は過小評価され，実際の ESS が n^* に満たないという問題が生じる．そこで， $\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}$ の分散をパラメトリックブートストラップ法によって $\tilde{V}(\hat{\theta}^*)$ として推定し，次に(5)式の制約条件を

$$\frac{(\sum_{s=0}^S n_s w^{(s)})^2}{\sum_{s=0}^S n_s w^{(s)^2}} = \varphi n^*$$

と更新することでこの問題に対処する． $\varphi = \frac{\tilde{V}(\hat{\theta}^*)}{V(\hat{\theta}^*)}$ は重み付き平均の推定量の分散の過小評価を補正する係数である．この更新のあと，同様の最適化問題を解くことで，最適な重みは

$$\begin{aligned} & (\hat{w}^{(1)}, \dots, \hat{w}^{(S)}) \\ &= \arg \max_{(w^{(1)}, \dots, w^{(S)})} \left[\sum_{s=0}^S w^{(s)} \left[-\frac{n_s}{2} \log(2\pi\sigma^2) \right. \right. \\ & \quad \left. \left. - \frac{1}{2\sigma^2} \{ (n_s - 1)\hat{\sigma}^{2(s)} + n_s(\bar{X}^{(s)} - \theta^*)^2 \} \right] \right], \end{aligned}$$

$$\text{subject to } \frac{(\sum_{s=0}^S n_s w^{(s)})^2}{\sum_{s=0}^S n_s w^{(s)^2}} = \varphi n^*$$

と得られる.

パラメトリックブートストラップでは, 各対照情報源の標本平均を $\bar{X}_b^{(s)} \sim N(\bar{X}^{(s)}, \frac{\sigma^2}{n_s})$, $b = 1, \dots, B$ に基づいてリサンプリングし, $\hat{\theta}_b^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}$ および $\tilde{V}(\hat{\theta}^*)$ を推定する. すなわち, 重み付き平均の推定量の分散を経験的に推定し, 対応する制約条件を用いることで, $\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}$ の ESS が n^* を達成することを意図している. また $\theta^{(0)} = \dots = \theta^{(S)}$ においては $E(\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}) = \theta^{(0)}$ であり, (6) 式の $\hat{\theta}^*$ は $\theta^{(0)}$ の不偏推定量である. したがって, 母集団が同質な対照情報源のもとで, 群間差 $\delta = \theta_{trt} - \theta^*$ の検定統計量

$$Z = \frac{\hat{\delta}}{\sqrt{\frac{\sigma^2}{n_{trt}} + \frac{\sigma^2}{n^*}}} \sim N(0, 1^2) | \theta^{(0)} = \dots = \theta^{(S)}$$

に基づいた検定は第一種の過誤確率と検出力を名目水準に制御する.

上記の議論は母集団の同質性に基づいているが, 外部情報の特定の実現値によって条件付けているわけではない. 実際の臨床試験において, 外部情報は計画段階で通常既知であるため, 外部情報源の特定の実現値の組み合わせである $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(S)})$ のもとでの動作特性に焦点が当たる場合がある (Lim et al., 2020; Kopp-Schneider et al., 2024). この場合, $\theta^{(0)} = \bar{x}^{(1)} = \dots = \bar{x}^{(S)}$ であれば $\hat{\theta}^*$ は $\theta^{(0)}$ の不偏推定量である. 加えて, 提案法はブートストラップ法に基づいて ESS を再調整するため, $\hat{\theta}^* | \hat{w}^{(1)}, \dots, \hat{w}^{(S)}$ の ESS は同様に n^* に保たれる. すなわち, 群間差 δ の検定統計量

$$Z = \frac{\hat{\delta}}{\sqrt{\frac{\sigma^2}{n_{trt}} + \frac{\sigma^2}{n^*}}} \sim N(0, 1^2) | \theta^{(0)} = \bar{x}^{(1)} = \dots = \bar{x}^{(S)}$$

に基づいた検定は, 第一種の過誤確率と検出力を名目水準に制御する. したがって, 特定の外部情報で条件付けるか否かに関わらず, 提案法は必要な仮定のもとで同じ動作特性を提供する. 必要な仮定とは, 特定の外部情報を想定する場合には $\theta^{(0)} = \bar{x}^{(1)} = \dots = \bar{x}^{(S)}$, そうでない場合は $\theta^{(0)} = \dots = \theta^{(S)}$ である.

提案法では, 同時対照に与える重みを $w^{(0)} = 1$ と固定するため, 各外部情報源のサンプルサイズと標本分散が互いに等しいのであれば, 標本平均が同時対照平均により近い外部情報源により大きな重みを与えることが尤度の観点から有利である. すなわち提案法は, 同時対照群と類似した外部情報源により大きな重みを与える傾向を有する. また, 標本分散がより小さい情報源は尤度への寄与が

相対的に大きいため、提案法は標本分散がより小さい情報源により大きな重みを与える傾向を有する。さらに、制約条件のもとで、よりサイズが小さい外部情報源に大きな重みを与える方が尤度の観点で有利であるため、提案法の最適な重みは各外部情報源の標本平均や標本分散だけでなくサンプルサイズによっても調整される。提案法の重み付けの実際の挙動は2節以降示す。

既存の適応的借用法は、情報源間の同質性の高さに応じて借用量を動的に変化させるアプローチを採用している。対照的に、提案法は借用後の ESS を一定に保ち、第一種の過誤確率と検出力が同質な対照情報源のもとで名目水準に一致することを保証したうえで、各外部情報源に対する重みを動的に配分するというアプローチを採用している。この点で、提案法は既存法とは異なる設計思想に基づいている。

以降は実用上の観点から提案法を補足する。提案法は前述の通り等式制約付きの最適化問題に基づいており、R パッケージ `nloptr` によって数値的に解くことができる。また、最適な重みを用いた重み付きサンプルサイズ $n_s \hat{w}^{(s)}$ を用いて、情報量の観点での各情報源の寄与を比較することができる。ただし、重み付きサンプルサイズは(5)式で示した $V(\hat{\theta}^*)$ に基づく ESS とは異なるため、 $n_s \hat{w}^{(s)}$ の合計が所望のサンプルサイズ n^* に一致するわけではない。

提案法におけるブートストラップ法はパラメトリックではなくノンパラメトリックとすることもできるが、その場合は外部情報が個人レベルのデータとして利用可能である必要がある。外部情報は論文などにおいて要約統計量の形で公表される状況は多いものの、必ずしも個人レベルのデータが利用できるわけではない。このことを考慮し、本研究ではパラメトリックブートストラップを用いている。

提案法における重みは、ブートストラップ信頼区間に基づいてばらつきを評価することもできる。ただしこのブートストラップは前述した制約条件の再定義のためのブートストラップとは異なることに注意されたい。

1.2. 既存法

1.2.1. メタアナリシスのアプローチ

提案法との比較のために、複数の外部情報源が利用できる状況で頻繁に議論され (FDA, 2010; Gsteiger et al., 2013; Neuenschwander et al., 2010; Kopp-Schneider et al., 2020; Pang and Zhu, 2024; Schmidli et al., 2014; Viele et al., 2014), また実際の臨床試験でも用いられた (Baeten et al., 2013; Hueber et al., 2012) メタアナリシスのアプローチを説明する。前節の記法を用い、さらに各対照パラメータは $\theta^{(s)} \sim N(\mu, \tau^2), s = 0, \dots, S$ の事前分布に従い、 (μ, τ) は既知の事前分布に従うとする。

MAP 法は、試験のデザイン段階に外部情報源から導いた MAP 事前分布を、試験終了後にベイズの定理によって同時対照データと組み合わせて推論する二段階アプローチである。すなわち、当該試験データの収集前の時点で外部情報源を用いて、 $\theta^{(0)}$ に関する MAP 事前分布を

$$P(\theta^{(0)}|\mathbf{X}^{(-0)}) \propto \int \dots \int \prod_{s=1}^S \left(L(\theta^{(s)}|\mathbf{X}^{(s)})P(\theta^{(s)}|\mu, \tau) \right) P(\mu, \tau)P(\theta^{(0)}|\mu, \tau)d\mu d\tau$$

と構成する。ただし $\mathbf{X}^{(-0)} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(S)})$, $\theta^{(-0)} = (\theta^{(1)}, \dots, \theta^{(S)})$ である。この MAP 事前分布を同時対照データに基づく尤度関数と組み合わせることで、 $\theta^{(0)}$ の周辺事後分布が

$$P(\theta^{(0)}|\mathbf{X}) \propto L(\theta^{(0)}|\mathbf{X}^{(0)}) \int \dots \int \prod_{s=1}^S \left(L(\theta^{(s)}|\mathbf{X}^{(s)})P(\theta^{(s)}|\mu, \tau) \right) P(\mu, \tau)P(\theta^{(0)}|\mu, \tau)d\mu d\tau$$

と得られる。このモデルでは、情報源間の変動である τ の事前分布（半正規分布や半 Cauchy 分布などが用いられる）のパラメータが借用の程度を制御するチューニングパラメータとなる。

また Schmidli et al. (2014)は、prior-data conflict をさらに考慮するため、MAP 事前分布を曖昧な分布との混合によってロバスト化する R-MAP 法を提案した (Schmidli et al., 2014)。この方法では、 $\theta^{(0)}$ の推測に用いる事前分布を、外部情報に基づく MAP 成分と外部情報に基づかないロバスト成分の両方によって構成することができる。すなわち、ロバスト成分の混合割合を w_R 、 $\theta^{(0)}$ に対する曖昧な分布を $P_R(\theta^{(0)})$ とすると、ロバスト MAP 事前分布は

$$P(\theta^{(0)}|\mathbf{X}^{(-0)}) = (1 - w_R)P(\theta^{(0)}|\mathbf{X}^{(-0)}) + w_R P_R(\theta^{(0)})$$

と表される。

メタアナリシス的アプローチでは、対照情報源全体における類似性をもとに τ が推定される。そのため、試験治療の有効性の証明の観点で有望な外部情報源とそうでない外部情報源が混在している場合は対照パラメータの推定精度が上昇しにくく、結果として借用量が抑えられやすいという性質をもっている。

1.2.2. Power prior 法

Power prior 法は、外部情報に基づく尤度関数を 0 から 1 の値を取り得るパラメータでべき乗する（すなわち対数尤度関数を重み付けする）ことによって外部情報を利用する手法である (Ibrahim and Chen, 2000)。このべき乗パラメータは power parameter と呼ばれ、定数を事前指定するアプローチと、データに基づいて推定するアプローチが存在する。前者に基づく power prior は conditional power prior (CPP) と呼ばれる (Ibrahim and Chen, 2000)。後者に基づく power prior は、

もともと power parameter を興味のあるパラメータと同時に推測する joint power prior (JPP) として提案された (Ibrahim and Chen, 2000).

ここでは複数の外部情報源を利用するための power prior を説明する. 本研究では power parameter を $\omega = (\omega^{(1)}, \dots, \omega^{(S)})$ と表記する. 各対照情報源が $X_i^{(s)} \sim N(\theta_c, \sigma^2), s = 0, \dots, S, i = 1, \dots, n_s$ の正規分布に従い, 対照パラメータ θ_c と power parameter $\omega^{(s)}$ は既知の初期事前分布に従うとする. メタアナリシスのアプローチとは異なり, ここでは各対照情報源について共通のパラメータ θ_c を仮定していることに注意されたい. この仕様のもとで, CPP は

$$P(\theta_c | \mathbf{X}^{(-0)}, \omega) \propto \left\{ \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} \right\} P(\theta_c)$$

と定義される. JPP では ω も同時に推定するため, 初期事前分布 $P(\omega)$ を用いて

$$P(\theta_c, \omega | \mathbf{X}^{(-0)}) \propto \left\{ \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} P(\omega^{(s)}) \right\} P(\theta_c)$$

表される.

1.2.2.1. Modified power prior

JPP において, 外部情報に基づく尤度関数を定数倍すると power prior が異なる分布になるという問題, すなわち尤度原理への違反が指摘され, これを解決するための正規化定数を導入した手法である modified power prior (MPP) が提案された (Duan et al., 2006; Neuenschwander et al., 2009). MPP は

$$P(\theta_c, \omega | \mathbf{X}^{(-0)}) \propto \frac{\left\{ \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} P(\omega^{(s)}) \right\} P(\theta_c)}{g(\omega)}$$

と定義される. 尤度原理を満たすための正規化定数は

$$g(\omega) = \int_{-\infty}^{\infty} \left\{ \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} \right\} P(\theta_c) d\theta_c \quad (7)$$

で与えられる. (7)式の積分は一般的に困難であるが, 本稿で考慮する設定では解析的に導出することができる. この導出は付録で示している.

MPP を同時対照に基づく尤度関数と組み合わせることで, 同時事後分布は

$$P(\theta_c, \omega | \mathbf{X}) \propto L(\theta_c | \mathbf{X}^{(0)}) P(\theta_c, \omega | \mathbf{X}^{(-0)})$$

となる. このモデルでは $\omega^{(s)}$ の初期事前分布 (通常は Beta 分布が用いられる) のパラメータが借用の程度を制御するチューニングパラメータとなる.

MPP では外部情報源ごとに個別の power parameter が設定され, それらの事後分布には同時対照データセットと各外部情報源の類似性が反映される. この性

質は、対照情報源全体における類似性に基づいて借用が促進または抑制されるメタアナリシスのアプローチとは対照的である。

1.2.2.2. Empirical Bayes power prior

EBPP は経験ベイズ推定に基づいた power prior である (Gravestock and Held, 2017; Gravestock and Held, 2019)。前節で示した MPP は、power parameter を θ_c と同時に推定するフルベイズ推定に基づいているが、EBPP は power parameter の周辺尤度の最大化によって最適な重みを推定し、得られた重みを定数の power parameter として用いる。前節と同じ記法を用いると、power parameter の周辺尤度は

$$\begin{aligned} L(\boldsymbol{\omega}|\mathbf{X}) &= \int_{-\infty}^{\infty} L(\theta_c|\mathbf{X}^{(0)})P(\theta_c|\boldsymbol{\omega}, \mathbf{X}^{(-0)}) d\theta_c \\ &= \int_{-\infty}^{\infty} L(\theta_c|\mathbf{X}^{(0)}) \frac{\prod_{s=1}^S L(\theta_c|\mathbf{X}^{(s)})^{\omega^{(s)}} P(\theta_c)}{g(\boldsymbol{\omega})} d\theta_c \end{aligned} \quad (8)$$

と表される。(8)式の積分は一般的に困難であるが、(7)式と同様に、本稿で考慮する設定では解析的に導出することができる。この導出の詳細は付録で示して

いるが、 θ_c に一様な初期事前分布を仮定し、また $\rho^{(s)} = \frac{n_s}{\sigma^2}$ とおくと

$$L(\boldsymbol{\omega}|\mathbf{X}) \propto \sqrt{\frac{2\pi \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}{\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}} \exp\left(-\frac{\rho^{(0)}\{\sum_{s=1}^S \omega^{(s)} \rho^{(s)}(\bar{X}^{(s)} - \bar{X}^{(0)})\}^2}{2(\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)})(\sum_{s=1}^S \omega^{(s)} \rho^{(s)})}\right)$$

と表すことができる。この周辺尤度に対する最適解

$$\hat{\boldsymbol{\omega}} = \arg \max_{\boldsymbol{\omega} \in [0,1]^S} L(\boldsymbol{\omega}|\mathbf{X})$$

を用いて、EBPP は

$$P(\theta_c|\hat{\boldsymbol{\omega}}, \mathbf{X}^{(-0)}) \propto \prod_{s=1}^S L(\theta_c|\mathbf{X}^{(s)})^{\hat{\omega}^{(s)}} P(\theta_c)$$

と得られる。この手法は power parameter の事前分布を有さず、したがって借用性に関するチューニングパラメータも有さない。

2. 数値例

本節では、数値例を用いて、前節までに示した提案法と既存法が様々な外部情報源の組み合わせのもとでどのような重み付けをするかを示す。既存法では、提案法と同様に外部情報の解析への寄与を重み付けする手法である MPP と EBPP

を対象とする。

連続アウトカムに基づいて試験治療群と対照群の平均値を比較する RCT を想定する。同時対照のサンプルサイズが 50 であり，さらに 50 人分の情報を 2 つの外部情報源（外部情報源 1 と 2）から補強したいとする。また同時対照の標本平均は 0 であり，母分散と標本分散はいずれの情報源も 1 とする。

ここでは表 1 で示す 6 つのシナリオを考慮する。シナリオ 1—3 では，各外部情報源のサンプルサイズがいずれも 100 の場合に，外部情報源 1 の標本平均 $\bar{x}^{(1)}$ を $[-2,2]$ の区間内で 0.01 ずつ動かしながら，各外部情報源に対する最適な重み $(\hat{w}^{(1)}, \hat{w}^{(2)})$ と重み付きサンプルサイズ $(\hat{w}^{(1)}n_1, \hat{w}^{(2)}n_2)$ を求める。また，提案法では補正係数 φ も求める。シナリオ 4—6 では，各外部情報源の標本平均を固定したうえで，外部情報源 1 のサンプルサイズ n_1 を $[1,100]$ の区間内で 1 ずつ動かしながら，同様に最適な重みと重み付きサンプルサイズ，および提案法の補正係数を求める。

手法ごとの設定は下記のとおりである。提案法では，対照群に対する所望のサンプルサイズ n^* を 100 とした。MPP では， θ_c と $\omega^{(s)}$ の初期事前分布をそれぞれ $N(0, 10000^2)$ と $Beta(1, 1)$ とした。EBPP では， θ_c の初期事前分布を $N(0, 10000^2)$ とした。提案法と EBPP では計算された最適な重みを用い，MPP では power parameter の事後平均を用いた。

表 1. 数値例：各外部情報源のシナリオ.

シナリオ	標本平均*	サンプルサイズ†
#1	$(\bar{x}_1, 0)$	$(100, 100)$
#2	$(\bar{x}_1, 0.2)$	$(100, 100)$
#3	$(\bar{x}_1, 0.4)$	$(100, 100)$
#4	$(0, 0)$	$(n_1, 100)$
#5	$(0, 0.2)$	$(n_1, 100)$
#6	$(0, 0.4)$	$(n_1, 100)$

* $\bar{x}_1 \in \{-2, -1.99 \dots, 2\}$

† $n_1 \in \{1, 2 \dots, 100\}$

3. モンテカルロシミュレーション

連続アウトカムに基づいて試験治療と対照治療の平均値を比較する仮想的な RCT をシミュレートし，提案法と既存法の動作特性を比較評価した。試験計画時の見積もり値として，試験治療群と対照群の平均値はそれぞれ 0.4 と 0 であり，期待される治療効果 $\theta_{trt} - \theta^{(0)} = 0.4$ を想定したとする。また母分散は各群ともに 1 であるとする。このとき，片側有意水準 0.025 および検出力 0.8 とすると，

正規近似に基づくサンプルサイズ設計により、均等割付の場合は各群 100 人ずつ（合計 200 人）の登録が必要となる。しかしながら、当該試験で登録可能な対象者数は合計 150 人であり、50 人分のサンプルサイズ不足が見込まれるとする。ここで、当該試験の対照群を補足し得る 2 つの外部情報源（いずれもサンプルサイズ 100 かつ母分散 1）が利用できるとする。これを受けて、試験治療群と対照群にそれぞれ 100 人と 50 人を割り付け、対照群で不足する 50 人分の情報を外部情報源から補うハイブリッド対照試験デザインを考える。

表 2 に示す特定の外部情報源の組み合わせのもとで、手法ごとの群間差 δ の帰無仮説 $H_0: \delta \leq 0$ に対する検出力と第一種の過誤確率、 δ の推定量のバイアスと RMSE（root mean squared error）、および対照群の ESS を、外部情報源 1 の標本平均 $\bar{x}^{(1)} \in \{0, 0.1, \dots, 0.4\}$ の関数として評価した。ESS は(4)式と同様の分散比に基づいて算出した。各シナリオおよび各 $\bar{x}^{(1)}$ の値に対して合計 10 万回の反復を行った。各反復において、治療群および同時対照群のデータはそれぞれ $N(0.4, 1)$ と $N(\theta^{(0)}, 1)$ から生成した。 $\theta^{(0)}$ は対立仮説のもとでは 0 に、帰無仮説のもとでは 0.4 に固定した。

各手法に必要な設定は下記のとおりである。提案法における対照群の目標サンプルサイズは 100 とした。R-MAP では、方法 1.2.1 節の記法のもとで、 $\theta^{(s)} \sim N(\mu, \tau^2)$, $\mu \sim N(0, 10000^2)$, $\tau \sim N^+(0, 0.1^2)$ とし、ロバスト割合は 0.5 とした。

$N^+(\cdot, \cdot)$ は非負の実数値を取り得る半正規分布である。MPP では、1.2.2.1 節の記法のもとで、 θ_c と $\omega^{(s)}$ の初期事前分布をそれぞれ $N(0, 10000^2)$ と $Beta(1, 1)$ とした。EBPP では θ_c の初期事前分布を $N(0, 10000^2)$ とした。

すべてのベイズ法において、試験治療群の平均値 θ_{trt} の事前分布は $N(0, 10000^2)$ とし、治療効果 δ に対して、

$$P(\delta > 0 | \text{all information}) > 1 - 0.025 \rightarrow \text{reject } H_0: \delta \leq 0$$

の基準をもとに検定を行った。またベイズ法における点推定量には事後平均を用いた。

上記に加えて、同一の設定のもとで、表 3 に示す 4 つの外部情報源を利用する場合についても検討した。この追加検討は、提案法の重要な性質（同質な対照情報源のもとでの検出力と第一種の過誤確率の一貫した制御）が、外部情報源の数や合計外部サンプルサイズに依存しないことを示すためのものである。この追加検討は 1 万回の反復により実施した。

表 2. シミュレーション：外部情報源のシナリオ（外部情報源：2 つ）。

シナリオ	標本平均*	サンプルサイズ
#1	$(\bar{x}_1, 0)$	(100, 100)

#2	$(\bar{x}_1, 0.2)$	(100, 100)
#3	$(\bar{x}_1, 0.4)$	(100, 100)
#4	$(\bar{x}_1, 0)$	(10, 190)
#5	$(\bar{x}_1, 0.2)$	(10, 190)
#6	$(\bar{x}_1, 0.4)$	(10, 190)

* $\bar{x}_1 \in \{0, 0.1, \dots, 0.4\}$

表 3. シミュレーション：外部情報源のシナリオ（外部情報源：4 つ）.

シナリオ	標本平均*	サンプルサイズ
#A1	$(\bar{x}_1, 0, 0, 0)$	(100, 100, 100, 100)
#A2	$(\bar{x}_1, 0.4, 0.4, 0.4)$	(100, 100, 100, 100)

* $\bar{x}_1 \in \{0, 0.1, \dots, 0.4\}$

4. 臨床試験データへの適用

本節では、Hueber et al.が実施した臨床試験（Hueber et al., 2012）のデータを用いて、提案法と既存法がどのように機能するかを実演する．同試験は、中等度から重度のクローン病患者に対して secukinumab とプラセボを比較した RCT である．有効性の主要評価項目はベースラインから 6 週間後のクローン病活動性指数（CDAI）の変化量（ Δ CDAI）であった．各群で 40 人と 20 人を登録し、さらに 6 つの過去試験（Hanauer et al., 2006; Sandborn et al., 2001a; Sandborn et al., 2001b; Sandborn et al., 2007a; Sandborn et al., 2007b; Winter et al., 2004）のデータから 20 人分の情報を外部対照として用いるハイブリッド対照試験として計画された．表 4 に、これらの過去試験におけるプラセボ群のサンプルサイズと Δ CDAI の要約統計量（平均値または中央値）を示す．これらの数値は R パッケージ RBesT に収録されている（Weber et al., 2021）．Hueber 試験の結果としては、表 5 に示すように、secukinumab はプラセボ治療よりも CDAI を減少させず、クローン病に対する有効性は示されなかった．

本節では、Hueber 試験および 6 つの過去試験の結果を提案法および既存法に適用し、 Δ CDAI の平均値の群間差とその 95%信頼区間または信用区間を推定する．また、提案法と重み付け型の既存法である MPP と EBPP については、各外部情報源への重みとその 95%区間も推定する．重みの区間推定は、提案法と EBPP ではパーセンタイルに基づくブートストラップ信頼区間、MPP では重みの信用区間を用いて実施する．

提案法では、対照群に対する目標サンプルサイズを Hueber 試験の計画のとおり 40 とし、既存法における設定はすべて前節同様とした．なお、同試験の原著論文では同時対照群単独での標本平均は数値としては報告されていないため、

本節では報告された事後平均（表 5）を標本平均とみなして解析した。また各対照群の標本標準偏差も報告されていないため、同試験において真の標準偏差として仮定された値である 88 を各対照群の標本標準偏差として用いた。

表 4. Hueber et al. (2012)で外部対照として用いられた過去試験のプラセボ群.

研究	サンプルサイズ	標本平均または中央値
Sandborn et al. (2007a)	166	-49
Sandborn et al. (2007b)	328	-36
Sandborn et al. (2001a)	20	-47
Sandborn et al. (2001b)	58	-54
Hanauer et al. (2006)	74	-51
Winter et al. (2004)	25	-90

表 5. Hueber et al. (2012)による主要解析の結果.

群	サンプルサイズ	事後平均 (SD)
Secukinumab	39	-29.2 (14.0)
Placebo	20	-63.1 (13.9)
Δ CDAI	—	33.9 (19.7)

結果

1. 数値例

数値例の結果を図 1—3 に示す．図 1 と図 2 ではそれぞれ各手法における重みと重み付きサンプルサイズを示しており，図 3 では提案法に焦点を当て，重みと補正係数 ϕ を示している．図 1 および図 2 では，外部情報源を色で区別し，手法を線種で区別している（実線：提案法，破線：MPP，薄い実線：EBPP）．

最初に図 1 の結果を説明する．シナリオ 1 から，すべての手法において $\bar{x}^{(0)} = \bar{x}^{(1)} = \bar{x}^{(2)} = 0$ のときに $\hat{w}^{(1)} = \hat{w}^{(2)}$ であり，同じサンプルサイズのもとで，各手法は妥当な重み付けがされている．また，いずれの手法も $\bar{x}^{(1)}$ が $\bar{x}^{(0)} = \bar{x}^{(2)} = 0$ から離れるにつれて外部情報源 2 の相対的な影響が左右対称に増加している．提案法では， $\bar{x}^{(1)}$ の絶対値が 1 程度までは $\hat{w}^{(2)}$ だけでなく $\hat{w}^{(1)}$ も増加している．これは不均等重みによる推定効率の減少を補填するための挙動である．一方， $\bar{x}^{(1)}$ と $\bar{x}^{(0)} = \bar{x}^{(2)} = 0$ の乖離が 1 程度以上大きくなると， $\hat{w}^{(1)}$ は絶対的に減少に転じる．MPP では， $\bar{x}^{(1)}$ が $\bar{x}^{(0)} = \bar{x}^{(2)} = 0$ から離れるにつれ $\hat{w}^{(1)}$ は絶対的に減少し， $\hat{w}^{(2)}$ は絶対的に増加する． $\hat{w}^{(2)}$ の増加は $\bar{x}^{(1)}$ の絶対値が 1 程度まで続き，以降の重みはほぼ一定となっている．EBPP では， $\bar{x}^{(1)}$ が $\bar{x}^{(0)} = \bar{x}^{(2)} = 0$ から離れるにつれ $\hat{w}^{(1)}$ は急速に減少し， $\hat{w}^{(2)}$ は常に 1 である．

シナリオ 2 と 3 は $\bar{x}^{(0)} \neq \bar{x}^{(2)}$ であり，対照情報源全体で常にある程度の異質性が存在する状況を示している．これらのシナリオにおいても，シナリオ 1 と同様に， $\bar{x}^{(1)}$ の $\bar{x}^{(0)} = 0$ からの乖離に応じて各手法の重みが適応的に調整される性質が確認できる．具体的には， $\bar{x}^{(1)} \in (-\bar{x}^{(2)}, \bar{x}^{(2)})$ の領域では，外部情報源 1 の方が同時対照からの乖離が小さいため，いずれの手法でも $\hat{w}^{(1)} > \hat{w}^{(2)}$ となる．また，いずれの手法も $\bar{x}^{(1)} = \bar{x}^{(2)}$ のときに $\hat{w}^{(1)} = \hat{w}^{(2)}$ となる．一方，いずれの適応的方法においても，同時対照と外部情報が整合しない場合の重みが 0 になるとは限らない．例えば，シナリオ 2 と 3 の $\bar{x}^{(1)} = \bar{x}^{(2)}$ において，提案法は各外部情報源に約 0.115 の重みを与えており，一定の借用が行われている．これは必ずしも望ましい性質ではないが，提案法は異質性がかなり大きい場合の借用を防ぐという性質を有する．例えば，これらのシナリオの $\bar{x}^{(1)} \approx -2$ において，提案法は外部情報源 2（標本平均は 0.2 または 0.4 で固定）に約 0.25 の重みを与える一方，同時対照との乖離がかなり大きい外部情報源 1 に対する重みは 0 であり，まったく借用していない．対照的に，これらの状況で，MPP と EBPP は外部情報源 1 に 0.1 から 0.2 程度，外部情報源 2 に 0.7 以上の重みを与えており，推定量の大きなバイアスにつながる事が想定される．EBPP は，シナリオ 1 では類似した外部情報源に全借用に相当する重みを，類似していない外部情報源には無借用

に相当する重みを与えており、望ましい性質を有することが示唆されるものの、シナリオ 2 と 3 のように常に一定の異質性が存在する状況では望ましい性質を示さないことが示唆される。

シナリオ 4—6 は異なる標本平均の組み合わせのもとで、 n_1 の関数としての重みを示している。提案法はサンプルサイズがより小さい外部情報源にわずかに大きい重みを与える傾向があった。これは(1)式からも確認できるように、サイズが小さい情報源により大きな重みを与える方が重み付き対数尤度が高くなるからである。MPP は提案法とは逆の傾向を有するようである。EBPP は、各外部情報源の標本平均が同値のシナリオ 4 ではサンプルサイズの影響はないが、シナリオ 5 と 6 のように常に一定の異質性が存在する場合では他の手法と同様にサンプルサイズの影響がみられる。

図 2 は重み付きサンプルサイズを示している。図 1 のシナリオ 4—6 において、サンプルサイズが重みに影響することが確認されたが、実際のサンプルサイズのスケールではその影響は小さいようである。

図 3 は提案法における重みと補正係数 φ を示している。補正係数 φ が 1 より大きいことは(3)式による分散とブートストラップ分散の乖離が大きいことを示している。すなわち、 φ は重みの推定誤差が大きくなるほど増加する。シナリオ 1—3 では、外部情報内の異質性が小さい場合または極端に大きい場合には重みの推定誤差が小さく、 φ はほぼ 1 に等しい。一方で、 $\bar{x}^{(1)}$ の絶対値が 1 に近いとき、重みの推定誤差が最も大きくなり、 φ は大きくなる。ESS が n^* を満たすためには、合計対照サンプルサイズとして最大 φn^* が必要となる。

全体的に、提案法は MPP および EBPP と比較してより安定した重み付けを示す傾向が確認できる。これは、重みの均一性が重み付き尤度の増加に寄与する要素の一つであるためである。

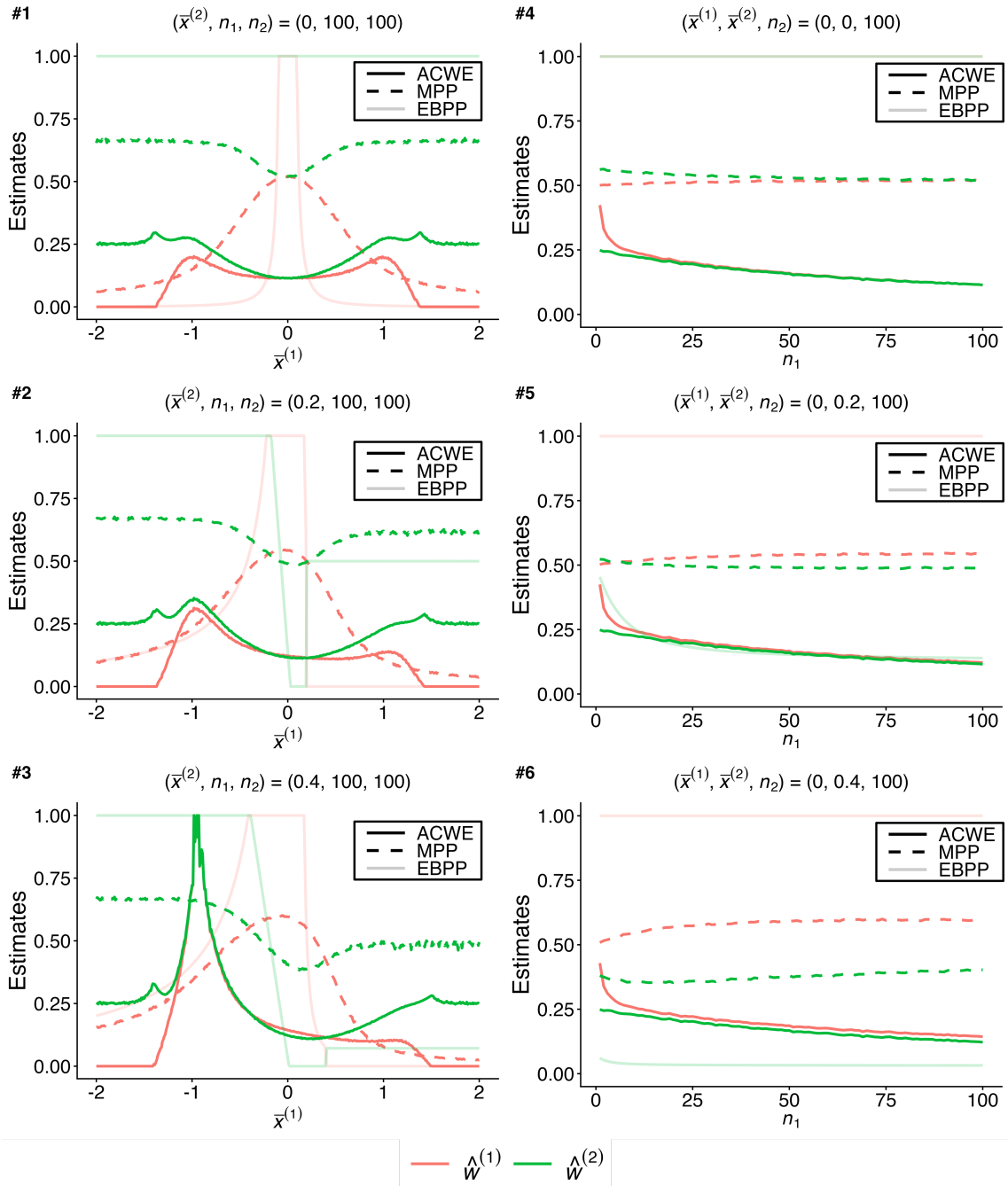


図 1. 提案法および既存法における最適な重み.

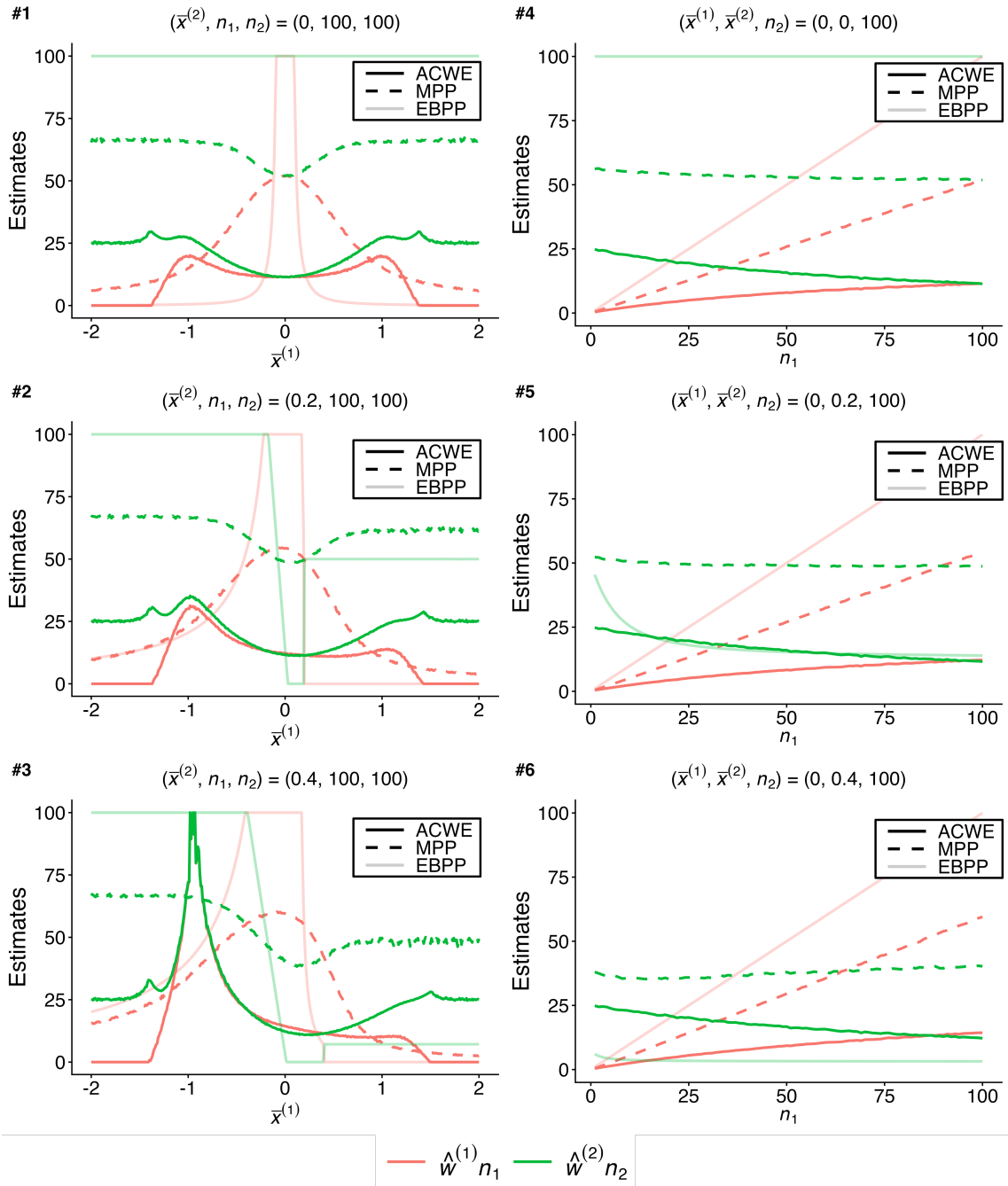


図 2. 提案法および既存法における重み付きサンプルサイズ.

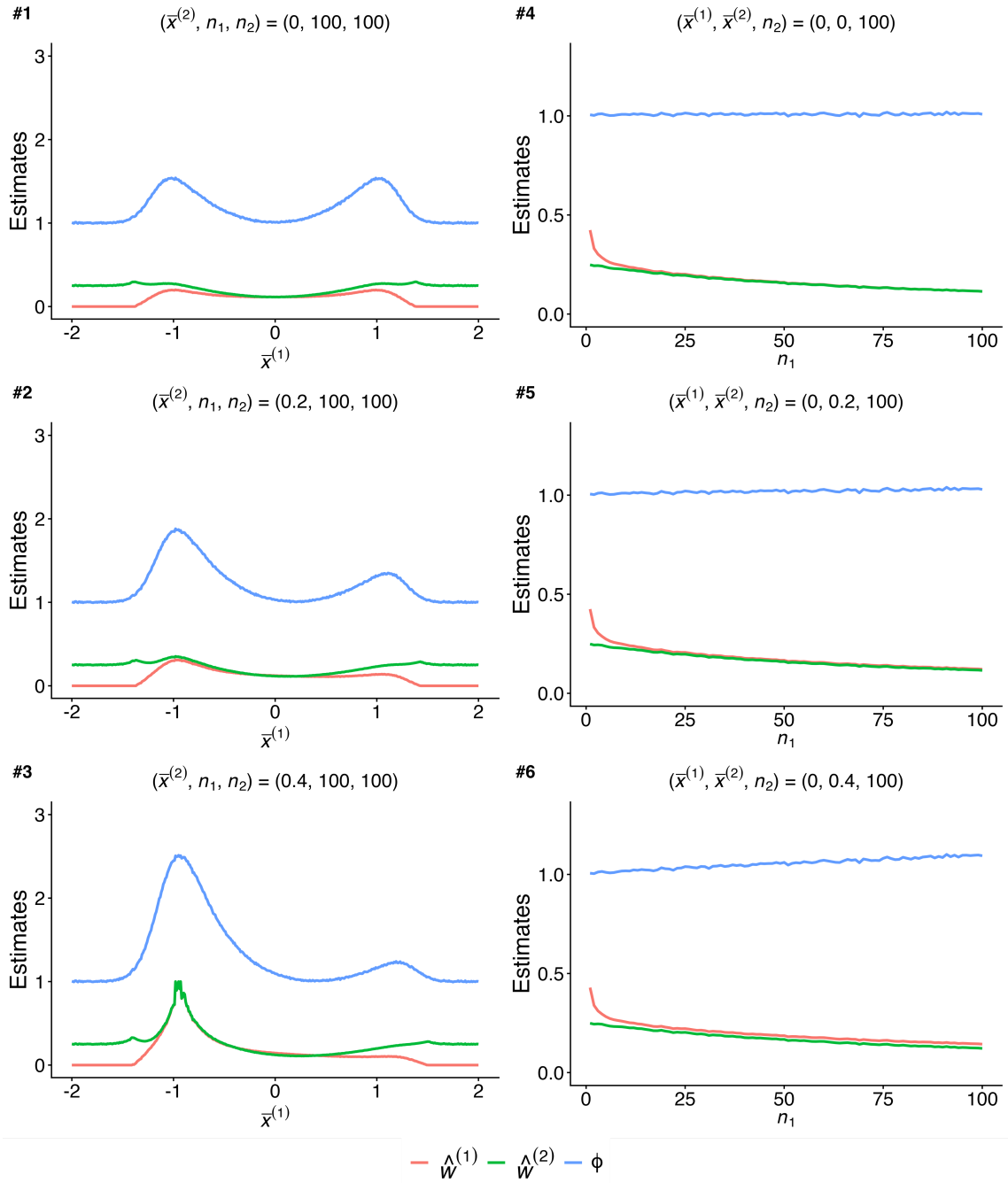


図 3. 提案法における最適な重みと補正係数 ϕ .

2. モンテカルロシミュレーション

図 4 と図 5 に検出力と第一種の過誤確率を示し、図 6 と図 7 には検出力と第一種の過誤確率の等高線図（図 6：シナリオ 1—3，図 7：シナリオ 4—6）を示す。図 8 には検出力と第一種の過誤確率の散布図を示す。図 9 には外部情報源の数を 4 とした追加検討の検出力と第一種の過誤確率を示す。すべてのシミュレーションにおける第一種の過誤確率と検出力の数値は付録表 1 と付録表 2 に

示した。また付録図 1 と付録図 2 に治療効果の推定量のバイアス（付録図 1：対立仮説下，付録図 2：帰無仮説下）を，付録図 3 と付録図 4 に治療効果の推定量の RMSE（付録図 3：対立仮説下，付録図 4：帰無仮説下）を，付録図 5 と付録図 6 に対照群の ESS（付録図 5：対立仮説下，付録図 6：帰無仮説下）を示した。

図 4 と図 5 より，提案法における検出力と第一種の過誤確率は， $\bar{x}^{(1)} = \bar{x}^{(2)} = 0$ のときに 0.8 に， $\bar{x}^{(1)} = \bar{x}^{(2)} = 0.4$ のときに 0.025 にそれぞれ制御されている。すなわち方法 1.1 節で説明したとおり，同質な対照情報源のもとで，提案法は検出力と第一種の過誤確率を名目水準に制御できていることが確認された。

今回は片側帰無仮説 $H_0: \delta \leq 0$ を設定しているため，外部標本平均が大きいほど帰無仮説が棄却されづらくなり，図 4 と図 5 においても， $\bar{x}^{(1)}$ が増加するにつれて情報借用法における検出力と第一種の過誤確率の両方が低下する傾向を確認することができる。提案法は同質な状況下で検出力 0.8 を維持することを目的としているため， $\theta^{(0)} = \bar{x}^{(1)} = \bar{x}^{(2)} = 0$ のように借用が適切とみられる場合でも過剰な借用を回避する。一方，中程度の異質性が存在する状況では，適応的な重み付けを通じて検出力の低下と第一種の過誤確率の増加の両方を緩和する。例えばシナリオ 2 では，提案法は検出力が比較的高く，また第一種の過誤確率の増加も比較的抑えられている。さらにシナリオ 3 における $\bar{x}^{(1)} \geq 0.2$ のような借用が望ましくないとみられる領域においても，提案法は検出力の大幅な低下を抑制している。EBPP は，帰無仮説または対立仮説のいずれかに対応する外部情報源が常に利用可能なシナリオ 1 および 3 において，適切な外部情報源の特定に優れた性能を発揮する。同手法はシナリオ 2 でも提案法と同等またはそれ以上の検出力を持つが，第一種の過誤確率の増加が顕著であり，借用の適切性が曖昧な状況には適さない可能性があることを示唆している。MPP は，シナリオ 1 における $\bar{x}^{(1)} \leq 0.2$ のような状況では提案法より高い検出力を示すが，多くの状況で第一種の過誤確率の増加が大きい。R-MAP は異質性が示唆される状況において借用の程度を制限する性能が高く，他の借用法よりも第一種の過誤確率の増加の抑制に優れていたが，検出力が低い傾向がみられた。

シナリオ 4—6 は，シナリオ 1—3 の外部サンプルサイズ構成を変化させたものであり， $(n_1, n_2) = (10, 190)$ である。これらのシナリオでもシナリオ 1—3 と同様に， $\bar{x}^{(1)}$ が大きくなるにつれ検出力が上昇し第一種の過誤確率が低下する傾向にあるが，サンプルサイズがかなり小さい外部情報源 1 は解析への影響が小さく，検出力と第一種の過誤確率は $\bar{x}^{(1)}$ の変化に対してほぼ一定であった。前節では外部サンプルサイズが最適な重みに影響することを示したが，実際の検出力および第一種の過誤確率における影響は限定的であり，このことは図 2 で確認した重み付きサンプルサイズにおける結果と同様であった。

上記で示した各手法の特徴は図 6 と図 7 の等高線図，および図 8 の散布図か

らも読み取ることができる．例えば提案法と R-MAP は，EBPP や MPP と比較すると，検出力の最大値は小さいものの検出力自体のばらつきが少なく安定した挙動を示す傾向がみられた．また図 8 の散布図からは，提案法は検出力と第一種の過誤確率の関係が比較的滑らかであることが確認された．

方法 1.1 節で示したとおり，提案法は借用を一定の水準に制御するため，同質な対照情報源のもとでは外部情報源の数が増えても検出力と第一種の過誤確率は変わらない．図 9 から，提案法における検出力と第一種の過誤確率は， $\bar{x}^{(1)} = \dots = \bar{x}^{(4)} = 0$ のときに 0.8 に， $\bar{x}^{(1)} = \dots = \bar{x}^{(4)} = 0.4$ のときに 0.025 にそれぞれ制御されていることが確認できる．一方，既存の情報借用法は，外部情報源の数やサンプルサイズによって借用の程度が変化するためこの性質は有さない．すなわち既存法は提案法とは異なり，検出力と第一種の過誤確率を事前に検討する際に，外部情報源の数やサンプルサイズも考慮に入れる必要がある．

その他，付録図 1—4 から，提案法のバイアスは比較的抑えられていたが，既存の情報借用法の中では中程度の水準であった．外部情報源間の異質性が高い状況では，EBPP は提案法よりも低いバイアスを示し，借用の適切性が曖昧な状況ではそれらは同程度のバイアスを示した．RMSE については提案法が優れている傾向があった．また，付録図 5—6 では提案法は ESS が常に一定であることが確認できる．

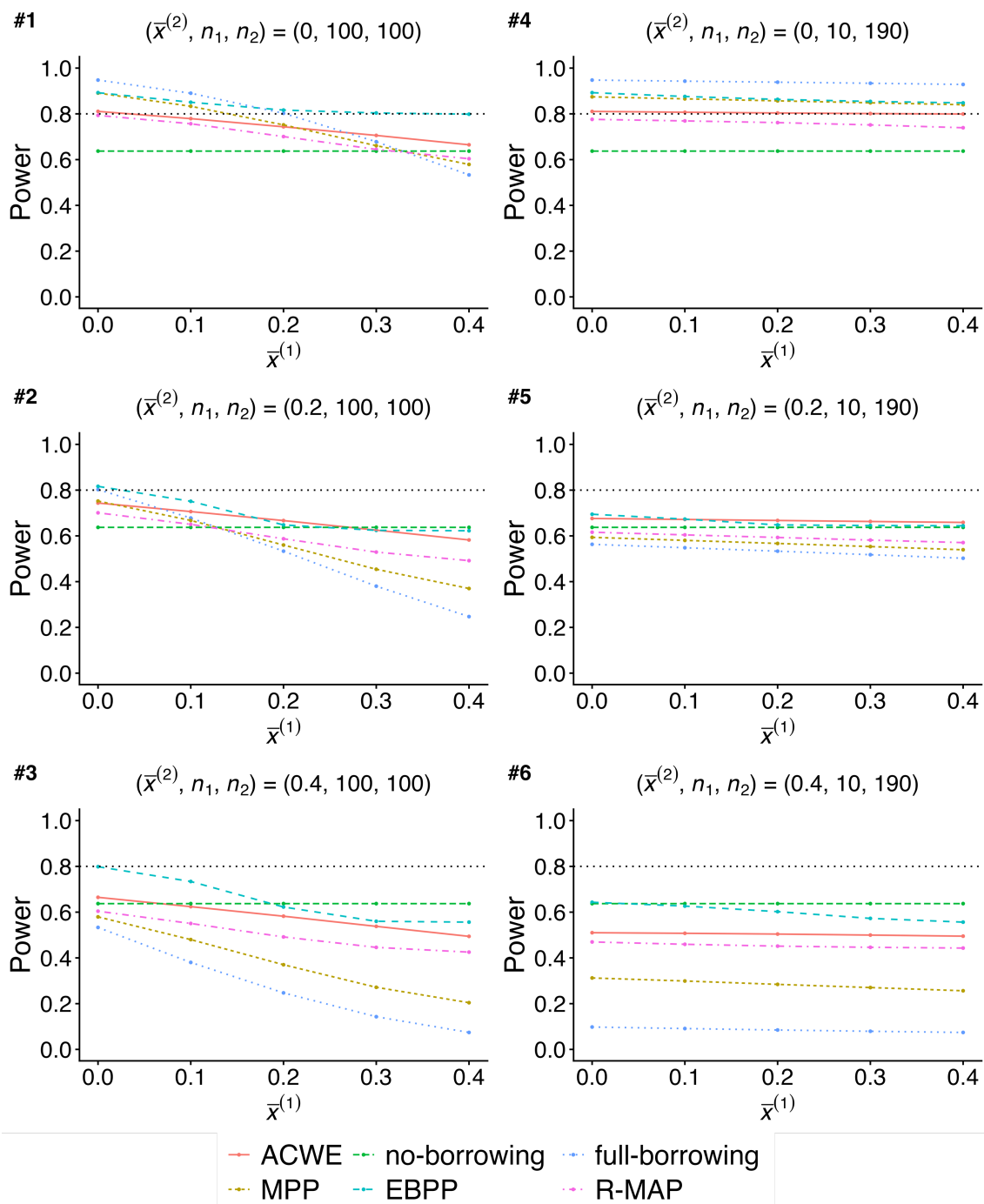


図 4. 検出力.

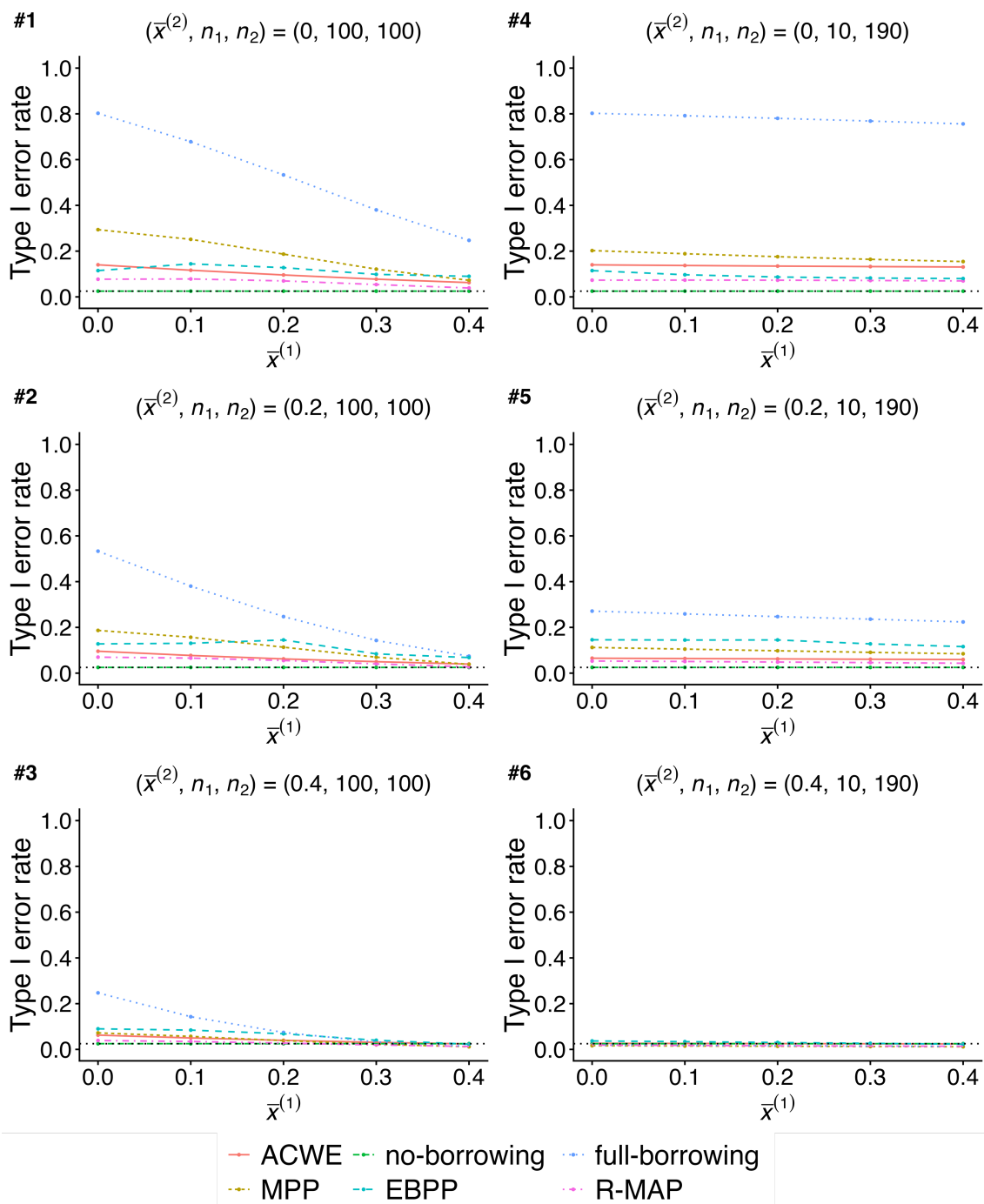


図 5. 第一種の過誤確率.

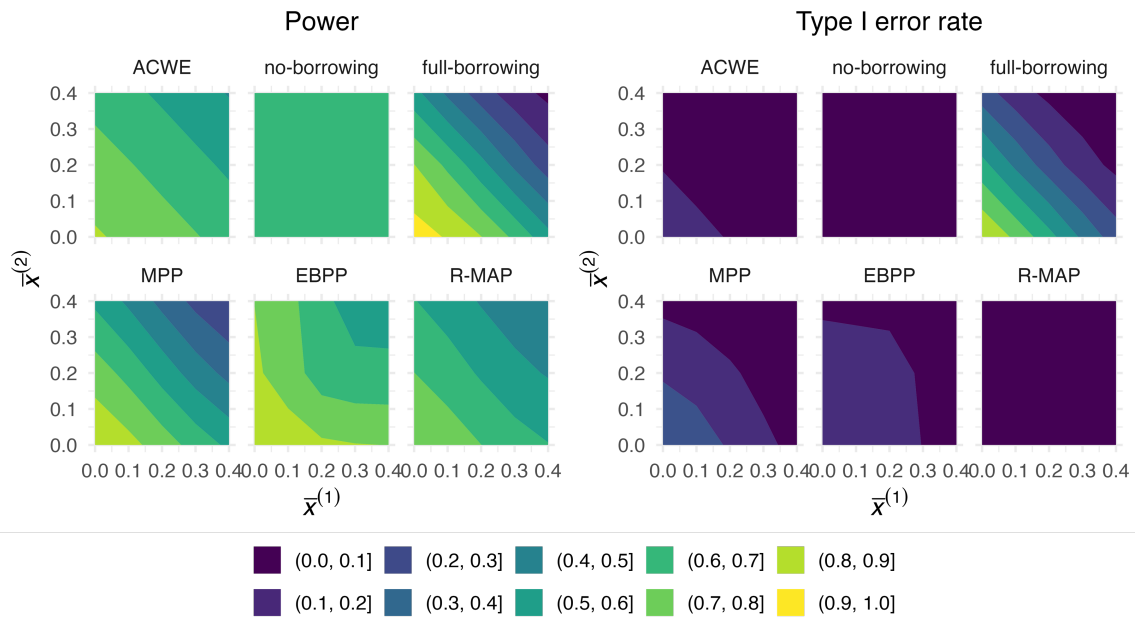


図 6. 等高線図（シナリオ 1—3）.

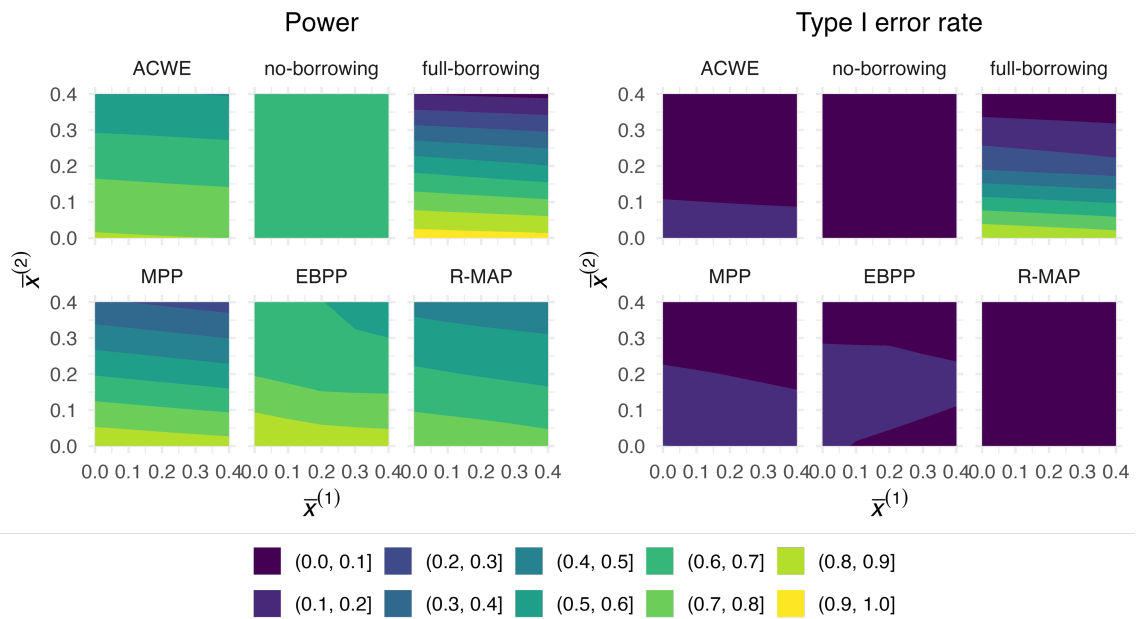


図 7. 等高線図（シナリオ 4—6）.

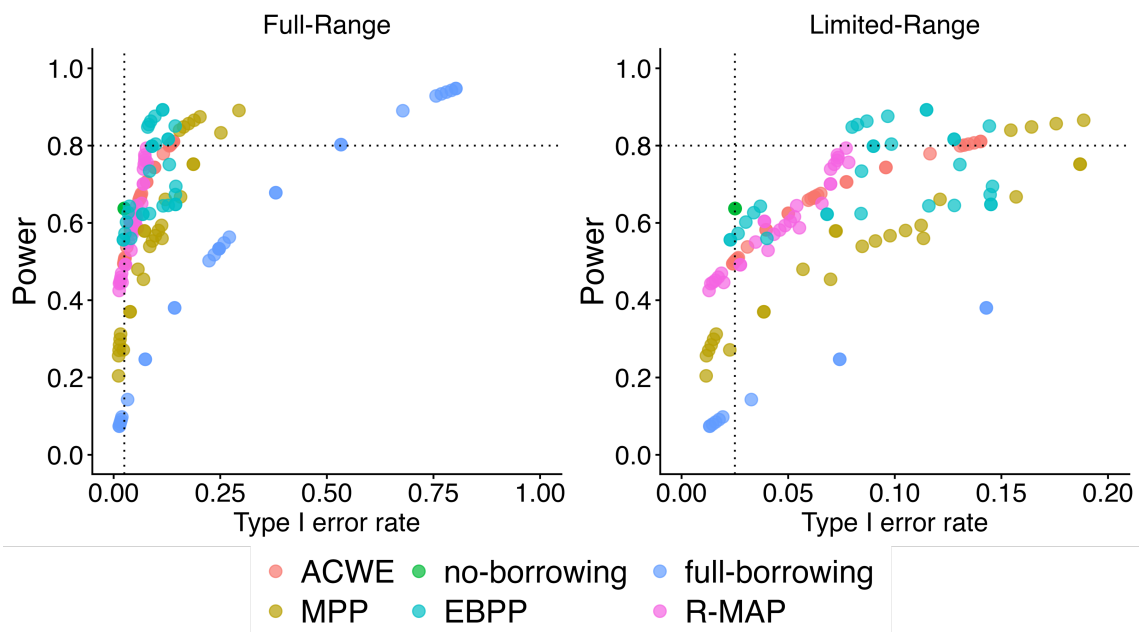


図 8. 検出力と第一種の過誤確率の散布図（左：全範囲，右：一部範囲）

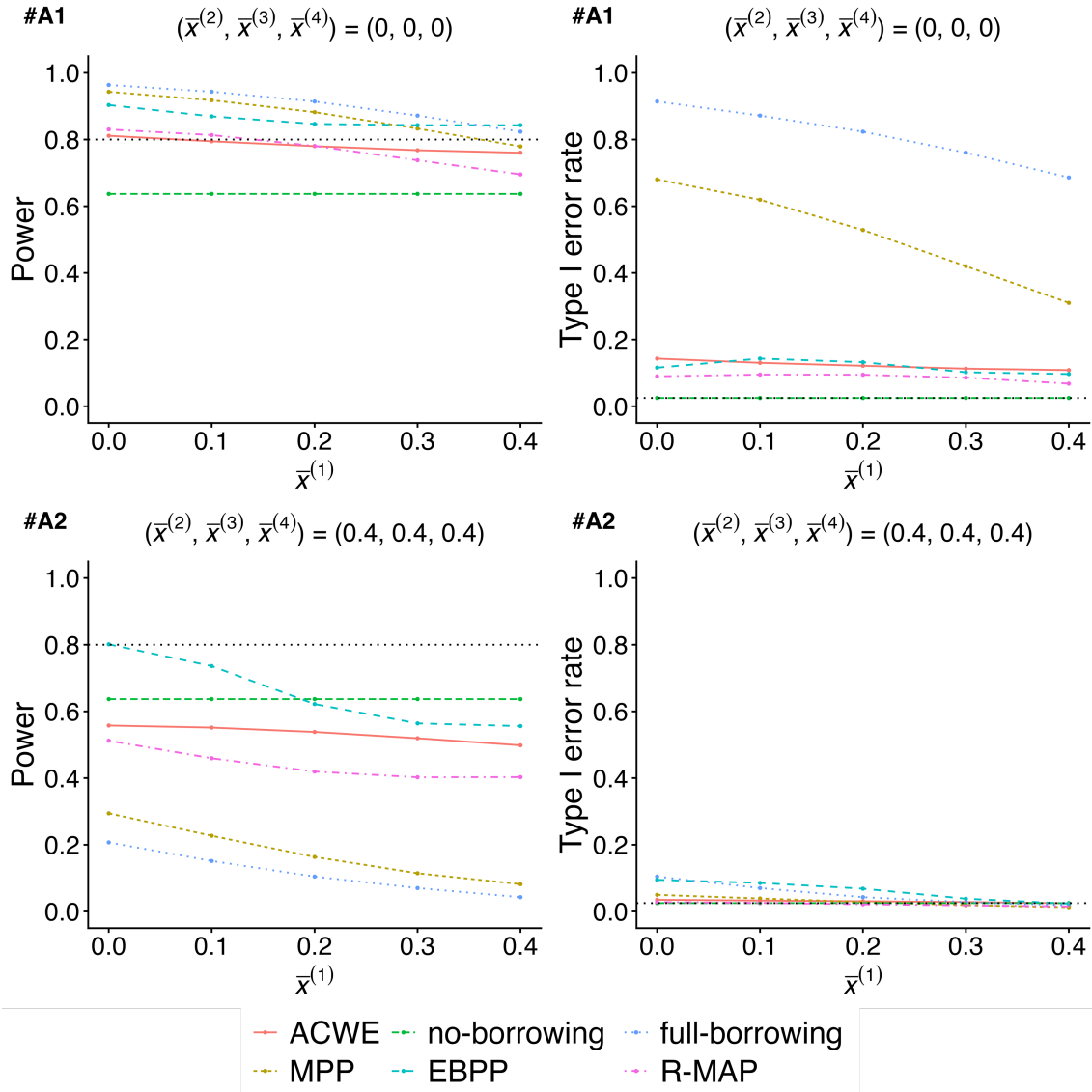


図 9. 検出力と第一種の過誤確率（追加検討； $S = 4$ ）.

3. 臨床試験データへの適用

各手法による $\Delta CDAI$ の平均値の群間差の点推定値と 95%CI を図 10 に示す. CI は頻度流の手法 (ACWE, 無借用, 全借用) では信頼区間, ベイズ法 (R-MAP, MPP, EBPP) では信用区間を表す. 表 6 に提案法と重み付け型の既存法である MPP と EBPP における各外部情報源への重みと重み付きサンプルサイズを示す. 付録図 7 に提案法, MPP, および EBPP の重みの 95%CI を示す.

図 10 から, 無借用と全借用の点推定値には 0.2 標準偏差ほどの差があり, 対照データ全体として中程度の異質性が存在することが示唆される. 外部情報からの借用の結果, 各適応的手法の点推定値は無借用と全借用との間に位置し, 区間幅も中間程度となった. 提案法と R-MAP の区間幅は MPP と EBPP の場合よ

りも大きく、より保守的な傾向が示唆される。提案法は対照群の目標サンプルサイズを 40 に制御しているためである。R-MAP では、外部データに基づかない曖昧な事前分布を混合させているためである。本試験例のように合計外部サンプルサイズが大きい場合、MPP のような既存法では、借用量を制限するためにはチューニングパラメータを極端な値に設定する必要があるようである。

表 6 から提案法の適応的な重み付けを確認することができる。例えば、当該試験と結果が最も類似している Sandborn et al. (2001b) には 0.0178, その次に類似している Hanauer et al. (2006) には 0.0174 が与えられ、続いて Sandborn et al. (2007a) には 0.0168, Sandborn et al. (2007b) には 0.0140 へと続く。そして最も類似していない Winter et al. (2004) には 0.0113 の重みが与えられている。一つだけ例外があり、Sandborn et al. (2001a) の点推定値は Hueber 試験の対照平均に最も近い値ではないものの、与えられた重みが 0.0189 と最も大きくなっている。これは同過去試験のサンプルサイズが 20 と小さいからである。方法 1.1 節で述べたように、ESS の制約条件下において、よりサイズの小さい外部情報源に大きな重みを割り当てることは尤度の観点から有利であるため、重みは標本平均だけでなくサンプルサイズによっても調整される。重み付きサンプルサイズにおいて、Sandborn et al. (2007b) および Sandborn et al. (2007a) はサイズが大きいため、借用に対する影響力がより大きい。ただし重み付きサンプルサイズは重み付き推定量の分散に基づく ESS とは異なるため、重み付きサンプルサイズの単純な合計は目標サンプルサイズ 40 とは一致しないことに留意する必要がある。方法 1.1 で述べたように、重みの推定誤差を考慮するためには(5)式の制約条件を φn^* に更新する必要がある。本試験例では $\varphi = 1.14162$ と計算され、したがって本来の目標サンプルサイズ 40 に対して解析上の目標サンプルサイズは $\varphi n^* = 45.665$ に更新する必要がある。提案法の重みの推定値を(5)式に代入すると同様に 45.665 が得られ、この値を達成していることが確認された。表 6 に示している重みは丸めているため一致しないが、これらを用いて(5)式を計算した場合でもほぼ同一の値である 45.72 が得られる。

表 6 に示すように、Winter et al. (2004) と Sandborn et al. (2007b) の標準化差の絶対値はほぼ同値であるものの、提案法は前者に対してより小さい重みを割り当てている。(1)式の最終項が示すように、提案法は重み付き推定量と各外部標本平均の差の目的関数に対する寄与を通じて重みを最適化する。本適用例では、Winter et al. を除く全ての研究が負の標準化差を有し、かつそれらのいくつかは互いに類似している。このような状況では、重み付き推定量は同時対照の標本平均から見ると負の方向に縮小しやすい。結果として、標準化差が正である唯一の外部情報である Winter et al. に大きな重みを割り当てると(1)式の最終項の観点で大きな損失が生じる。

また表 6 は，提案法が MPP や EBPP などの既存法よりも安定した重み付けを行う傾向があることを示している．これは結果 1 節で述べたように，重みの均一性が重み付き尤度の増加に寄与する要素の一つであるためである．MPP と EBPP，特に EBPP は，これまでの節でも示したように極端な重み付けを行う性質が確認された．

提案法の重み付けの安定した挙動は付録図 7 の信頼区間からも確認できる．MPP は全体的に区間幅が広く，同時対照との類似性が示唆される外部情報源の識別精度が低いことが示唆された．EBPP の重みはほぼ 0 または 1 の値しか取らず，区間幅は極端に広いか狭いかのいずれかであることから，妥当な区間推定とならないことが示唆された．

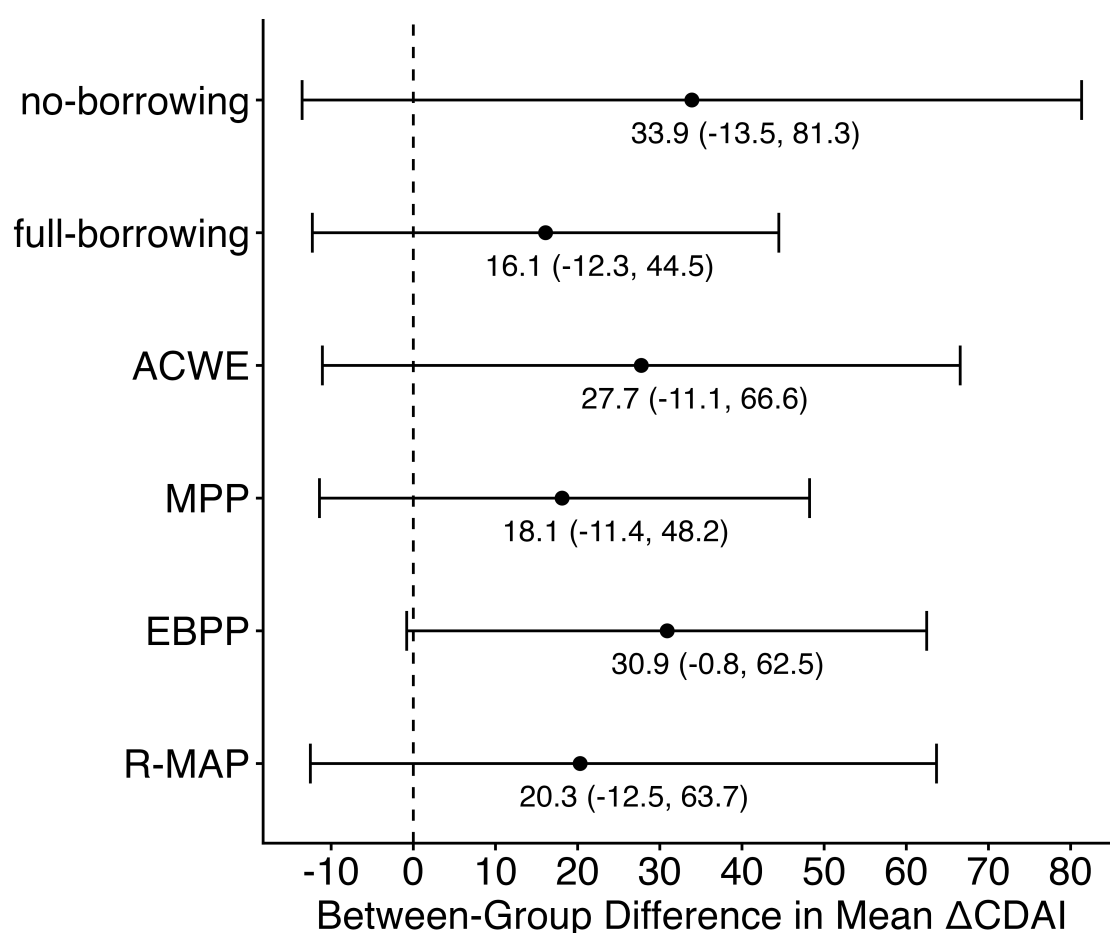


図 10. Δ CDAI の平均値の群間差.

表 6. 各過去試験への重みと重み付きサンプルサイズ.

Study	n	std	ACWE		MPP		EBPP	
			w	wn	w	wn	w	wn

S (2007a)	166	-0.160	0.0168	2.8	0.5184	86.1	0.0000	0.0
S (2007b)	328	-0.308	0.0140	4.6	0.4603	151.0	0.0000	0.0
S (2001a)	20	-0.183	0.0189	0.4	0.5002	10.0	0.0000	0.0
S (2001b)	58	-0.103	0.0178	1.0	0.5165	30.0	1.0000	58.0
H (2006)	74	-0.138	0.0174	1.3	0.5011	37.1	0.6400	47.4
W (2004)	25	0.306	0.0113	0.3	0.5160	12.9	1.0000	25.0

S: Sandborn et al., H: Hanauer et al., W: Winter et al., n: サンプルサイズ, std: Hueber
試験からの標準化差, w: 重み, wn: 重み付きサンプルサイズ.

考察

本研究では、適応的な重み付き最尤推定に ESS に関する制約条件を導入することで、第一種の過誤確率と検出力の名目水準への制御と適応的な情報借用を両立する新たな方法である ACWE を提案した。提案法は頻度論に基づいており、数学的に単純であり、数値例、シミュレーション、臨床試験データへの適用により良好な動作特性を持つことが示された。

提案法は外部情報源から得られる尤度を割り引く形式をとっており、“Equal but discounted”モデル (Neuenschwander et al., 2010; Spiegelhalter et al., 2004) に相当することから、power prior との共通点を有する。特に EBPP は、チューニングパラメータを指定する必要がなく、最適化によって重みを推定する手法であるため提案法との共通点が多い。一方、(1)式と(5)式で示している提案法の目的関数と制約条件、そして方法 1.2.2.2 節で示した EBPP の目的関数からも確認できるように、これらの手法は明確に異なる。提案法は、頻度論に基づく制約付き最適化問題として定式化した点に新規性がある。

情報借用法としての提案法の特長は、第一種の過誤確率と検出力をデザインによって制御できることである。これには同時対照と外部情報の同質性を仮定する必要がある。動作特性を一般に制御できるわけではない。しかし、特定の条件下で明示的に制御できることは試験計画の観点から望ましい。また、現実的な応用では異質な対照情報源のもとでの性能や挙動の安定性も同時に考慮する必要がある。要約すると、提案法は次の二つの観点から設計されている。

- (1) 同質な対照情報源のもとでは、第一種の過誤確率と検出力を名目水準に制御する
- (2) 異質な対照情報源のもとでは、適応的な重み付けを通じて第一種の過誤確率の増加を抑制しつつ、検出力を上昇させる

この整理された設計は既存法には見られない。提案法は制約付き借用を行うため、結果 2 節で示したように、同時対照パラメータと整合する外部情報が利用可能な場合でも、必ずしも検出力の大幅な向上を達成するわけではない。しかし本稿を通じて示したとおり、提案法は設計の一貫性に重点を置いている。

既存の適応的借用法は、対照情報源同士の同質性に基づいて借用量を適応させるアプローチを採用しているため、第一種の過誤確率と検出力を明示的に制御することができない。これは試験計画の観点では望ましくないため、これらの重要な計画事項をデザインによって保証しながら、適応的情報借用の利点（すな

わち情報源間の同質性の解析への反映)を引き出すためにはどうしたらよいか、という問いが本研究の起点となった。

その結果, ESS を一定水準に制約付けるという, 既存法とは逆の発想につながった。緒言 5 節で述べたように, 標準的な RCT で設計されるサンプルサイズは第一種の過誤確率と検出力の名目水準に基づいている。そのため, ESS がこのサンプルサイズに一致するように制約付けることで, 同質な対照情報源のもとでは第一種の過誤確率と検出力は名目水準に制御されると考えた。そしてその制約のもとで重み付き最尤推定を設計すれば, 適応的な情報借用を制約付き最適化問題として頻度論の枠組みで定式化できると考えた。これまで示してきたように, 提案法は標準的な RCT の計画事項 (有意水準, 検出力, サンプルサイズ) に基づいて設計したものであるため, ハイブリッド対照試験を標準的な RCT の自然な拡張として位置付けることに寄与するかもしれない。

ESS を一定水準に制約付けるという設計は, すべての外部情報源が同時対照と異なる場合にも借用が行われることを意味し, 実際に数値例からも確認された。これは必ずしも望ましい性質ではないが, 結果 1 節でも述べたように, 同時対照と整合しない外部情報源に対する重みが 0 になるとは限らないという問題は既存法でも見られ, 適応的な方法でこれを完全に防ぐことは難しいと考えられる。現実的に重要なことは, 各手法が様々な種類の異質性のもとでどのような挙動を示すかを事前に理解していることである。異質な対照情報源のもとの提案法の性能として, シミュレーションでは下記が観察された。

- 一部の外部情報源のみが同時対照と同質である場合は, EBPP のような既存法よりも性能が低い傾向があった (検出力がより低く, 第一種の過誤確率はさほど変わらない)
- 同時対照と外部情報の乖離が大きいとも小さいとも言えない曖昧な状況では, 既存法よりも性能が高い傾向があった (検出力がより高く, 第一種の過誤確率もより低い)

方法 1.1 節で説明したように, 重みの推定誤差の存在下でも目標サンプルサイズ n^* に相当する推定精度を達成するには, 対照群全体で φn^* のサンプルサイズが必要となる。実用上は, 補正係数 φ は同時対照群の要約統計量の見積もり値に基づいて計算する必要がある。対照群全体の合計サンプルサイズが十分であることを事前に検討する必要がある。これは提案法に固有の要件である。この事前検討では図 3 のようなプロットを作成することが有用である。

本研究のシミュレーションでは, 特定の外部情報で条件付けたもとの動作特性を評価した。この理由は, 外部情報は試験の計画段階で通常既知であるとい

う意味で、条件付き推論の方が現実を反映するためである。特定の外部情報を想定しない無条件の性能に関心が当たる場合も考えられるが、方法 1.1 節で述べたように、提案法は外部情報の条件付けの有無に関わらず、必要な仮定のもとでは同一の検定統計量を導く。したがって、無条件の外部情報を考慮した場合でも提案法の重要な設計（同質な対照情報源のもとでの検出力と第一種の過誤確率の明示的な制御）は保たれる。

また、シミュレーションでは主に 2 つの外部情報源のみを考慮した。4 つの外部情報源を用いた追加シミュレーションによって、提案法の重要な性質は外部情報源の数に依存しないことは確認できたが、より多くの異質な外部情報源のもとでの性能評価は今後の課題である。

今後のさらなる研究として、提案法の非正規アウトカムへの適用と共変量調整を検討している。提案法は尤度に基づく手法であるため、分布が指定できるのであれば非正規アウトカムにも適用できると考えられる。また、情報源間の患者背景の違いは同質性の評価に直結する重要な要素であるため、解析によって調整できることが望ましい。患者背景の違いを調整するためにはすべての情報源において個人レベルのデータが利用可能である必要があるが、回帰モデルまたは傾向スコアによる交絡調整と重み付き対数尤度に基づく重み推定を組み合わせることは想定できる。その他、提案法を実行する R プログラムは開発済みであるが、実務家の視点を考慮すると、パッケージまたはアプリケーションとして利用可能である方が望ましい。この開発も今後の課題の一つである。

結論

本研究では、臨床試験において複数の外部情報源を利用する新たな適応的情報借用法である ACWE を開発した。本研究から得られた知見は以下のとおりである。

- 適応的な重み付き最尤推定に ESS を一定に保つ制約条件を導入することで、頻度論の枠組みのもとで、対照情報源が同質である場合の第一種の過誤確率と検出力の名目水準への制御と適応的な借用を両立する
- 重みの推定誤差による ESS の過小評価は、ブートストラップ分散を用いた制約条件の再定義によって解決する
- モンテカルロシミュレーションに基づき、提案法は設計どおりの統計的性能、すなわち同質な対照情報源における第一種の過誤確率と検出力の名目水準への制御を達成することが示された
 - 各対照情報源が同質でない場合は、適応的な重み付けを通じて第一種の過誤確率の増加を抑制しつつ、検出力を上昇させる傾向が確認された
- 数値例では、同時対照との乖離が著しく大きい外部情報源に対する重みは 0 となり、適応的な情報借用法として望ましい性質を有することが示された

適応的な情報借用は、情報源間の同質性を解析に動的に反映することができるという意味で望ましい。一方、既存法は情報源同士の同質性に基づいて借用量を適応させるアプローチを採用しているため、第一種の過誤確率と検出力を明示的に制御することができず、これは試験計画の観点では望ましくない。

本研究の意義は、この問題点を解決するための手法を数理的に設計したことである。提案法は、興味のあるパラメータの ESS を一定に制御する制約のもとで適応的な重み付き推定を行うことで、同質な対照情報源のもとでは第一種の過誤確率と検出力が名目水準に制御され、異質性のある状況では適応的な重み付けに基づく借用が行われる。この体系的な設計、および ESS を一定水準に制約付けるという既存法とは逆の発想には新規性がある。また、情報借用を頻度論に基づく制約付き最適化問題として定式化した点にも新規性がある。

提案法の根幹的な性質はシミュレーションによって数値的にも確認されたが、より多くの異質な外部情報源のもとでの性能評価は今後の課題である。また、提案法の非正規アウトカムへの適用と共変量調整についても確立する必要がある。

本研究で開発した手法は標準的な RCT の計画事項（有意水準、検出力、サンプルサイズ）に基づいて設計したものであるため、ハイブリッド対照試験を標準

的な臨床試験の自然な拡張として位置付けることに寄与するかもしれない。これによって、情報借用の適用対象となりやすい希少な疾患領域を対象とした研究の発展につながることを期待される。

謝辞

本研究のご指導ならびにご激励を賜りました北海道大学大学院医学研究院医学統計学教室の横田勲准教授に、心より感謝の意を表します。

北海道大学病院医療・ヘルスサイエンス研究開発機構の佐藤典宏機構長には、生物統計担当者としての勤務と博士課程在籍の両立をご快諾いただき、研究遂行の機会をご提供くださいましたことに、謹んで感謝申し上げます。

滋賀大学データサイエンス・AI イノベーション研究推進センターの佐藤俊哉特任教授（京都大学大学院医学研究科社会健康医学系専攻医療統計学 名誉教授）には、専門職学位課程在学中に多大なご指導を賜り、修了後も変わらぬご助言とお励ましをいただきました。深く御礼申し上げます。

京都大学大学院医学研究科社会健康医学系専攻臨床統計学の田中司朗特定教授には、専門職学位課程では指導教員としてご指導を賜り、博士課程では共同研究者として多くの有益な示唆を賜りました。心より感謝申し上げます。

医薬品医療機器総合機構新薬審査第五部の岡田和史氏は、身近な研究仲間として度重なる助言と建設的な議論をいただきました。本研究の着想を最初に持ちかけた際の前向きな議論が、今日の成果へとつながっております。ここに感謝の意を表します。

本研究は、直接的な研究活動のみならず、日常的な管理業務や環境整備をご支援くださった方々にも支えられてきました。北海道大学大学院医学研究院医学統計学教室の平馬永恵氏ならびに日當裕子氏に深く感謝いたします。

最後に、研究生活を間近で支えてくれた妻と娘に心より感謝いたします。

利益相反

以下の研究費の補助を受けている。

Japan Agency for Medical Research and Development. JP21lk0201702 JSPS KAKENHI.
21K17712

引用文献

- Baeten D, Baraliakos X, Braun J, Sieper J, Emery P, van der Heijde D, McInnes I, van Laar JM, Landewé R, Wordsworth P, et al (2013) Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. *Lancet* 382, 1705–1713.
- Banbeta A, van Rosmalen J, Dejardin D and Lesaffre E (2019) Modified power prior with multiple historical trials for binary endpoints. *Stat Med* 38, 1147–1169.
- Banbeta A, Lesaffre E and van Rosmalen J (2022) The power prior with multiple historical controls for the linear regression model. *Pharm Stat* 21, 418–438.
- Bodey GP, Fainstein V, Elting LS, Anaissie E, Rolston K, Khardori N, Kantarjian H, Plager C, Murphy WK, Holmes F, et al (1990) Beta-lactam regimens for the febrile neutropenic patient. *Cancer* 65, 9–16.
- Burger HU, Gerlinger C, Harbron C, Koch A, Posch M, Rochon J and Schiel A (2021) The use of external controls: To what extent can it currently be recommended? *Pharm Stat* 20, 1002–1016.
- Burns PB, Rohrich RJ and Chung KC (2011) The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg* 128, 305–310.
- Chen WC, Wang C, Li H, Lu N, Tiwari R, Xu Y and Yue LQ (2020) Propensity score-integrated composite likelihood approach for augmenting the control arm of a randomized controlled trial by incorporating real-world data. *J Biopharm Stat* 30, 508–520.
- Duan Y, Ye K and Smith EP (2006) Evaluating water quality using power priors to incorporate historical information. *Environmetrics* 17, 95–106.
- Food and Drug Administration (2010) Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials. <https://www.fda.gov/media/71512/download>. Accessed January 6, 2025.
- Food and Drug Administration (2020) Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products. <https://www.fda.gov/media/130897/download>. Accessed January 6, 2025.
- Food and Drug Administration (2023) Rare Diseases: Considerations for the Development of Drugs and Biological Products Guidance for Industry. <https://www.fda.gov/media/119757/download>. Accessed January 6, 2025.
- Ghadessi M, Tang R, Zhou J, Liu R, Wang C, Toyozumi K, Mei C, Zhang L, Deng C Q and Beckman RA (2020) A roadmap to using historical controls in clinical trials - by

- Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG). *Orphanet J Rare Dis* 15, 69.
- Gravestock I and Held L (2017) Adaptive power priors with empirical Bayes for clinical trials. *Pharm Stat* 16, 349–360.
- Gravestock I and Held L (2019) Power priors based on multiple historical studies for binary outcomes. *Biom J* 61, 1201–1218.
- Gsteiger S, Neuenschwander B, Mercier F and Schmidli H (2013) Using historical control information for the design and analysis of clinical trials with overdispersed count data. *Stat Med* 32, 3609–3622.
- Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P and Schünemann HJ (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336, 924–926.
- Hanauer SB, Sandborn WJ, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh D, Panaccione R, Wolf D and Pollack P (2006) Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial. *Gastroenterology* 130, 323–333.
- Holmes DR Jr, Teirstein P, Satler L, Sketch M, O'Malley J, Popma JJ, Kuntz RE, Fitzgerald PJ, Wang H, Caramanica E, et al (2006) Sirolimus-eluting stents vs vascular brachytherapy for in-stent restenosis within bare-metal stents: the SISR randomized trial. *JAMA* 295, 1264–1273.
- Hueber W, Sands BE, Lewitzky S, Vandemeulebroecke M, Reinisch W, Higgins PD, Wehkamp J, Feagan BG, Yao MD, Karczewski M, et al (2012) Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* 61, 1693–1700.
- Ibrahim JG and Chen MH (2000) Power prior distributions for regression models. *Stat Sci* 15, 46–60.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (1998) Statistical principles for clinical trials: E9. https://database.ich.org/sites/default/files/E9_Guideline.pdf. Accessed January 6, 2025.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (2000) Choice of control group and related issues in clinical trials: E10. https://database.ich.org/sites/default/files/E10_Guideline.pdf. Accessed January 6, 2025.
- Jahanshahi M, Gregg K, Davis G, Ndu A, Miller V, Vockley J, Ollivier C, Franolic T and Sakai S (2021) The use of external controls in FDA regulatory decision making. *Ther Innov Regul Sci* 55, 1019–1035.

- Kish L (1965) Survey Sampling (John Wiley & Sons: New York, Chichester, Brisbane, Toronto)
- Li R, Lin R, Huang J, Tian L and Zhu J (2023) A frequentist approach to dynamic borrowing. *Biom J* 65, e2100406.
- Li W, Liu F and Snavely D (2020) Revisit of test-then-pool methods and some practical considerations. *Pharm Stat* 19, 498–517.
- Lim J, Wang L, Best N, Liu J, Yuan J, Yong F, Zhang L, Walley R, Gosselin A, Roebeling R, et al (2020) Reducing Patient Burden in Clinical Trials Through the Use of Historical Controls: Appropriate Selection of Historical Data to Minimize Risk of Bias. *Ther Innov Regul Sci* 54, 850-860.
- Neuenschwander B, Branson M and Spiegelhalter DJ (2009) A note on the power prior. *Stat Med* 28, 3562–3566.
- Neuenschwander B, Capkun-Niggli G, Branson M and Spiegelhalter DJ (2010) Summarizing historical information on controls in clinical trials. *Clin Trials* 7, 5–18.
- Okada K, Tanaka S, Matsubayashi J, Takahashi K and Yokota I (2024) Decoupling power and type I error rate considerations when incorporating historical control data using a test-then-pool approach. *Biom J* 66, e2200312.
- Pang H and Zhu J (2024) A Case Study of a Hybrid Control Design in Diffuse B-Cell Lymphoma. In *Advancing the Use of Complex Innovative Designs in Clinical Trials Workshop* (Food and Drug Administration, host), <https://www.fda.gov/media/182238/download?attachment>. Accessed January 6, 2025.
- Pennello G and Thompson L (2008) Experience with reviewing Bayesian medical device trials. *J Biopharm Stat* 18, 81-115.
- Pocock SJ (1976) The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases* 29, 175-188.
- Salloway S, Farlow M, McDade E, Clifford DB, Wang G, Llibre-Guerra JJ, Hitchcock JM, Mills SL, Santacruz AM, Aschenbrenner AJ, et al (2021) A trial of gantenerumab or solanezumab in dominantly inherited Alzheimer's disease. *Nat Med* 27, 1187–1196.
- Sandborn WJ, Hanauer SB, Katz S, Safdi M, Wolf DG, Baerg RD, Tremaine WJ, Johnson T, Diehl NN and Zinsmeister AR (2001a) Etanercept for active Crohn's disease: a randomized, double-blind, placebo-controlled trial. *Gastroenterology* 121, 1088–1094.
- Sandborn WJ, Feagan BG, Hanauer SB, Present DH, Sutherland LR, Kamm MA, Wolf DC, Baker JP, Hawkey C, Archambault A, et al (2001b) An engineered human antibody to TNF (CDP571) for active Crohn's disease: a randomized double-blind placebo-controlled trial. *Gastroenterology* 120, 1330–1338.
- Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Colombel JF, Panaccione R, D'Haens G,

- Li J, Rosenfeld MR, Kent JD, et al (2007a) Adalimumab induction therapy for Crohn disease previously treated with infliximab: a randomized trial. *Ann Intern Med* 146, 829–838.
- Sandborn WJ, Feagan BG, Stoinov S, Honiball PJ, Rutgeerts P, Mason D, Bloomfield R and Schreiber S (2007b) Certolizumab pegol for the treatment of Crohn's disease. *N Engl J Med* 357, 228–238.
- Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D and Neuenschwander B (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 70, 1023–1032.
- Spiegelhalter DJ, Abrams KR and Myles JP (2004) Bayesian approaches to clinical trials and health-care evaluation (John Wiley & Sons: Chichester)
- Viele K, Berry S, Neuenschwander B, Amzal B, Chen F, Enas N, Hobbs B, Ibrahim JG, Kinnersley N, Lindborg S, et al (2014) Use of historical control data for assessing treatment effects in clinical trials. *Pharm Stat* 13, 41–54.
- Weber S, Li Y, Seaman JW III, Kakizume T and Schmidli H (2021) Applying meta-analytic-predictive priors with the R Bayesian evidence synthesis tools. *J Stat Softw* 100, 1–32.
- Winter TA, Wright J, Ghosh S, Jahnsen J, Innes A and Round P (2004) Intravenous CDP870, a PEGylated Fab' fragment of a humanized antitumour necrosis factor antibody, in patients with moderate-to-severe Crohn's disease: an exploratory study. *Aliment Pharmacol Ther* 20, 1337–1346.
- Yamaue H, Shimizu A, Hagiwara Y, Sho M, Yanagimoto H, Nakamori S, Ueno H, Ishii H, Kitano M, Sugimori K, et al (2017) Multicenter, randomized, open-label Phase II study comparing S-1 alternate-day oral therapy with the standard daily regimen as a first-line treatment in patients with unresectable advanced pancreatic cancer. *Cancer Chemother Pharmacol* 79, 813–823.
- Yin G (2014) 適応的ランダム化. 臨床試験デザイン: ベイズ流・頻度流の適応的方法 (初版), 手良向聡, 大門貴志訳 (メディカル・パブリケーションズ: 東京) 225–227.
- 武田健太郎, 大庭真梨, 柿爪智行, 坂巻顕太郎, 田栗正隆, 森田智視 (2015) 臨床試験におけるヒストリカルコントロールデータの利用. *計量生物学* 36, 25–50.
- 日本製薬工業協会 (2022) Rare disease の治療効果の推測法. https://www.jpma.or.jp/information/evaluation/results/allotment/jtrngf000000085m-att/ds_202212_rare.pdf. Accessed January 6, 2025.
- 日本製薬工業協会 (2022) 薬事申請に Real World Data を外部対照として利用する場合の留意点 .

https://www.jpma.or.jp/information/evaluation/results/allotment/gbkspa0000000mw7-att/DS_202207_RWD.pdf. Accessed January 6, 2025.

野村尚吾, 大東智洋, 澤本涼 (2022) Hybrid control アプローチを用いるランダム化比較試験の計画と解析: 外部データが要約統計量の場合. 計量生物学 43, 63-96.

付録

(7)式の導出

θ_c の初期事前分布を $N(0, \eta^2)$ とすると,

$$\begin{aligned} g(\boldsymbol{\omega}) &= \int_{-\infty}^{\infty} \left\{ \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} \right\} p(\theta_c) d\theta_c \\ &= \left\{ \prod_{s=1}^S \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{\omega^{(s)} n_s} \right\} \frac{1}{\sqrt{2\pi\eta^2}} \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2\sigma^2} \sum_{s=1}^S \left\{ \omega^{(s)} \sum_{i=1}^{n_s} (X_i^{(s)} \right. \right. \\ &\quad \left. \left. - \theta_c \right)^2 \right\} - \frac{\theta_c^2}{2\eta^2} \right) d\theta_c \end{aligned}$$

と表される．指数部を U とおくと,

$$\begin{aligned} U &= -\frac{1}{2\sigma^2} \sum_{s=1}^S \omega^{(s)} \left(\sum_{i=1}^{n_s} X_i^{(s)2} \right) + \frac{\theta_c}{\sigma^2} \sum_{s=1}^S \omega^{(s)} \left(\sum_{i=1}^{n_s} X_i^{(s)} \right) \\ &\quad - \frac{\theta_c^2}{2} \left(\frac{1}{\sigma^2} \sum_{s=1}^S \omega^{(s)} n_s + \frac{1}{\eta^2} \right) \end{aligned}$$

と変形でき,

$$\begin{aligned} A &= \frac{1}{\sigma^2} \sum_{s=1}^S \omega^{(s)} n_s + \frac{1}{\eta^2}, \\ B &= \frac{1}{\sigma^2} \sum_{s=1}^S \omega^{(s)} \sum_{i=1}^{n_s} X_i^{(s)} \\ &= \frac{1}{\sigma^2} \sum_{s=1}^S \omega^{(s)} n_s \bar{X}^{(s)}, \\ C &= \frac{1}{2\sigma^2} \sum_{s=1}^S \omega^{(s)} \sum_{i=1}^{n_s} X_i^{(s)2} \\ &= \frac{1}{2\sigma^2} \sum_{s=1}^S \omega^{(s)} \left\{ n_s \bar{X}^{(s)2} + (n_s - 1) \hat{\sigma}^{2(s)} \right\} \end{aligned}$$

とおくと, 平方完成により

$$\begin{aligned} U &= -\frac{\theta_c^2}{2} A + \theta_c B - C \\ &= -\frac{A}{2} \left(\theta_c - \frac{B}{A} \right)^2 + \frac{B^2}{2A} - C \end{aligned}$$

となる．

$$\exp(U) = \exp \left(-\frac{A}{2} \left(\theta_c - \frac{B}{A} \right)^2 \right) \exp \left(\frac{B^2}{2A} - C \right)$$

から,

$$g(\boldsymbol{\omega}) = \left\{ \prod_{s=1}^S \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{\omega^{(s)}n_s} \right\} \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(\frac{B^2}{2A} - C\right) \int_{-\infty}^{\infty} \exp\left(-\frac{A}{2}\left(\theta_c - \frac{B}{A}\right)^2\right) d\theta_c$$

が得られる. ここでガウス積分により

$$\int_{-\infty}^{\infty} \exp\left(-\frac{A}{2}\left(\theta_c - \frac{B}{A}\right)^2\right) d\theta_c = \sqrt{\frac{2\pi}{A}}$$

であるため, (7)式は

$$g(\boldsymbol{\omega}) = \left\{ \prod_{s=1}^S \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{\omega^{(s)}n_s} \right\} \frac{1}{\sqrt{A\eta^2}} \exp\left(\frac{B^2}{2A} - C\right)$$

と導出できる.

(8)式の導出

θ_c の初期事前分布を $N(0, \eta^2)$ とすると, (8)式の分子は

$$\begin{aligned} h(\boldsymbol{\omega}) &= \int_{-\infty}^{\infty} L(\theta_c | \mathbf{X}^{(0)}) \prod_{s=1}^S L(\theta_c | \mathbf{X}^{(s)})^{\omega^{(s)}} d\theta_c \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n_0} \left\{ \prod_{s=1}^S \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{\omega^{(s)}n_s} \right\} \frac{1}{\sqrt{A_+\eta^2}} \exp\left(\frac{B_+^2}{2A_+} - C_+\right) \end{aligned}$$

と表される. ここでは(7)式の導出にならい,

$$A_+ = \frac{1}{\sigma^2} \left(n_0 + \sum_{s=1}^S \omega^{(s)}n_s \right) + \frac{1}{\eta^2},$$

$$B_+ = \frac{1}{\sigma^2} \left(n_0 \bar{X}^{(0)} + \sum_{s=1}^S \omega^{(s)}n_s \bar{X}^{(s)} \right),$$

$$C_+ = \frac{1}{2\sigma^2} \left[n_0 \bar{X}^{(0)2} + (n_0 - 1) \hat{\sigma}^{2(0)} + \sum_{s=1}^S \omega^{(s)} \left\{ n_s \bar{X}^{(s)2} + (n_s - 1) \hat{\sigma}^{2(s)} \right\} \right]$$

と表記している. 上記で導出した $g(\boldsymbol{\omega})$ を用いると, (8)式は

$$\begin{aligned} L(\boldsymbol{\omega} | \mathbf{X}) &= \frac{h(\boldsymbol{\omega})}{g(\boldsymbol{\omega})} = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n_0} \sqrt{\frac{A}{A_+}} \exp\left(\frac{B_+^2}{2A_+} - \frac{B^2}{2A} - C_+ + C\right) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n_0} \sqrt{\frac{A}{A_+}} \exp\left(\frac{B_+^2}{2A_+} - \frac{B^2}{2A} - \frac{n_0 \bar{X}^{(0)2} + (n_0 - 1) \hat{\sigma}^{2(0)}}{2\sigma^2}\right) \end{aligned}$$

$$\propto \sqrt{\frac{A}{A_+}} \exp\left(\frac{B_+^2}{2A_+} - \frac{B^2}{2A}\right)$$

と得られる．ここで $\eta \gg 0$ とし，また Gravestock and Held (2019) の記法になら

い， $\rho^{(s)} = \frac{n_s}{\sigma^2}$ とすると

$$\begin{aligned} \frac{A}{A_+} &= \frac{\frac{1}{\sigma^2} \sum_{s=1}^S \omega^{(s)} n_s}{\frac{1}{\sigma^2} (n_0 + \sum_{s=1}^S \omega^{(s)} n_s)} = \frac{\sum_{s=1}^S \omega^{(s)} \rho^{(s)}}{\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}} \propto \frac{2\pi \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}{\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}, \\ \frac{B_+^2}{2A_+} - \frac{B^2}{2A} &= \frac{AB_+^2 - A_+B^2}{2AA_+} \\ &= -\frac{\rho^{(0)} \{\sum_{s=1}^S \omega^{(s)} \rho^{(s)} (\bar{X}^{(s)} - \bar{X}^{(0)})\}^2}{2(\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}) (\sum_{s=1}^S \omega^{(s)} \rho^{(s)})} + \frac{\rho^{(0)} \bar{X}^{(0)2}}{2} \\ &\propto -\frac{\rho^{(0)} \{\sum_{s=1}^S \omega^{(s)} \rho^{(s)} (\bar{X}^{(s)} - \bar{X}^{(0)})\}^2}{2(\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}) (\sum_{s=1}^S \omega^{(s)} \rho^{(s)})} \end{aligned}$$

より，(8)式は

$$L(\omega|X) \propto \sqrt{\frac{2\pi \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}{\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}}} \exp\left(-\frac{\rho^{(0)} \{\sum_{s=1}^S \omega^{(s)} \rho^{(s)} (\bar{X}^{(s)} - \bar{X}^{(0)})\}^2}{2(\rho^{(0)} + \sum_{s=1}^S \omega^{(s)} \rho^{(s)}) (\sum_{s=1}^S \omega^{(s)} \rho^{(s)})}\right)$$

と導出できる．

付録表 1. 第一種の過誤確率.

シナリオ	手法	外部対照平均	第一種の過誤確率
1	ACWE	(0, 0)	0.1403
1	ACWE	(0.1, 0)	0.11652
1	ACWE	(0.2, 0)	0.09579
1	ACWE	(0.3, 0)	0.07744
1	ACWE	(0.4, 0)	0.06248
1	no-borrowing	(0, 0)	0.02509
1	no-borrowing	(0.1, 0)	0.02509
1	no-borrowing	(0.2, 0)	0.02509
1	no-borrowing	(0.3, 0)	0.02509
1	no-borrowing	(0.4, 0)	0.02509
1	full-borrowing	(0, 0)	0.80252
1	full-borrowing	(0.1, 0)	0.67814
1	full-borrowing	(0.2, 0)	0.53317
1	full-borrowing	(0.3, 0)	0.38023
1	full-borrowing	(0.4, 0)	0.24714
1	R-MAP	(0, 0)	0.07715556
1	R-MAP	(0.1, 0)	0.07837778
1	R-MAP	(0.2, 0)	0.06985556
1	R-MAP	(0.3, 0)	0.05401111
1	R-MAP	(0.4, 0)	0.03887222
1	MPP	(0, 0)	0.29389
1	MPP	(0.1, 0)	0.2515
1	MPP	(0.2, 0)	0.18698
1	MPP	(0.3, 0)	0.12123
1	MPP	(0.4, 0)	0.07248
1	EBPP	(0, 0)	0.11492
1	EBPP	(0.1, 0)	0.14436
1	EBPP	(0.2, 0)	0.12784
1	EBPP	(0.3, 0)	0.09842
1	EBPP	(0.4, 0)	0.08997

2	ACWE	(0, 0.2)	0.09586
2	ACWE	(0.1, 0.2)	0.07726
2	ACWE	(0.2, 0.2)	0.06228
2	ACWE	(0.3, 0.2)	0.05005
2	ACWE	(0.4, 0.2)	0.03967
2	no-borrowing	(0, 0.2)	0.02509
2	no-borrowing	(0.1, 0.2)	0.02509
2	no-borrowing	(0.2, 0.2)	0.02509
2	no-borrowing	(0.3, 0.2)	0.02509
2	no-borrowing	(0.4, 0.2)	0.02509
2	full-borrowing	(0, 0.2)	0.53317
2	full-borrowing	(0.1, 0.2)	0.38023
2	full-borrowing	(0.2, 0.2)	0.24714
2	full-borrowing	(0.3, 0.2)	0.14294
2	full-borrowing	(0.4, 0.2)	0.07417
2	R-MAP	(0, 0.2)	0.06986111
2	R-MAP	(0.1, 0.2)	0.06582222
2	R-MAP	(0.2, 0.2)	0.05527778
2	R-MAP	(0.3, 0.2)	0.04056111
2	R-MAP	(0.4, 0.2)	0.02752222
2	MPP	(0, 0.2)	0.18685
2	MPP	(0.1, 0.2)	0.15692
2	MPP	(0.2, 0.2)	0.11344
2	MPP	(0.3, 0.2)	0.06988
2	MPP	(0.4, 0.2)	0.03854
2	EBPP	(0, 0.2)	0.12784
2	EBPP	(0.1, 0.2)	0.13053
2	EBPP	(0.2, 0.2)	0.1451
2	EBPP	(0.3, 0.2)	0.08418
2	EBPP	(0.4, 0.2)	0.06824
3	ACWE	(0, 0.4)	0.06249
3	ACWE	(0.1, 0.4)	0.05005

3	ACWE	(0.2, 0.4)	0.03966
3	ACWE	(0.3, 0.4)	0.0309
3	ACWE	(0.4, 0.4)	0.02399
3	no-borrowing	(0, 0.4)	0.02509
3	no-borrowing	(0.1, 0.4)	0.02509
3	no-borrowing	(0.2, 0.4)	0.02509
3	no-borrowing	(0.3, 0.4)	0.02509
3	no-borrowing	(0.4, 0.4)	0.02509
3	full-borrowing	(0, 0.4)	0.24714
3	full-borrowing	(0.1, 0.4)	0.14294
3	full-borrowing	(0.2, 0.4)	0.07417
3	full-borrowing	(0.3, 0.4)	0.03271
3	full-borrowing	(0.4, 0.4)	0.01325
3	R-MAP	(0, 0.4)	0.03883889
3	R-MAP	(0.1, 0.4)	0.03474444
3	R-MAP	(0.2, 0.4)	0.02757778
3	R-MAP	(0.3, 0.4)	0.01975556
3	R-MAP	(0.4, 0.4)	0.01288889
3	MPP	(0, 0.4)	0.07221
3	MPP	(0.1, 0.4)	0.05694
3	MPP	(0.2, 0.4)	0.03863
3	MPP	(0.3, 0.4)	0.02262
3	MPP	(0.4, 0.4)	0.01154
3	EBPP	(0, 0.4)	0.08997
3	EBPP	(0.1, 0.4)	0.08439
3	EBPP	(0.2, 0.4)	0.06824
3	EBPP	(0.3, 0.4)	0.04003
3	EBPP	(0.4, 0.4)	0.02288
4	ACWE	(0, 0)	0.14016
4	ACWE	(0.1, 0)	0.13721
4	ACWE	(0.2, 0)	0.1346
4	ACWE	(0.3, 0)	0.13247

4	ACWE	(0.4, 0)	0.1307
4	no-borrowing	(0, 0)	0.02509
4	no-borrowing	(0.1, 0)	0.02509
4	no-borrowing	(0.2, 0)	0.02509
4	no-borrowing	(0.3, 0)	0.02509
4	no-borrowing	(0.4, 0)	0.02509
4	full-borrowing	(0, 0)	0.80252
4	full-borrowing	(0.1, 0)	0.79204
4	full-borrowing	(0.2, 0)	0.78039
4	full-borrowing	(0.3, 0)	0.7685
4	full-borrowing	(0.4, 0)	0.75612
4	R-MAP	(0, 0)	0.07308889
4	R-MAP	(0.1, 0)	0.07340556
4	R-MAP	(0.2, 0)	0.073
4	R-MAP	(0.3, 0)	0.07165556
4	R-MAP	(0.4, 0)	0.06991111
4	MPP	(0, 0)	0.20234
4	MPP	(0.1, 0)	0.1886
4	MPP	(0.2, 0)	0.17571
4	MPP	(0.3, 0)	0.16414
4	MPP	(0.4, 0)	0.15443
4	EBPP	(0, 0)	0.11485
4	EBPP	(0.1, 0)	0.09671
4	EBPP	(0.2, 0)	0.0869
4	EBPP	(0.3, 0)	0.08246
4	EBPP	(0.4, 0)	0.08003
5	ACWE	(0, 0.2)	0.06529
5	ACWE	(0.1, 0.2)	0.06381
5	ACWE	(0.2, 0.2)	0.06227
5	ACWE	(0.3, 0.2)	0.06075
5	ACWE	(0.4, 0.2)	0.05947
5	no-borrowing	(0, 0.2)	0.02509

5	no-borrowing	(0.1, 0.2)	0.02509
5	no-borrowing	(0.2, 0.2)	0.02509
5	no-borrowing	(0.3, 0.2)	0.02509
5	no-borrowing	(0.4, 0.2)	0.02509
5	full-borrowing	(0, 0.2)	0.27123
5	full-borrowing	(0.1, 0.2)	0.25895
5	full-borrowing	(0.2, 0.2)	0.24714
5	full-borrowing	(0.3, 0.2)	0.23597
5	full-borrowing	(0.4, 0.2)	0.22381
5	R-MAP	(0, 0.2)	0.05293889
5	R-MAP	(0.1, 0.2)	0.05096667
5	R-MAP	(0.2, 0.2)	0.04851111
5	R-MAP	(0.3, 0.2)	0.04606111
5	R-MAP	(0.4, 0.2)	0.04332222
5	MPP	(0, 0.2)	0.11236
5	MPP	(0.1, 0.2)	0.10504
5	MPP	(0.2, 0.2)	0.09779
5	MPP	(0.3, 0.2)	0.09092
5	MPP	(0.4, 0.2)	0.08472
5	EBPP	(0, 0.2)	0.14583
5	EBPP	(0.1, 0.2)	0.14472
5	EBPP	(0.2, 0.2)	0.1451
5	EBPP	(0.3, 0.2)	0.12794
5	EBPP	(0.4, 0.2)	0.11598
6	ACWE	(0, 0.4)	0.02663
6	ACWE	(0.1, 0.4)	0.02615
6	ACWE	(0.2, 0.4)	0.02546
6	ACWE	(0.3, 0.4)	0.02472
6	ACWE	(0.4, 0.4)	0.02414
6	no-borrowing	(0, 0.4)	0.02509
6	no-borrowing	(0.1, 0.4)	0.02509
6	no-borrowing	(0.2, 0.4)	0.02509

6	no-borrowing	(0.3, 0.4)	0.02509
6	no-borrowing	(0.4, 0.4)	0.02509
6	full-borrowing	(0, 0.4)	0.01932
6	full-borrowing	(0.1, 0.4)	0.01758
6	full-borrowing	(0.2, 0.4)	0.01597
6	full-borrowing	(0.3, 0.4)	0.01451
6	full-borrowing	(0.4, 0.4)	0.01325
6	R-MAP	(0, 0.4)	0.01867778
6	R-MAP	(0.1, 0.4)	0.01718889
6	R-MAP	(0.2, 0.4)	0.01583333
6	R-MAP	(0.3, 0.4)	0.01468333
6	R-MAP	(0.4, 0.4)	0.01371111
6	MPP	(0, 0.4)	0.01625
6	MPP	(0.1, 0.4)	0.01502
6	MPP	(0.2, 0.4)	0.01389
6	MPP	(0.3, 0.4)	0.01272
6	MPP	(0.4, 0.4)	0.01167
6	EBPP	(0, 0.4)	0.03696
6	EBPP	(0.1, 0.4)	0.03386
6	EBPP	(0.2, 0.4)	0.03014
6	EBPP	(0.3, 0.4)	0.02651
6	EBPP	(0.4, 0.4)	0.02287
A1	ACWE	(0, 0, 0, 0)	0.143
A1	ACWE	(0.1, 0, 0, 0)	0.1303
A1	ACWE	(0.2, 0, 0, 0)	0.1215
A1	ACWE	(0.3, 0, 0, 0)	0.1129
A1	ACWE	(0.4, 0, 0, 0)	0.1085
A1	no-borrowing	(0, 0, 0, 0)	0.025
A1	no-borrowing	(0.1, 0, 0, 0)	0.025
A1	no-borrowing	(0.2, 0, 0, 0)	0.025
A1	no-borrowing	(0.3, 0, 0, 0)	0.025
A1	no-borrowing	(0.4, 0, 0, 0)	0.025

A1	full-borrowing	(0, 0, 0, 0)	0.9141
A1	full-borrowing	(0.1, 0, 0, 0)	0.8718
A1	full-borrowing	(0.2, 0, 0, 0)	0.8235
A1	full-borrowing	(0.3, 0, 0, 0)	0.7606
A1	full-borrowing	(0.4, 0, 0, 0)	0.6862
A1	R-MAP	(0, 0, 0, 0)	0.0899
A1	R-MAP	(0.1, 0, 0, 0)	0.095
A1	R-MAP	(0.2, 0, 0, 0)	0.0946
A1	R-MAP	(0.3, 0, 0, 0)	0.0858
A1	R-MAP	(0.4, 0, 0, 0)	0.0677
A1	MPP	(0, 0, 0, 0)	0.6804
A1	MPP	(0.1, 0, 0, 0)	0.6194
A1	MPP	(0.2, 0, 0, 0)	0.5286
A1	MPP	(0.3, 0, 0, 0)	0.4201
A1	MPP	(0.4, 0, 0, 0)	0.3101
A1	EBPP	(0, 0, 0, 0)	0.1158
A1	EBPP	(0.1, 0, 0, 0)	0.1431
A1	EBPP	(0.2, 0, 0, 0)	0.1322
A1	EBPP	(0.3, 0, 0, 0)	0.102
A1	EBPP	(0.4, 0, 0, 0)	0.0967
A2	ACWE	(0, 0.4, 0.4, 0.4)	0.0349
A2	ACWE	(0.1, 0.4, 0.4, 0.4)	0.0326
A2	ACWE	(0.2, 0.4, 0.4, 0.4)	0.0306
A2	ACWE	(0.3, 0.4, 0.4, 0.4)	0.0282
A2	ACWE	(0.4, 0.4, 0.4, 0.4)	0.0244
A2	no-borrowing	(0, 0.4, 0.4, 0.4)	0.025
A2	no-borrowing	(0.1, 0.4, 0.4, 0.4)	0.025
A2	no-borrowing	(0.2, 0.4, 0.4, 0.4)	0.025
A2	no-borrowing	(0.3, 0.4, 0.4, 0.4)	0.025
A2	no-borrowing	(0.4, 0.4, 0.4, 0.4)	0.025
A2	full-borrowing	(0, 0.4, 0.4, 0.4)	0.1046
A2	full-borrowing	(0.1, 0.4, 0.4, 0.4)	0.07

A2	full-borrowing	(0.2, 0.4, 0.4, 0.4)	0.0428
A2	full-borrowing	(0.3, 0.4, 0.4, 0.4)	0.0281
A2	full-borrowing	(0.4, 0.4, 0.4, 0.4)	0.0182
A2	R-MAP	(0, 0.4, 0.4, 0.4)	0.028
A2	R-MAP	(0.1, 0.4, 0.4, 0.4)	0.0257
A2	R-MAP	(0.2, 0.4, 0.4, 0.4)	0.0215
A2	R-MAP	(0.3, 0.4, 0.4, 0.4)	0.0182
A2	R-MAP	(0.4, 0.4, 0.4, 0.4)	0.0151
A2	MPP	(0, 0.4, 0.4, 0.4)	0.0498
A2	MPP	(0.1, 0.4, 0.4, 0.4)	0.0389
A2	MPP	(0.2, 0.4, 0.4, 0.4)	0.0282
A2	MPP	(0.3, 0.4, 0.4, 0.4)	0.0189
A2	MPP	(0.4, 0.4, 0.4, 0.4)	0.0128
A2	EBPP	(0, 0.4, 0.4, 0.4)	0.0948
A2	EBPP	(0.1, 0.4, 0.4, 0.4)	0.0855
A2	EBPP	(0.2, 0.4, 0.4, 0.4)	0.0682
A2	EBPP	(0.3, 0.4, 0.4, 0.4)	0.0382
A2	EBPP	(0.4, 0.4, 0.4, 0.4)	0.0235

付録表 2. 検出力.

シナリオ	手法	外部対照平均	検出力
1	ACWE	(0, 0)	0.8108
1	ACWE	(0.1, 0)	0.7791
1	ACWE	(0.2, 0)	0.74349
1	ACWE	(0.3, 0)	0.70577
1	ACWE	(0.4, 0)	0.66476
1	no-borrowing	(0, 0)	0.63724
1	no-borrowing	(0.1, 0)	0.63724
1	no-borrowing	(0.2, 0)	0.63724
1	no-borrowing	(0.3, 0)	0.63724
1	no-borrowing	(0.4, 0)	0.63724
1	full-borrowing	(0, 0)	0.94767

1	full-borrowing	(0.1, 0)	0.89025
1	full-borrowing	(0.2, 0)	0.80252
1	full-borrowing	(0.3, 0)	0.67814
1	full-borrowing	(0.4, 0)	0.53317
1	R-MAP	(0, 0)	0.79349444
1	R-MAP	(0.1, 0)	0.75676111
1	R-MAP	(0.2, 0)	0.70072222
1	R-MAP	(0.3, 0)	0.64458889
1	R-MAP	(0.4, 0)	0.60384444
1	MPP	(0, 0)	0.89092
1	MPP	(0.1, 0)	0.83313
1	MPP	(0.2, 0)	0.75178
1	MPP	(0.3, 0)	0.66091
1	MPP	(0.4, 0)	0.57898
1	EBPP	(0, 0)	0.89257
1	EBPP	(0.1, 0)	0.85063
1	EBPP	(0.2, 0)	0.81651
1	EBPP	(0.3, 0)	0.80386
1	EBPP	(0.4, 0)	0.79856
2	ACWE	(0, 0.2)	0.74356
2	ACWE	(0.1, 0.2)	0.70625
2	ACWE	(0.2, 0.2)	0.66691
2	ACWE	(0.3, 0.2)	0.62514
2	ACWE	(0.4, 0.2)	0.58218
2	no-borrowing	(0, 0.2)	0.63724
2	no-borrowing	(0.1, 0.2)	0.63724
2	no-borrowing	(0.2, 0.2)	0.63724
2	no-borrowing	(0.3, 0.2)	0.63724
2	no-borrowing	(0.4, 0.2)	0.63724
2	full-borrowing	(0, 0.2)	0.80252
2	full-borrowing	(0.1, 0.2)	0.67814
2	full-borrowing	(0.2, 0.2)	0.53317

2	full-borrowing	(0.3, 0.2)	0.38023
2	full-borrowing	(0.4, 0.2)	0.24714
2	R-MAP	(0, 0.2)	0.70083889
2	R-MAP	(0.1, 0.2)	0.65046667
2	R-MAP	(0.2, 0.2)	0.58735556
2	R-MAP	(0.3, 0.2)	0.52916667
2	R-MAP	(0.4, 0.2)	0.4918
2	MPP	(0, 0.2)	0.75215
2	MPP	(0.1, 0.2)	0.66742
2	MPP	(0.2, 0.2)	0.56001
2	MPP	(0.3, 0.2)	0.45409
2	MPP	(0.4, 0.2)	0.36999
2	EBPP	(0, 0.2)	0.81651
2	EBPP	(0.1, 0.2)	0.75106
2	EBPP	(0.2, 0.2)	0.64793
2	EBPP	(0.3, 0.2)	0.62414
2	EBPP	(0.4, 0.2)	0.62248
3	ACWE	(0, 0.4)	0.66478
3	ACWE	(0.1, 0.4)	0.62402
3	ACWE	(0.2, 0.4)	0.58218
3	ACWE	(0.3, 0.4)	0.53785
3	ACWE	(0.4, 0.4)	0.49415
3	no-borrowing	(0, 0.4)	0.63724
3	no-borrowing	(0.1, 0.4)	0.63724
3	no-borrowing	(0.2, 0.4)	0.63724
3	no-borrowing	(0.3, 0.4)	0.63724
3	no-borrowing	(0.4, 0.4)	0.63724
3	full-borrowing	(0, 0.4)	0.53317
3	full-borrowing	(0.1, 0.4)	0.38023
3	full-borrowing	(0.2, 0.4)	0.24714
3	full-borrowing	(0.3, 0.4)	0.14294
3	full-borrowing	(0.4, 0.4)	0.07417

3	R-MAP	(0, 0.4)	0.60414444
3	R-MAP	(0.1, 0.4)	0.55051667
3	R-MAP	(0.2, 0.4)	0.49166667
3	R-MAP	(0.3, 0.4)	0.44597778
3	R-MAP	(0.4, 0.4)	0.42520556
3	MPP	(0, 0.4)	0.57931
3	MPP	(0.1, 0.4)	0.47999
3	MPP	(0.2, 0.4)	0.37011
3	MPP	(0.3, 0.4)	0.27153
3	MPP	(0.4, 0.4)	0.20425
3	EBPP	(0, 0.4)	0.79856
3	EBPP	(0.1, 0.4)	0.73393
3	EBPP	(0.2, 0.4)	0.62248
3	EBPP	(0.3, 0.4)	0.56
3	EBPP	(0.4, 0.4)	0.55636
4	ACWE	(0, 0)	0.81058
4	ACWE	(0.1, 0)	0.80708
4	ACWE	(0.2, 0)	0.80388
4	ACWE	(0.3, 0)	0.80126
4	ACWE	(0.4, 0)	0.7993
4	no-borrowing	(0, 0)	0.63724
4	no-borrowing	(0.1, 0)	0.63724
4	no-borrowing	(0.2, 0)	0.63724
4	no-borrowing	(0.3, 0)	0.63724
4	no-borrowing	(0.4, 0)	0.63724
4	full-borrowing	(0, 0)	0.94767
4	full-borrowing	(0.1, 0)	0.94293
4	full-borrowing	(0.2, 0)	0.93847
4	full-borrowing	(0.3, 0)	0.93378
4	full-borrowing	(0.4, 0)	0.92849
4	R-MAP	(0, 0)	0.77595
4	R-MAP	(0.1, 0)	0.7692

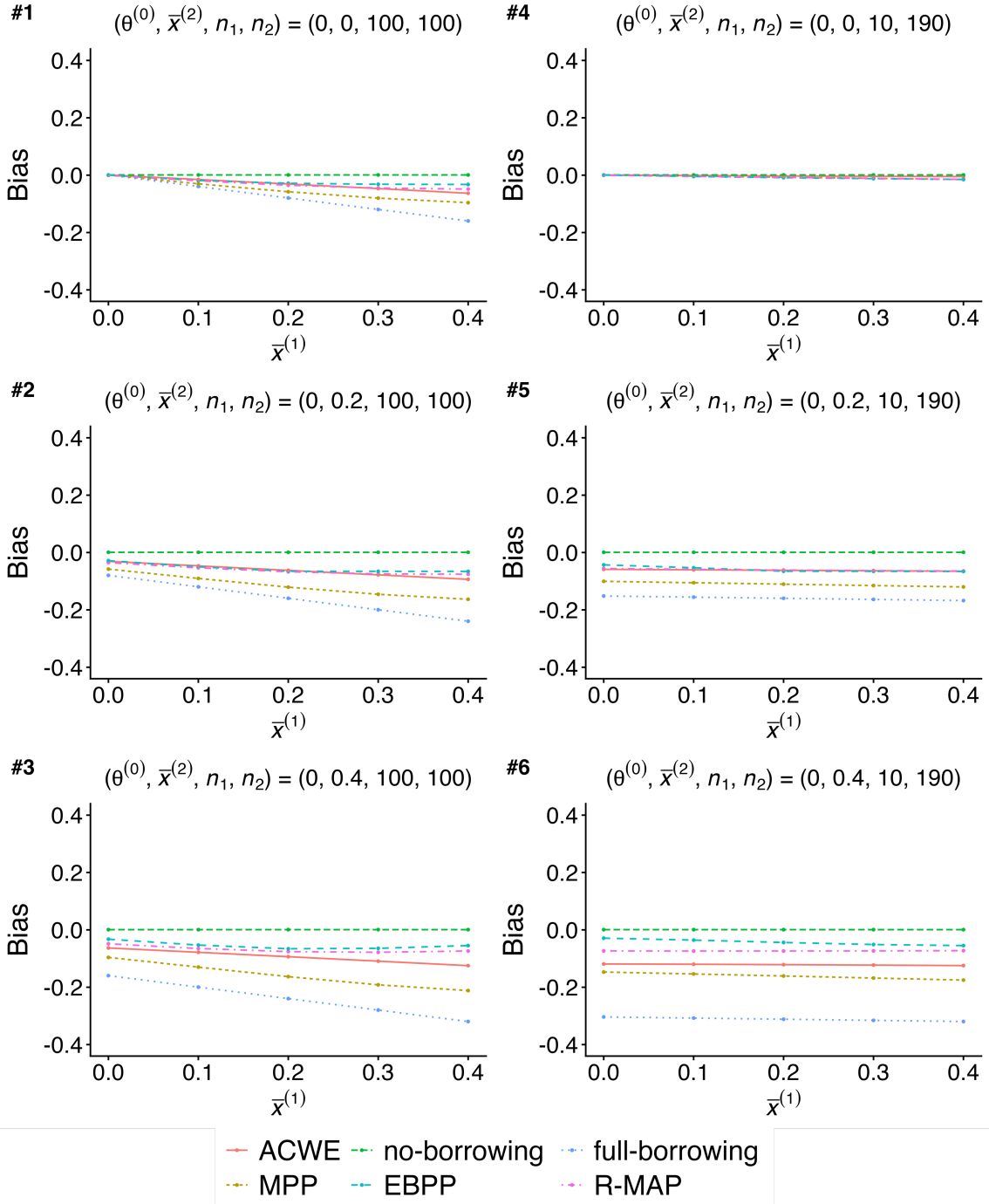
4	R-MAP	(0.2, 0)	0.76177222
4	R-MAP	(0.3, 0)	0.75168889
4	R-MAP	(0.4, 0)	0.73932222
4	MPP	(0, 0)	0.87445
4	MPP	(0.1, 0)	0.86563
4	MPP	(0.2, 0)	0.85682
4	MPP	(0.3, 0)	0.8483
4	MPP	(0.4, 0)	0.83993
4	EBPP	(0, 0)	0.89257
4	EBPP	(0.1, 0)	0.87593
4	EBPP	(0.2, 0)	0.86325
4	EBPP	(0.3, 0)	0.85431
4	EBPP	(0.4, 0)	0.84797
5	ACWE	(0, 0.2)	0.67609
5	ACWE	(0.1, 0.2)	0.67188
5	ACWE	(0.2, 0.2)	0.66733
5	ACWE	(0.3, 0.2)	0.66285
5	ACWE	(0.4, 0.2)	0.65879
5	no-borrowing	(0, 0.2)	0.63724
5	no-borrowing	(0.1, 0.2)	0.63724
5	no-borrowing	(0.2, 0.2)	0.63724
5	no-borrowing	(0.3, 0.2)	0.63724
5	no-borrowing	(0.4, 0.2)	0.63724
5	full-borrowing	(0, 0.2)	0.56313
5	full-borrowing	(0.1, 0.2)	0.54815
5	full-borrowing	(0.2, 0.2)	0.53317
5	full-borrowing	(0.3, 0.2)	0.51775
5	full-borrowing	(0.4, 0.2)	0.50227
5	R-MAP	(0, 0.2)	0.61598889
5	R-MAP	(0.1, 0.2)	0.60436667
5	R-MAP	(0.2, 0.2)	0.59312222
5	R-MAP	(0.3, 0.2)	0.5816

5	R-MAP	(0.4, 0.2)	0.57065
5	MPP	(0, 0.2)	0.5936
5	MPP	(0.1, 0.2)	0.58036
5	MPP	(0.2, 0.2)	0.5669
5	MPP	(0.3, 0.2)	0.55325
5	MPP	(0.4, 0.2)	0.53938
5	EBPP	(0, 0.2)	0.69454
5	EBPP	(0.1, 0.2)	0.67327
5	EBPP	(0.2, 0.2)	0.64791
5	EBPP	(0.3, 0.2)	0.64523
5	EBPP	(0.4, 0.2)	0.64416
6	ACWE	(0, 0.4)	0.51006
6	ACWE	(0.1, 0.4)	0.5075
6	ACWE	(0.2, 0.4)	0.50417
6	ACWE	(0.3, 0.4)	0.49978
6	ACWE	(0.4, 0.4)	0.49503
6	no-borrowing	(0, 0.4)	0.63724
6	no-borrowing	(0.1, 0.4)	0.63724
6	no-borrowing	(0.2, 0.4)	0.63724
6	no-borrowing	(0.3, 0.4)	0.63724
6	no-borrowing	(0.4, 0.4)	0.63724
6	full-borrowing	(0, 0.4)	0.09792
6	full-borrowing	(0.1, 0.4)	0.09118
6	full-borrowing	(0.2, 0.4)	0.08484
6	full-borrowing	(0.3, 0.4)	0.07929
6	full-borrowing	(0.4, 0.4)	0.07417
6	R-MAP	(0, 0.4)	0.46974444
6	R-MAP	(0.1, 0.4)	0.45947222
6	R-MAP	(0.2, 0.4)	0.45146111
6	R-MAP	(0.3, 0.4)	0.44627778
6	R-MAP	(0.4, 0.4)	0.44296111
6	MPP	(0, 0.4)	0.31228

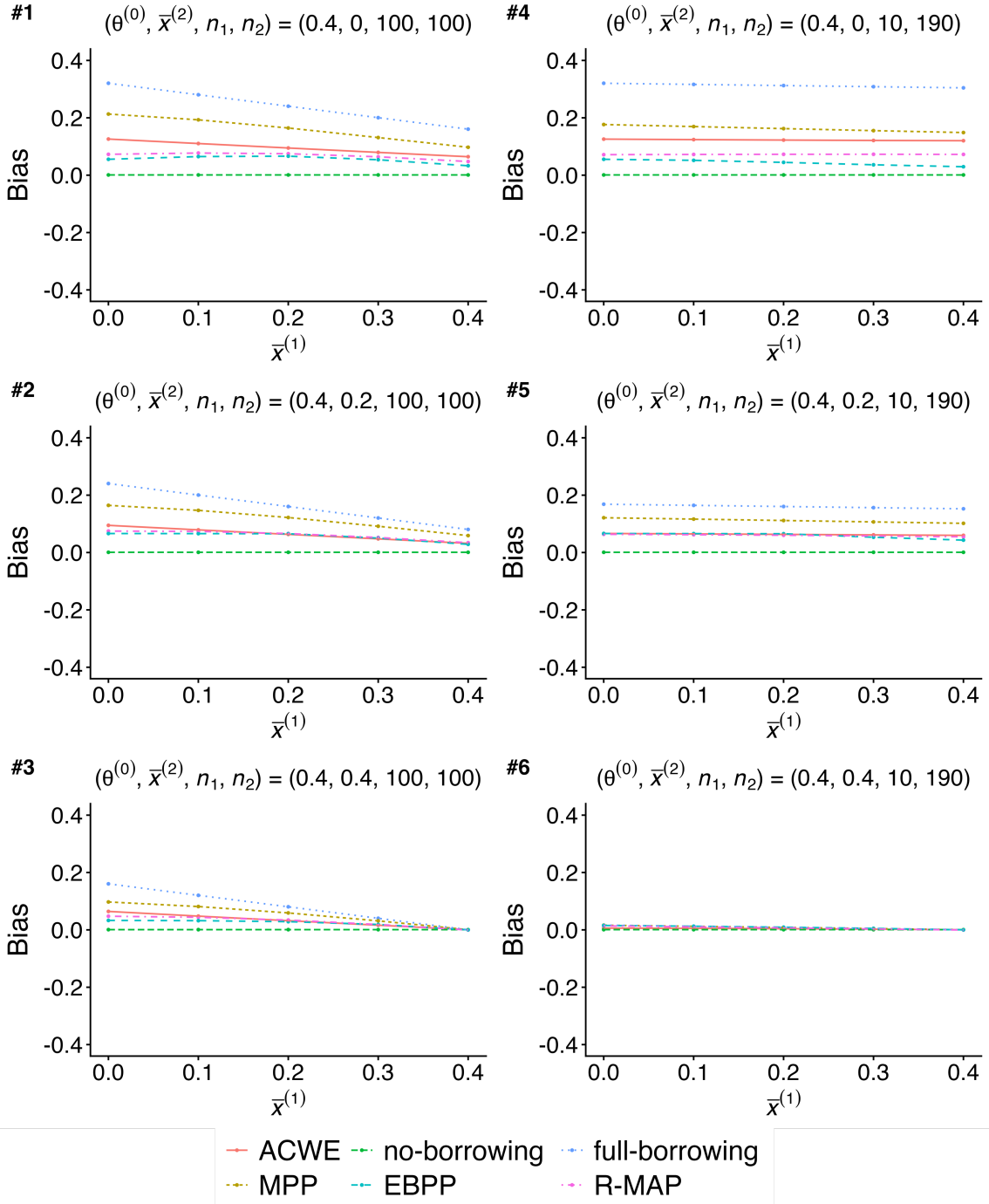
6	MPP	(0.1, 0.4)	0.29902
6	MPP	(0.2, 0.4)	0.28432
6	MPP	(0.3, 0.4)	0.27039
6	MPP	(0.4, 0.4)	0.25635
6	EBPP	(0, 0.4)	0.64308
6	EBPP	(0.1, 0.4)	0.62646
6	EBPP	(0.2, 0.4)	0.60214
6	EBPP	(0.3, 0.4)	0.57267
6	EBPP	(0.4, 0.4)	0.55637
A1	ACWE	(0, 0, 0, 0)	0.8113
A1	ACWE	(0.1, 0, 0, 0)	0.7947
A1	ACWE	(0.2, 0, 0, 0)	0.7801
A1	ACWE	(0.3, 0, 0, 0)	0.768
A1	ACWE	(0.4, 0, 0, 0)	0.7604
A1	no-borrowing	(0, 0, 0, 0)	0.6369
A1	no-borrowing	(0.1, 0, 0, 0)	0.6369
A1	no-borrowing	(0.2, 0, 0, 0)	0.6369
A1	no-borrowing	(0.3, 0, 0, 0)	0.6369
A1	no-borrowing	(0.4, 0, 0, 0)	0.6369
A1	full-borrowing	(0, 0, 0, 0)	0.9637
A1	full-borrowing	(0.1, 0, 0, 0)	0.9435
A1	full-borrowing	(0.2, 0, 0, 0)	0.9141
A1	full-borrowing	(0.3, 0, 0, 0)	0.8718
A1	full-borrowing	(0.4, 0, 0, 0)	0.8235
A1	R-MAP	(0, 0, 0, 0)	0.8303
A1	R-MAP	(0.1, 0, 0, 0)	0.8134
A1	R-MAP	(0.2, 0, 0, 0)	0.7806
A1	R-MAP	(0.3, 0, 0, 0)	0.7379
A1	R-MAP	(0.4, 0, 0, 0)	0.6954
A1	MPP	(0, 0, 0, 0)	0.9432
A1	MPP	(0.1, 0, 0, 0)	0.9179
A1	MPP	(0.2, 0, 0, 0)	0.8821

A1	MPP	(0.3, 0, 0, 0)	0.8328
A1	MPP	(0.4, 0, 0, 0)	0.7793
A1	EBPP	(0, 0, 0, 0)	0.9035
A1	EBPP	(0.1, 0, 0, 0)	0.8695
A1	EBPP	(0.2, 0, 0, 0)	0.8468
A1	EBPP	(0.3, 0, 0, 0)	0.8431
A1	EBPP	(0.4, 0, 0, 0)	0.8431
A2	ACWE	(0, 0.4, 0.4, 0.4)	0.5577
A2	ACWE	(0.1, 0.4, 0.4, 0.4)	0.5515
A2	ACWE	(0.2, 0.4, 0.4, 0.4)	0.5384
A2	ACWE	(0.3, 0.4, 0.4, 0.4)	0.5195
A2	ACWE	(0.4, 0.4, 0.4, 0.4)	0.4983
A2	no-borrowing	(0, 0.4, 0.4, 0.4)	0.6369
A2	no-borrowing	(0.1, 0.4, 0.4, 0.4)	0.6369
A2	no-borrowing	(0.2, 0.4, 0.4, 0.4)	0.6369
A2	no-borrowing	(0.3, 0.4, 0.4, 0.4)	0.6369
A2	no-borrowing	(0.4, 0.4, 0.4, 0.4)	0.6369
A2	full-borrowing	(0, 0.4, 0.4, 0.4)	0.2071
A2	full-borrowing	(0.1, 0.4, 0.4, 0.4)	0.1513
A2	full-borrowing	(0.2, 0.4, 0.4, 0.4)	0.1046
A2	full-borrowing	(0.3, 0.4, 0.4, 0.4)	0.07
A2	full-borrowing	(0.4, 0.4, 0.4, 0.4)	0.0428
A2	R-MAP	(0, 0.4, 0.4, 0.4)	0.5122
A2	R-MAP	(0.1, 0.4, 0.4, 0.4)	0.4594
A2	R-MAP	(0.2, 0.4, 0.4, 0.4)	0.4195
A2	R-MAP	(0.3, 0.4, 0.4, 0.4)	0.4024
A2	R-MAP	(0.4, 0.4, 0.4, 0.4)	0.4029
A2	MPP	(0, 0.4, 0.4, 0.4)	0.2942
A2	MPP	(0.1, 0.4, 0.4, 0.4)	0.227
A2	MPP	(0.2, 0.4, 0.4, 0.4)	0.1634
A2	MPP	(0.3, 0.4, 0.4, 0.4)	0.1145
A2	MPP	(0.4, 0.4, 0.4, 0.4)	0.0817

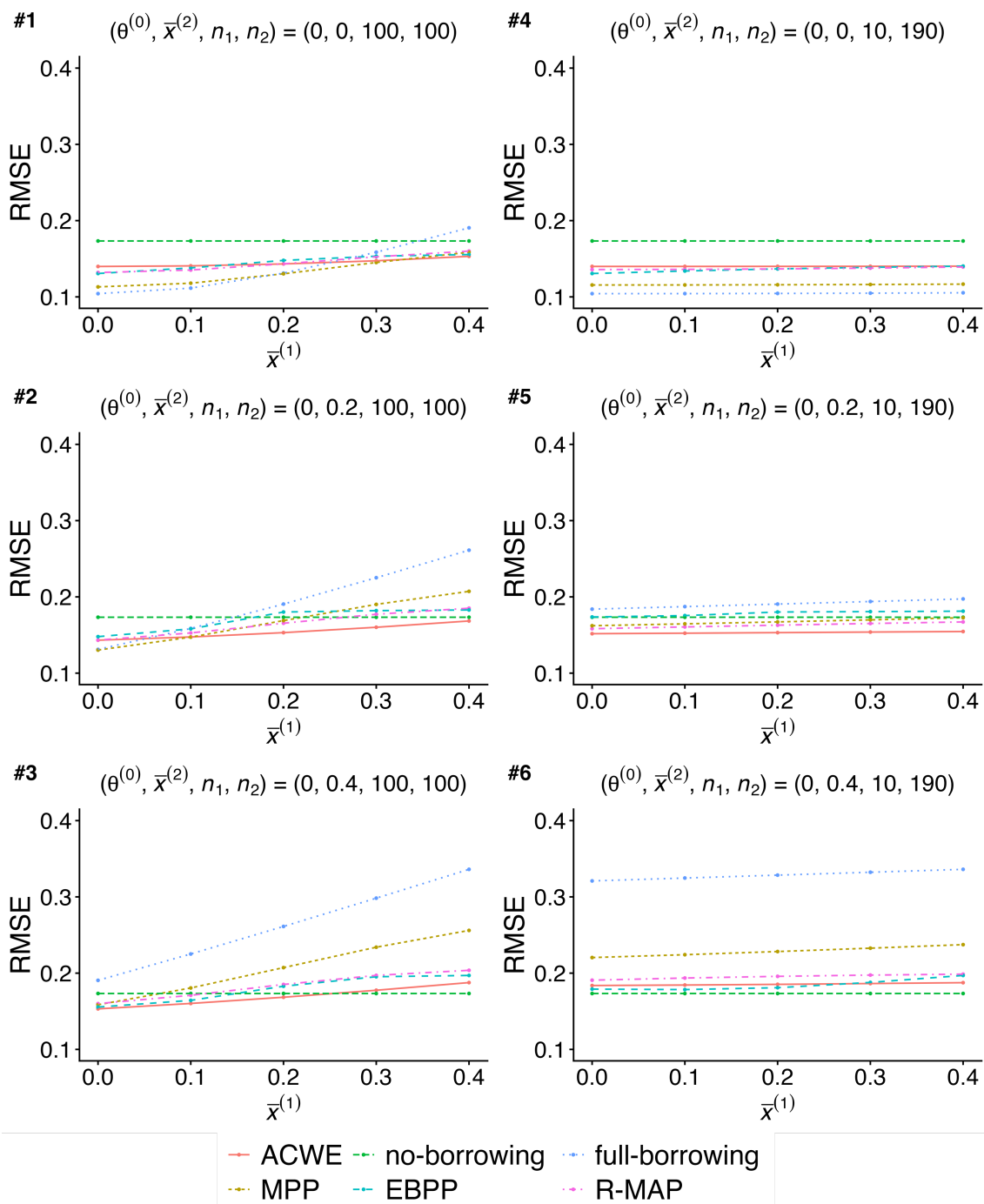
A2	EBPP	(0, 0.4, 0.4, 0.4)	0.8013
A2	EBPP	(0.1, 0.4, 0.4, 0.4)	0.7358
A2	EBPP	(0.2, 0.4, 0.4, 0.4)	0.6221
A2	EBPP	(0.3, 0.4, 0.4, 0.4)	0.564
A2	EBPP	(0.4, 0.4, 0.4, 0.4)	0.556



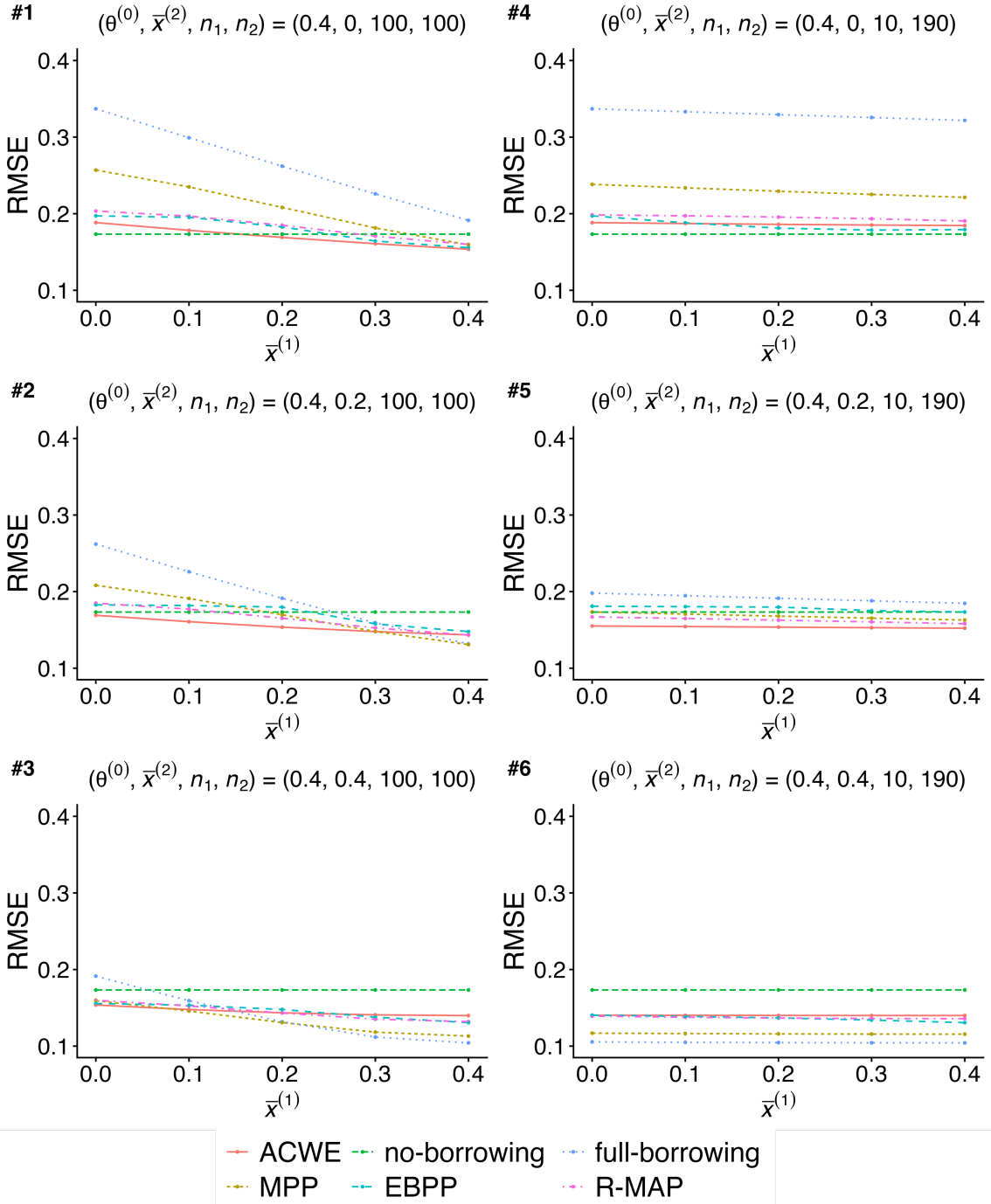
付録図 1. δ の推定量のバイアス (対立仮説下)



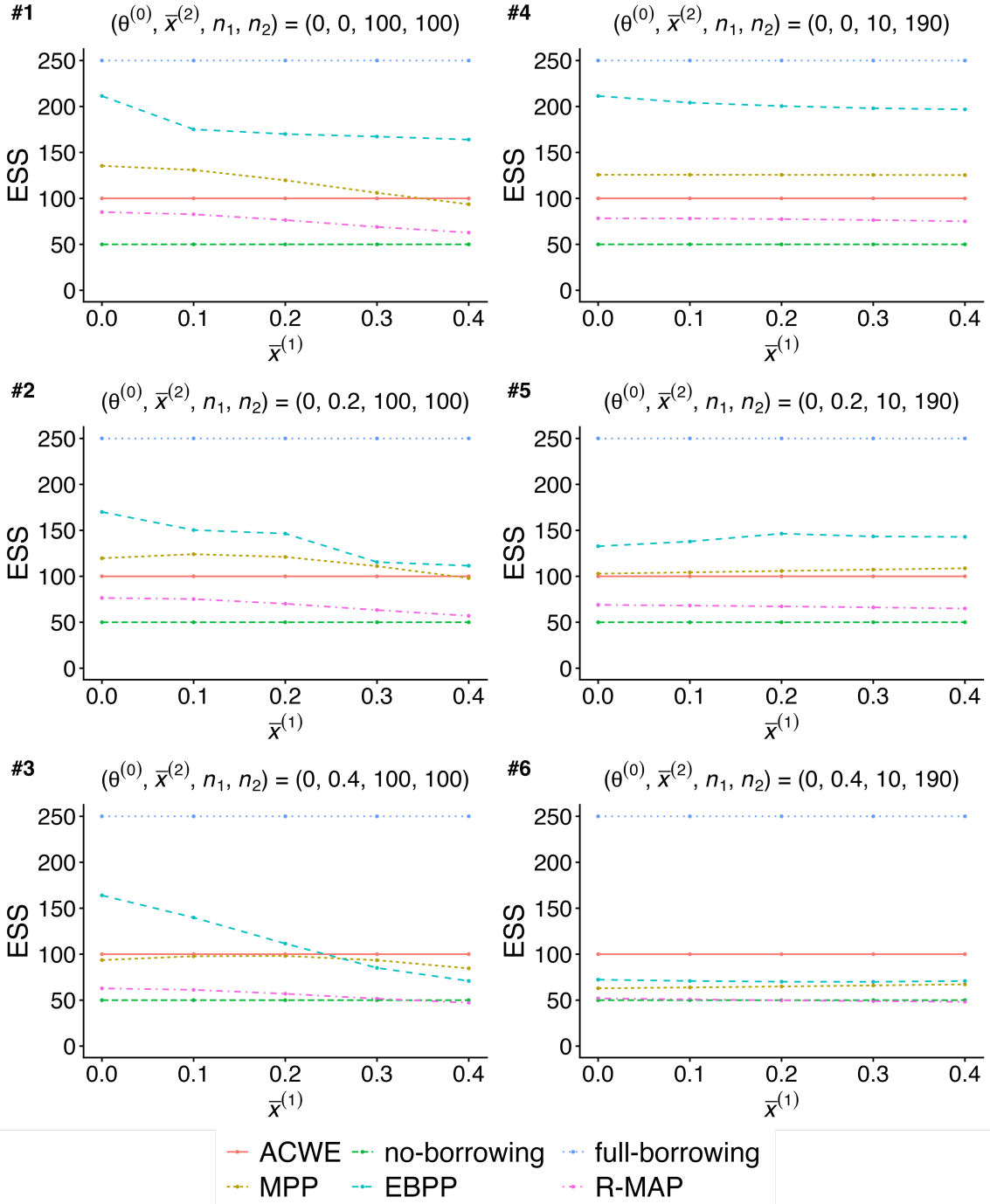
付録図 2. δ の推定量のバイアス（帰無仮説下）



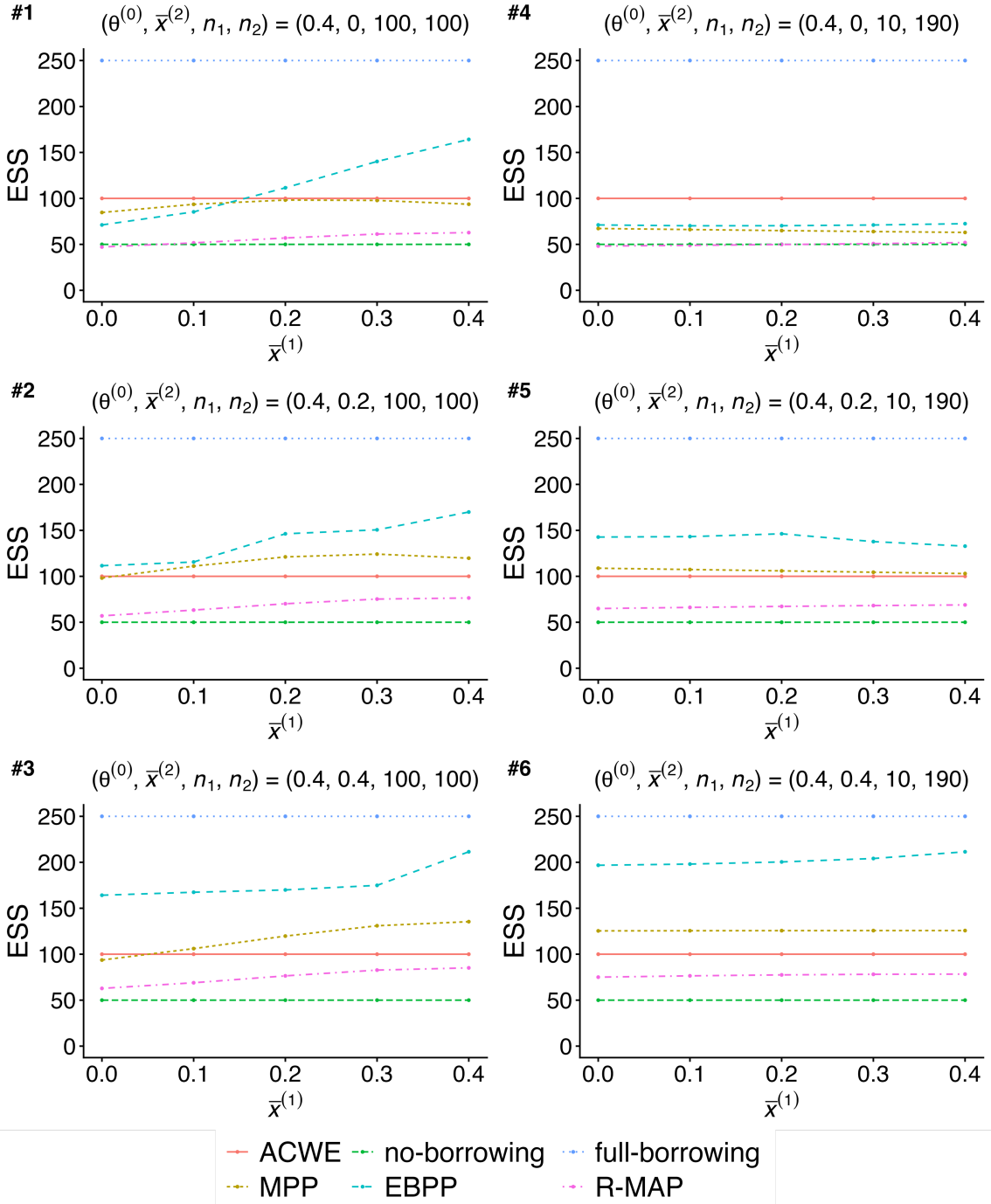
付録図 3. δ の推定量の RMSE (対立仮説下)



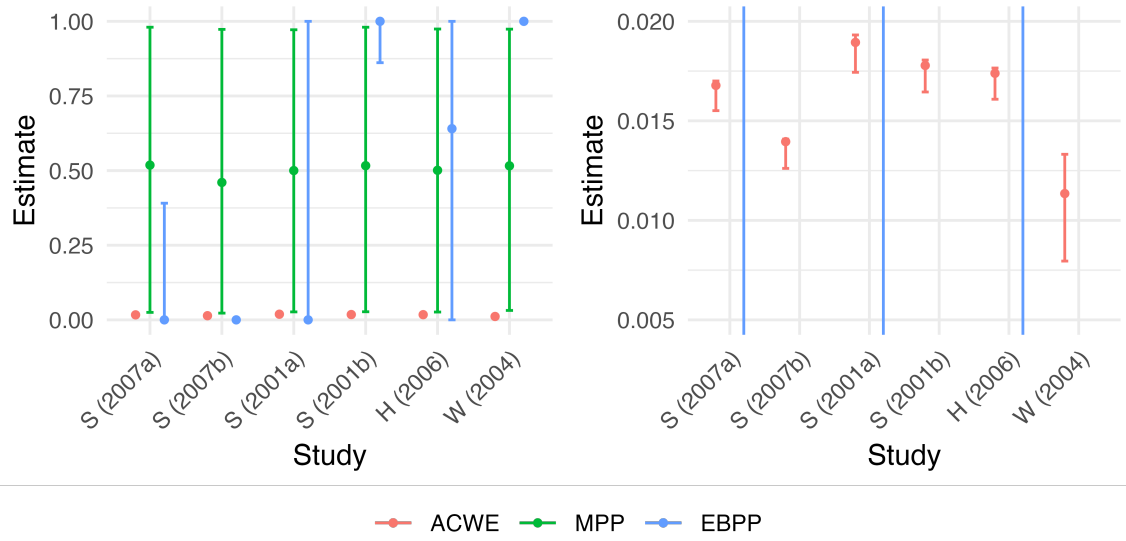
付録図 4. δ の推定量の RMSE (帰無仮説下)



付録図 5. 対照群の ESS (対立仮説下)



付録図 6. 対照群の ESS (帰無仮説下) .



付録図 7. 重みの点推定値と 95%区間（左：全範囲，右：一部範囲）。