

令和 4 年 6 月 17 日現在

機関番号：32660

研究種目：若手研究

研究期間：2019～2021

課題番号：19K20227

研究課題名(和文) 臨床研究における傾向スコアモデルの選択と因果効果の推測

研究課題名(英文) Selection of propensity score models and inference of causal effects in clinical studies

研究代表者

安藤 宗司 (Ando, Shuji)

東京理科大学・工学部情報工学科・助教

研究者番号：40803226

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：臨床研究において、因果効果を偏りなく推定するために、次の3つのデータ解析に関する方法を検討・開発した。課題(1)傾向スコアモデルに含めるべき変数の種類の重要度の検討、課題(2)傾向スコアの推定法の開発、課題(3)因果効果を偏りなく推定するための推定量の開発。課題(1)と(2)については、研究成果を学会にて発表をした。課題(3)については、研究成果を論文にまとめ、学術雑誌に投稿済み(査読対応中)である。

研究成果の学術的意義や社会的意義

ランダムに処置の割り当てを行わない臨床研究では、因果効果を推定する際、交絡を調整する必要がある。交絡の調整法として、傾向スコア法がよく用いられている。交絡因子を特定できれば偏りなく因果効果を推定できるが、近年は、観測できる共変量の数が多くなり、交絡因子の特定がさらに難しくなっている。本研究では、観測されたデータから、交絡因子を特定することができる傾向スコアに対応した変数選択法と推定法を開発した。本研究の成果により、観測できる共変量の数が多い状況でも、偏りなく因果効果を推定できることになる。

研究成果の概要(英文)：In order to estimate causal effects without bias in clinical research, we evaluated and developed the following three methods for data analysis. (1) Investigation of the importance of the types of variables to be included in the propensity score model, (2) Development of a method for estimating a propensity score, and (3) Development of unbiased estimators of causal effects. The results of the research on issues (1) and (2) were presented at several conferences. For issue (3), the research results have been summarized in a paper and submitted to a journal (under review).

研究分野：統計科学

キーワード：統計的因果推論 変数選択法 モデル選択 モデル選択

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

ランダム化割付を行わない臨床研究では、比較群間の共変量の分布が異なるため、交絡の影響を取り除いて因果効果(治療効果)を推定する必要がある。交絡の調整方法として、近年、傾向スコア法が注目されている。従来、傾向スコアモデルに交絡因子のみ含めていた。しかし、近年、交絡因子と結果にのみ関連のある変数(結果予測因子)の両方を選択することで治療効果の推定量の分散が小さくなるという研究成果(Brookhart et al., 2006; Zhu et al., 2015)がある。しかし、この先行研究の評価方法は、限定的な状況でしか行っていない。また、傾向スコアモデルに対して、交絡因子と結果予測因子を変数選択するための手法として、Outcome-Adaptive Lasso が提案されている(Shortreed and Ertefaie, 2017)。

2. 研究の目的

本研究では、臨床研究において因果効果を偏りなく推定するために、次の3つの課題を解決することを目的とした。課題(2)は2年目に追加したものである。

課題(1) 傾向スコアモデルに含めるべき変数の種類の重要度の検討

交絡因子に加えて結果予測因子を傾向スコアモデルに含めることが推奨されている。しかし、この結果は、交絡因子、結果予測因子、処置予測因子が2つずつのみ設定した数値実験により評価された結果に基づいている。さらに、共変量間に相関がある状況は考慮されていない。このような背景から、共変量の数が増えた状況、共変量間に相関がある状況を想定して、傾向スコアモデルに含めるべき変数を詳細に再検討する。

課題(2) 傾向スコアの推定法の開発

傾向スコアモデルに対して、交絡因子と結果予測因子を変数選択するための手法として、Outcome-Adaptive Lasso が提案されている。しかし、Outcome-Adaptive Lasso は、処置予測因子も僅かではあるが、選択される傾向が先行研究の結果から認められた。本研究では、交絡因子と結果予測因子の選択割合は維持しつつ、処置予測因子の選択割合を軽減するために Outcome-Adaptive Bridge を提案する。別のアプローチとして、Outcome-Adaptive Lasso の重みを修正した方法も提案する。

課題(3) 因果効果を偏りなく推定するための推定量の開発

傾向スコアモデルを誤特定した場合、因果効果の推定に偏りが生じる。そのため、二重頑健推定法を用いて、因果効果を推定することが多くなっている。本研究では、傾向スコアの推定に Outcome-Adaptive Lasso、結果回帰の推定に Adaptive Lasso を用いた二重頑健推定法を提案する。

3. 研究の方法

課題(1)

交絡因子、結果予測因子、処置予測因子をそれぞれ3つずつ、4つずつ設定し、結果変数、処置変数への影響度も考慮した。さらに、それらの共変量間に相関がある状況を設定する。傾向スコアモデルに、交絡因子のみ、交絡因子と結果予測因子のみ、交絡因子と処置予測因子のみ、交絡因子、結果予測因子、処置予測因子のすべて、を含めた4つの場合を検討した。因果効果の推定法として逆確率重み付き法を用いた。

課題(2)

Outcome-Adaptive Lasso は、処置予測因子も僅かではあるが、選択される傾向が先行研究の結果から認められたため、方法を Lasso から Bridge に変更することでこの課題を解決することができる考えた。そこで、Adaptive Bridge を傾向スコアに対応するように拡張した Outcome-Adaptive Bridge を提案した。一方、Outcome-Adaptive Lasso の罰則項の重みを交絡因子と結果予測因子がより選択されやすいように変更することで、交絡因子と結果予測因子の選択割合は維持しつつ、処置予測因子の選択割合を軽減できる傾向スコアの推定法も提案した。

課題(3)

傾向スコアの推定に Outcome-Adaptive Lasso、結果回帰の推定に Adaptive Lasso を用いる。これらの推定法は、オラクル性(変数選択の一致性、漸近正規性)を有しており、交絡因子と結果予測因子を選択できることを、数理的に保証できる方法である。オラクル性を満たすためには、ある条件を満たす必要があるため、傾向スコアモデル、結果回帰モデルを誤特定する可能性はある。そのため、二重頑健推定法を用いることで傾向スコアモデル、結果回帰モデルのどちらか一方が正しく特定できれば、因果効果を偏りなく推定できる方法を提案した。

4. 研究成果

課題（１）

交絡因子，結果予測因子，処置予測因子をそれぞれ３つずつ，４つずつ設定した場合でも，交絡因子と結果予測因子のみを傾向スコアモデルに含めたときが，因果効果の推定量の偏り，平均二乗誤差が最小にした．先行研究の結果と同様に，本研究でも傾向スコアモデルには交絡因子に加えて，結果予測因子も含めるべきという結論になった．得られた研究成果を国内学会（２０１９年５月（ポスター），２０２１年５月（ポスター））で発表した．

課題（２）

Outcome-Adaptive Bridge と Outcome-Adaptive Lasso の重み改良法の性能を数値実験により評価した．数値実験の設定は，Shortreed and Ertefaie (2017) と同様にした．その結果，Outcome-Adaptive Bridge では，Outcome-Adaptive Lasso と比較して，処置予測因子の選択割合を大幅に減らすことができたが，交絡因子と結果予測因子の選択割合も若干ではあるが減った．その結果，因果効果の推定量の偏り，平均二乗誤差は，ほぼ同様の結果となり，改善は認められなかった．Outcome-Adaptive Lasso の重み改良法では，交絡因子と結果予測因子の選択割合は維持しつつ，処置予測因子の選択割合を減らす結果を得ることができた．その結果，因果効果の推定量の偏り，平均二乗誤差でも改善が認められた．得られた研究成果を国際学会（２０１９年８月（ポスター）），国内学会（２０１９年５月（口頭），２０２０年９月（口頭），２０２０年１１月（口頭））で発表した．さらに，投稿論文の草稿を作成した．

課題（３）

数値実験により，提案した因果効果の二重頑健推定法の性能を評価した．数値実験では，傾向スコアモデルのみを正しく特定できる状況，結果回帰モデルのみを正しく特定できる状況，傾向スコアモデルと結果回帰モデルの両モデルを正しく特定できる状況，傾向スコアモデルと結果回帰モデルの両モデルを正しく特定できない状況を想定した．その結果，傾向スコアモデルのみを正しく特定できる状況，結果回帰モデルのみを正しく特定できる状況では，提案した二重頑健推定法の性能が最も優れていた．傾向スコアモデルと結果回帰モデルの両モデルを正しく特定できる状況では，Outcome-Adaptive Lasso による傾向スコア法，Adaptive Lasso による結果回帰法と提案した二重頑健推定法の性能は同程度であった．傾向スコアモデルと結果回帰モデルの両モデルを正しく特定できない状況では，すべての推定法の性能が悪かった．得られた研究成果を国内学会（２０２１年５月（ポスター））で発表した．さらに，学術雑誌に投稿し，査読の結果を受け，投稿論文の内容を修正している．

< 引用文献 >

- Brookhart, A. M., Schneeweiss, S., Rothman, L. K., Glynn, J. R., Avorn, J. and Sturmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology* 163, 1149-1156.
- Zhu, Y., Schonbach, M., Coffman, L. D. and Williams, S. J. (2015). Variable selection for propensity score estimation via balancing covariates. *Epidemiology* 26, e14-e15.
- Shortreed, S. M. and Ertefaie, A. (2017). Outcome-Adaptive Lasso: Variable Selection for Causal Inference. *Biometrics* 73, 1111-1122.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計7件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 森田敦也, 本江渡, 福島龍一, 安藤宗司
2. 発表標題 処置効果の推定精度と傾向 スコアモデルに含める変数やC 統計量との関係
3. 学会等名 応用統計学会
4. 発表年 2021年

1. 発表者名 本江渡, 安藤宗司, 土田潤, 寒水孝司
2. 発表標題 Outcome-adaptive lasso と adaptive lasso を用いた二重頑健推定量の性能評価
3. 学会等名 応用統計学会
4. 発表年 2021年

1. 発表者名 福島龍一, 安藤宗司, 土田潤, 寒水孝司
2. 発表標題 傾向スコアモデルにおける変数の重要度を考慮したOutcome-Adaptive Lasso の構築
3. 学会等名 2020年度統計関連学会連合大会
4. 発表年 2020年

1. 発表者名 福島龍一, 安藤宗司, 土田潤, 寒水孝司
2. 発表標題 傾向スコアモデルにおける変数の重要度を考慮したOutcome-Adaptive Lasso の構築と実データへの適用
3. 学会等名 日本分類学会第39回大会
4. 発表年 2020年

1. 発表者名 福島 龍一, 本江 渡, 安藤 宗司, 寒水 孝司
2. 発表標題 平均処置効果の推定における傾向スコアモデルに含める変数の種類の検討
3. 学会等名 応用統計学会2019年年会
4. 発表年 2019年

1. 発表者名 安藤 宗司
2. 発表標題 Adaptive Bridge を用いた傾向スコアの変数選択
3. 学会等名 応用統計学会2019年年会
4. 発表年 2019年

1. 発表者名 Shuji Ando
2. 発表標題 Variable Selection in Propensity Score Using Adaptive Bridge for Confounding Adjustment
3. 学会等名 DSSV 2019 (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------