

論文の内容の要旨

論文題目 Robust and efficient estimation of structural nested mean models in randomized trials with partial compliance using auxiliary covariates

(治療の不遵守を伴うランダム化試験における補助変数を用いた頑健かつ効率的な構造ネスト平均モデルの推定)

氏名 篠崎 智大

1. 序論

ランダム化試験においては、割付治療のスイッチや試験外治療の開始による治療の不遵守 (non-compliance) が生じ得る。治療不遵守が生じると、ランダム化により割付群間の比較可能性が保たれていても、実際に従った治療で定義されるグループ同士が比較可能であることは保証されない。治療不遵守の存在下でも割付群に従って比較を行う **intent-to-treat** の原則は、ランダム化のみに基づいて「全ての治療が全ての人のアウトカムを変化させない」という狭義の因果帰無仮説に対する適切なサイズの検定を与えることから一般に推奨されるものの、不遵守が存在しないような「理想的な」試験で得られるはずの有効性 (efficacy) について不可知であるという点で、科学的に興味のある比較法とは言い難い。そこで近年、このような治療不遵守を適切に調整する方法として因果解析手法の一つである **g-推定法** が注目を集めている。

g-推定法 は構造ネストモデルと呼ばれる因果モデルを推定する方法の総称である。ランダム化試験における **g-推定法** の特長として、交絡変数の測定を要せずランダム化のみに基づく構造ネストモデルの推定が可能である点が挙げられる。さらに、補助モデルや構造ネストモデルを正しく特定できているかどうかに関わらず、**intent-to-treat** 検定と同様に狭義の因果帰無仮説に対する妥当な検定を得ることができる。一方、**g-推定法** は「交絡変数を全て測定している」という仮定の下で共変量効果の明示的なモデル化を通じた交絡調整解析も可能であり、ランダム化が意図されていない治療に対する柔軟な因果効果のモデル化を通して、複雑な治療レジメの推測に適した手法でもある。

上記の「ランダム化に基づく **g-推定法**」(randomized **g-estimation**) と「未測定の交絡がない仮定に基づく **g-推定法**」(observational **g-estimation**) はそれぞれ利用する仮定やデータが異なるため、下記のようなトレード・オフがある：

- 1) **randomized g-estimation** はランダム化のみに基づくため頑健な推定量を得られるが、**observational g-estimation** は検証不能な「未測定の交絡がない」仮定の下であれば補助的な共変量情報を用いた効率の良い推定が可能である (バイアス=分散トレード・オフ)
- 2) **observational g-estimation** では補助的な共変量を解析に含める際に、補助モデルの正しい特定が必要となるが、**randomized g-estimation** ではそもそも識別可能な構造ネストモデルが単純なものに制約されてしまう (モデル誤特定=識別不能性トレード・オフ)

本研究では、**g-推定法** が推定方程式に基づく推定法であることより、一般化モーメント法 (generalized method of moments) を利用することで、これらのトレード・オフを折衷した推定方法を提案した。提案手法の優れた特性は、ランダム化試験における治療不遵守として複数の状況を想定したシミュレーション実験により検証した。さらに、転移・再発性乳がんに対する **S-1** とタキサン系薬剤の一次療法としての効果を比較したランダム化比較試験 (**SELECT BC** 試験) データへ本手法を適用して、臨床的に解釈可能な結果を得た。

2. 方法

アウトカム Y に対する治療 A の効果を表す加法的構造平均モデル

$$E(Y|A = a, R, X) - E(Y^0|A = a, R, X) = c(a, R, X; \psi)$$

を推定する g-推定法を考える。ただし、 R はランダム割付を表す指示変数、 X はベースライン共変量とし、 Y^a は治療が $A=a$ の場合に観測されるであろう潜在アウトカムを表し、 $c(\cdot)$ は $c(0, r, x; \psi) = c(a, r, x; 0) \equiv 0$ を満たす任意の関数とする。

randomized g-estimation では $Y^a \perp\!\!\!\perp R \mid V$ (層 $V \subseteq X$ 内でのランダムな割付) および $Y^a = Y^{r^a} = Y^a$ (割付は治療のみを通してアウトカムに影響) の 2 条件の下で、アウトカム回帰モデル $E(Y|A, R, X) = m(A, R, X; \alpha)$ の当てはめ値を代入した、対象者 $i = 1, \dots, n$ のデータ $O_i = (A_i, R_i, X_i, Y_i)$ に対する推定関数

$$\tau_{\text{rand}}^*(O_i; \psi) = \{R_i - h(V_i)\} \{Y_i - c(A_i, R_i, X_i; \psi) - m(0, R_i, X_i; \hat{\alpha})\}$$

より、 $\sum_i \tau_{\text{rand}}^*(O_i; \psi) = 0$ を満たす ψ を g-推定量とする。補助モデル $m(A, R, X; \alpha)$ の特定の正誤は g-推定量の一致性に影響しないが、構造モデルのパラメータ ψ の次元は R の次元 (割付群数 - 1) を超えることができない。

observational g-estimation では $Y^a \perp\!\!\!\perp A \mid R, X$ (未測定の交絡がない) の仮定の下で、アウトカム回帰モデル $E(Y|A, R, X) = m(A, R, X; \alpha)$ および治療 A の傾向スコアモデル $\Pr(A = a \mid R, X) = p_a(R, X; \beta)$ の当てはめ値を代入した推定関数

$$\begin{aligned} \tau_{\text{obs}}^*(O_i; \psi) = & \{ \partial c(A_i, R_i, X_i; \psi) / \partial \psi - \sum_a p_a(R_i, X_i; \hat{\beta}) \partial c(a, R_i, X_i; \psi) / \partial \psi \\ & \times \{ Y_i - c(A_i, R_i, X_i; \psi) - m(0, R_i, X_i; \hat{\alpha}) \} \end{aligned}$$

より、 $\sum_i \tau_{\text{obs}}^*(O_i; \psi) = 0$ を満たす ψ を g-推定量とする。未測定の交絡がない仮定に加え、 $m(A, R, X; \alpha)$ または $p_a(R, X; \beta)$ の一方が正しく特定されていれば g-推定量の一致性が保たれる。この推定関数では、構造平均モデル $c(\cdot)$ のパラメータ ψ の識別可能な空間に次元の制約は無い。

提案する一般化モーメント法は、 $\tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta}) = \{ \tau_{\text{obs}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta})^T, \tau_{\text{rand}}^*(O_i; \psi, \hat{\beta})^T \}^T$ の二次形式

$$Q_n(\psi) = \sum_i \{ \tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta}) \}^T \Omega_n(\mathbf{O}) \sum_i \{ \tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta}) \}$$

を最小にする ψ を推定量とする。ただし、 $\Omega_n(\mathbf{O}) = \{ n^{-1} \sum_i \tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta}) \tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta})^T \}^{-1}$ は推定関数 $\tau_{\text{GMM}}^*(O_i; \psi, \hat{\alpha}, \hat{\beta})$ の標本共分散行列とする。推定量の一致性には randomized g-estimation と observational g-estimation で仮定した条件が共に必要だが、前者は試験内での治療法を正確に測定されているランダム化試験においては蓋然性が高い。従って、以下のシミュレーション実験では、後者の「未測定の交絡がない」仮定に対する感度を評価する。

g-推定量の漸近分散あるいは共分散行列は、それぞれの推定関数から一意に導かれる影響関数の標本分散より推定する。本手法は、乗法的な構造平均モデルや、2 時点以上に渡って与えられる治療の効果に対する構造ネスト平均モデルへの適用も直ちに可能である。

3. シミュレーション実験

3.1. 設定

サイズ $n = 1000$ の 2 群比較のランダム化比較試験を想定し、1) 割付治療のスイッチのみ (1 パラメータ構造平均モデル)、2) 割付治療から試験外治療への変更あり (2 パラメータ構造平均モデル)、3) 割付治療のスイッチに続く試験後の後治療の開始 (2 パラメータ構造ネスト平均モデル) の 3 つのシナリオを想定したシミュレーション実験を行った。いずれの実験でも 1900 回の繰り返しにより推定量の標本分布を評価した。

ランダム割付の不遵守メカニズムとして、シナリオ 1 と 3 では割付治療スイッチの割合を変化させ、シナリオ 2 と 3 では試験外治療確率を割付群間や一次治療間で変化させた。「未測定の交絡がない」仮定への感度として、全てのシナリオで 1) 未測定の交絡がない、2) 弱い交絡変数が未測定、3) 強い交絡変数が未測定であった場合の推定値を得た。さらに、補助モデル $m(A, R, X; \alpha)$ および $p_a(R, X; \beta)$ の正しい特定と誤った特定の組み合わせを考えた。

3.2. 結果

割付治療のスイッチ、あるいは試験外治療への変更を調整した割付治療効果に対して、提案法は observational g-estimation で一致性が保たれている状況でも推定効率を向上し、さらに「未測定の変数がない」仮定あるいは補助モデルの誤特定による observational g-estimation のバイアスを半減することが示された。特にバイアスと分散を合わせて評価する平均二乗誤差基準では提案法は一貫して既存の g-推定法に劣ることが無かった。

一方で、治療開始確率がランダム割付 R に依存しないような後治療や試験外治療の効果に対しては、バイアス・分散の観点から必ずしも推定性能向上が見られなかった。これら試験外治療の効果推定はランダム化試験において二義的ではあるが、ランダム割付 R がこれらの治療開始にも影響する状況を設定することで、平均二乗誤差を大きく減らすことが分かった。

4. データ解析

4.1. 方法

SELECT BC 試験は、転移・再発性乳がんに対する S-1 とタキサン系薬剤の補助化学療法の一治療としての効果を比較したランダム化比較試験である。適格患者のうち 306 名が S-1 群へ、286 名がタキサン群へ割り付けられたが、S-1 群で 46 名、タキサン群で 27 名がそれぞれ割付治療をスイッチし、さらに全体の過半の 323 名が二次治療として研究計画書に指定された化学療法を開始した。アウトカムを 5 年以内の死亡とし、一次治療としての S-1 (vs. タキサン) の効果と、二次治療開始による効果を乗法的構造ネスト平均モデルでパラメータ化してリスク比を求めた。交絡変数として、年齢、肝転移、エストロゲンおよびプロゲステロン受容体発現、リンパ節転移数、切除後から再発までの時間、補助化学療法の有無 (以上、ベースライン共変量)、全身症状、腫瘍増悪、および増悪までの時間 (以上、時間依存性共変量) を補助変数として用いて、アウトカム回帰と傾向スコアをモデル化した。推定される構造ネスト平均モデルには、ベースライン共変量および時間依存性共変量による効果修飾を許容した複数パラメータモデルを用いた。

4.2. 結果

治療効果一様の最も単純な 2 パラメータ構造ネスト平均モデルでは、二次治療リスク比 (95%信頼区間) は g-推定法 (observational g-estimation) で 0.89 (0.71–1.10)、提案法で 0.88 (0.71–1.09) であった。二次治療を開始しなかった場合に発現するであろう S-1 の一次治療リスク比 (95%信頼区間) は g-推定法で 0.84 (0.60–1.17)、提案法で 0.78 (0.59–1.04) であった。続いて、サブグループ間で治療効果を比較した結果、エストロゲン受容体 (ER) が一次治療の効果修飾因子に、がん増悪 (PD) が二次治療の効果修飾因子になる可能性が示唆された。これらの効果修飾を構造ネスト平均モデルの推定結果を表に示す: 提案法では、ER(-)の一次治療リスク比 (95%信頼区間) は 1.13 (0.65–1.97)、ER(+)で 0.73 (0.54–0.98)、PD(-)の二次治療リスク比 (95%信頼区間) は 1.52 (0.83–2.79)、PD(+)で 0.74 (0.60–0.92) となった。

表: SELECT BC 試験データにおける構造ネスト平均モデルの推定結果

効果	g-推定法 (observational g-estimation)		提案法	
	リスク比	95%信頼区間	リスク比	95%信頼区間
一次治療 S-1 (ER 陰性)	2.15	0.39, 11.88	1.13	0.65, 1.97
一次治療 S-1 (ER 陽性)	0.75	0.55, 1.03	0.72	0.54, 0.98
二次治療 (腫瘍増悪なし)	1.82	0.87, 3.83	1.52	0.83, 2.79
二次治療 (腫瘍増悪あり)	0.72	0.58, 0.89	0.74	0.60, 0.92

5. 考察

SELECT BC 試験のデータ解析では、二次治療の調整を行うことで S-1 の一次治療としての生存延長効果がタキサン系薬剤に比べて大きく推定された。このように S-1 の一次治療効果がエストロゲン受容体のある対象者で大きく推定された点は、**intent-to-treat** によるサブグループ解析の結果でも観察されており、臨床的にも自然な結果と考えられる。二次治療は腫瘍が増悪した患者のリスクを大きく減少させ、腫瘍増悪の無い患者では生存延長効果が見られない点も臨床的に自然な推定値ではあるものの、シミュレーション実験で見たように、調整しきれていない交絡要因・モデル誤特定によるバイアスの影響が排除できない点は本解析の限界である。

一般化モーメント法やそれに併せて用いられる操作変数推定法は、経済学や遺伝疫学、薬剤疫学分野で応用されてきたが、提案法は特に治療不遵守を伴うランダム化比較試験において有用であると考えられる。特に生存時間などいわゆるハードなエンドポイントに対するランダム化試験は、ランダム化および試験内での密な治療法の記録により **randomized g-estimation** に必要な操作変数の仮定に対する蓋然性が極めて高い研究デザインである。さらに、重要なデータ測定が計画された上で実施されるため、単なる観察研究に比べても **observational g-estimation** に必要な変数を扱いやすい研究デザインだと考えられる。

提案法が最も推定性能を発揮できる研究デザインとして、多段階ランダム化試験が挙げられる。多段階ランダム化試験は、治療の結果（生体指標の値や奏効有無）に応じて適応的に次の治療を選択するような「動的な」治療レジメから最適な組み合わせをプラグマティックに検証する個別化医療のための試験デザインだが、観測データからレジメを選択するというデザインの特性上、割付治療の不遵守は避け難いと考えられる。多段階ランダム化試験の解析方法は、最適レジメ構造ネスト平均モデルが標準となりつつあるが、提案法を応用することで効率的かつバイアスの少ない最適レジメの選択が期待される。そのための拡張法は本論文内で示した。

本研究の限界として、1) 構造ネスト平均モデルのクラスとして加法的・乗法的なものに限定した点、2) 生存時間アウトカムに対する打ち切りの影響評価を徹底できていない点、3) ランダム化の仮定を所与としており、その破綻に対する感度を評価できていない点、4) SELECT BC 試験データにおける未測定交絡の存在可能性が挙げられる。

6. 結論

構造ネスト平均モデルに対する g-推定法に一般化モーメント法を応用し、「未測定の交絡がない」仮定および補助モデルをランダム化情報に組み合わせることで、既存の g-推定法の性能を向上させた。シミュレーション実験およびデータ解析から、よく計画・実施されたランダム化試験では、提案法によりバイアス減少と推定効率上昇が期待できることを示した。最適レジメを選択するための多段階ランダム化試験において提案法の有用性が特に期待される。