

令和元年5月31日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16015

研究課題名(和文) 最適な治療レジメを選択する多段階ランダム化試験デザインと推測手法の開発

研究課題名(英文) Development of design and analysis for multistage-randomized trials for selecting optimal treatment regimes

研究代表者

篠崎 智大 (Shinozaki, Tomohiro)

東京大学・大学院医学系研究科(医学部)・助教

研究者番号：60644482

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究課題は、個々の患者により良い治療を届ける「個別化医療」のひとつとして、治療に応じて変化するバイオマーカー情報や治療へのレスポンスに基づいて適応的に以降の治療レジメを選択する「動的な治療レジメ」に注目した。候補となる動的レジメの中で、長期的に標的とする臨床アウトカムへの効果を最大化するような「最適な治療レジメ」を選択する際には、適切な統計解析手法と、その手法の適切なデザインで行われた臨床研究(特に多段階にわたるランダム化試験デザインなど)への適用が必要となる。本研究課題を通して、特に統計的因果推論の立場からいくつかの方法論的な成果を得ることができた。

研究成果の学術的意義や社会的意義

適切な治療法の確立には、基礎研究だけでなく現実の患者を対象とした臨床データとその適切な統計解析が必要である。特に、長期にわたって治療法を計画するためには、最初に計画した治療法を遵守するだけでなく、臨床データの経過・変化に応じて、適応的に以降の治療を変更するような治療方針を取るべきである。このような一連の治療の流れを「治療レジメ」と呼び、中でも長期的なアウトカムを最も改善するものを「最適な治療レジメ」と呼ぶが、現実の医療データからこのような最適レジメを選ぶことは至難の業である。本研究課題では、「統計的因果推論」のアプローチを通して、統計的観点からこの目的を達成するための方法論的研究を行った。

研究成果の概要(英文)：As an approach to "personalized medicine" for delivering better treatments to each patient, this research project focused on "dynamic treatment regimes," which use biomarker information varying with treatment sequence and/or response to previous treatment to adaptively select subsequent treatment strategy. Among the candidate dynamic regimes, we should select the "optimal treatment regime" that maximizes the efficacy on targeted long-term clinical outcomes through validated statistical methods applied to adequately designed clinical studies including multistage-randomized clinical trials. I have made some methodological contribution for tackling challenges in estimating optimal treatment regimes from a statistical causal inference perspective.

研究分野：統計的因果推論、医療統計学

キーワード：統計的因果推論 臨床試験 個別化医療 動的治療レジメ セミパラメトリック推測

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

近年の日常診療データを利用した臨床研究の広がりや、医療費削減に対する社会的要求から、より効果の大きい対象者に治療を限定する「個別化医療」への期待が高まっている。個別化医療に向けたアプローチとしては、(a)「効果を発現することが期待できる治療対象者」を生体指標や遺伝子で識別し、選択的に治療を行うことと一般に捉えられることが多いが、(b) 日常診療の経過に伴って変化する生体指標や治療へのレスポンスに応じて以降の治療を適応的に選択するような治療レジメ（治療戦略）の探索も現実の医療において果たす役割が大きい。後者のように、あらかじめ治療法を固定するのではなく、患者情報の経過に応じて適応的に治療方針を変えていくルールのことを「動的な治療レジメ」という。さらに、動的治療レジメの中から、中長期的に最も興味のある臨床アウトカムへの効果を最大化するような「最適な治療レジメ」の選択が、より良い医療の確立のために重要と見なされている。

しかし、最適な治療レジメの選択問題は 2000 年前後までは広く認知されるような形では明確に定義されておらず、観察研究はもちろん、臨床試験のような実験研究からも、統計学的に妥当な方法論で推測を行うことができない状況であった。この解決には単に既存の統計手法を臨床データに適用するだけでは不十分で、近年になって方法論的な整備の進んできた「統計的因果推論」と呼ばれるアプローチによる定式化とデータ収集・解析が極めて有効であることが分かってきた。本研究課題は、統計的因果推論の方法論の開発を通して、最適な治療レジメを選択するためのデータ解析手法と、データ収集のための研究デザインの開発を目的としたものである。これらの方法論を広く実用に堪える形で応用し、さらに必要に応じて手法自体を発展させることは、統計科学領域において感心を集めるテーマであったと考えている。

2. 研究の目的

上記のような「動的な治療レジメ」の効果推定を行うための統計的方法論を発展させることを目的とした。本課題では、そのための統計的因果推論の方法論的な基礎研究も行うことを目的とした。

3. 研究の方法

本研究は、主に統計解析手法と研究デザインの方法論的研究であり、次に定める手順に従って、研究を進めることとなった：(1) 治療後の共変量を用いた「治療目標値を定める動的レジメ」の記述、および治療目標値を定めた動的レジメの効果の定式化、(2) 既存の推定方法を拡張したモデルの推定、(3) シミュレーション実験による推定方法の評価、(4) 臨床研究データ解析を通じた結果の解釈、(5) 論文作成、国際学会での成果報告。ただし、上記 1 から 4 は一つの研究（論文）内で網羅されることが普通なので、論文ごとに上記の手順で研究を進めた。なお、自ら臨床データを取得することは本課題の範疇外である。

4. 研究成果

(1) ランダム化比較試験における治療変更を考慮した解析法（g-推定法）の拡張

構造ネストモデルは、統計的因果推論で用いられる「潜在アウトカムモデル」に含まれる因果効果を定義するための統計モデルのひとつで、Harvard 大学の James Robins 教授が提唱した。非常に汎用性が高く、モデル化に置く仮定の観点からもバイアス・分散のトレードオフのバランスが良く、統計的な性能も優れている。構造ネストモデルの特長のひとつが、a) 観察データに基づいて多数の共変量を用いる g-推定法と、b) ランダム化のみに基づきほぼデータに仮定をおかない g-推定法の 2 種類を使い分けることができる点である。つまり、ランダム化試験においては、構造ネストモデルを用いて、ほとんど無仮定に正しい推測を行うことができる。一方で、共変量情報を用いないことによる、統計的な効率低下、識別可能なモデル空間の制限が問題点であるので、必ずしも b（ランダム化に基づく g-推定法）が a（観察データに基づく g-推定法）より望ましいわけではないというジレンマがある。

そこで、本研究では図 1 のように繰り返し治療が行われる臨床データを想定し、ランダム化試験における構造ネストモデルの推定法として、従来の g-推定法の定義を広げた推定方法を提案・定式化し、シミュレーション実験を通して性能評価を行った。なお、この手法は、最適な「動的レジメ」選択を目的として実施される多段階ランダム化試験への適用可能性が高い。研究内容は他の文献に例を見ないものであり、臨床的にも非常に有用な方法論研究だと考えている。

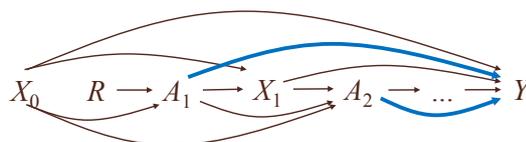


図 1. 繰り返し治療データ (A_k) および通常の統計的調整のできない時間依存性交絡変数 (X_k) を含むデータ構造

以上の研究成果は国内外の学会・シンポジウムで発表した。本研究課題の期間内での論文出版には至らなかったが、がん臨床試験データへ適用を行い、それを盛り込んだ形で論文にはま

とめてあり、生物統計学の専門誌に現在投稿中である。

(2) 動的レジメの直接効果推定

高齢者糖尿病患者を対象として「血清 LDL コレステロール値が x g/dL を下回ったら抗高脂血症薬アトルバスタチン治療を開始する」という動的レジメの中で、最も長期的な心血管疾患発症を抑える x を求めるデータ解析を行った。このようなレジメの選択は、それ自体臨床的に有用ではあるものの、そのメカニズムがブラックボックス化されてしまう危険がある。そこで、上記(1)のように直接・間接効果識別を通してメカニズムを推定しようという目的で行った研究である。

推定方法として、パラメトリックモデルを用いた g-計算アルゴリズムを用いた成果は本課題開始前に海外の学会で発表された (Shinozaki, T., Matsuyama, Y., Iimuro, S., Ohashi, Y. and Ito, H. (July 2012). “Estimation of controlled direct effects for dynamic treatment regimes.” At: 2012 Joint Statistical Meetings, San Diego, CA. Abstract 37)。しかし「直接・間接効果識別」については、近年の因果推論理論の発展に伴い、現実のデータからの原理的な識別不能性が論じられており、代替案が提案されている。そのひとつが介入主義的な直接・間接効果 (interventional direct/indirect effect) で、所属研究室の大学院生との共同研究として新しい統計手法として定式化し、同データを解析した。この研究内容は、現在論文文化を進めているところである。

(3) 傾向スコア法を用いた観察研究データ解析の pitfall の指摘

因果推論の手法としてよく用いられる統計解析手法に傾向スコア (propensity score) がある。傾向スコアは上記(1)の構造ネストモデルの推定にも用いられ、特に(1)のような複雑な構造をもつデータの統計的因果推論において、中心的な役割を果たすもののひとつである。

翻って、1 時点のみの治療から成る単純なデータ解析での交絡調整にも、傾向スコアが用いられることが多くなってきた。このような解析法には、傾向スコアマッチング、逆確率重み付け (inverse probability weighting, IPW)、傾向スコアの層別・回帰調整があるが、いずれもその計算法と群間での共変量バランス確認の容易さから、こと疫学分野では人気のある手法である。ただし、当然ながらこれらの手法が上手く働くためには、交絡調整に十分な共変量が全て測定され、傾向スコアはそれらを条件付けた曝露確率として少なくとも近似的に推定される必要がある。

上記3通りの「傾向スコア調整」(狭義には層別や回帰で「傾向スコアを条件付けた比較」を指すが、本稿ではこう呼ぶことにする) に続く群間比較においては、共変量を調整しない単純比較と、層別や回帰によって共変量を調整するやり方の2通りあり、いずれも解析原理上は擁護し得るものである。特に後者においては、傾向スコアとアウトカム回帰両方をモデルで近似する際に、いわゆる「二重ロバスト性」を有するような推定量を構成することができ、理論・応用の両面で研究が賑わっている。しかし、傾向スコア調整と、その後の層別・回帰において異なる共変量サブセットを調整するというアドホックなアプローチがしばしば医学・疫学分野で行われていることが本研究課題を進める中で明らかになり、しかも驚くことにこのようなアプローチに伴うバイアスの存在は、応用・統計理論の境界でまったく知られていなかった。そこで、このような「傾向スコア調整」に続く「共変量調整」における統計的ピットフォールを指摘し、バイアスの生じる状況を因果ダイアグラム (図 2) と解析的事実の両面から理論的に明らかにする研究を行い、疫学分野で最も影響力の強い雑誌である *Epidemiology* 誌に掲載予定である。

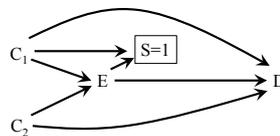


図 2. 共変量の一部を傾向スコアマッチング等で調整した状況を表す因果ダイアグラム

(4) マッチングを伴う研究デザインにおけるハザード比推定量についての統計学的考察

上記のように、傾向スコアを用いて比較するデータどうしマッチングさせて共変量バランスを取ることは、医学研究でよく行われている。傾向スコア以外にも、血縁関係や居住区など、より「個別的な」変数でマッチングをとることで、未測定の変数まで比較群間でそろえようとする研究デザインもあり、sibling マッチングと呼ばれている。これらマッチング後には通常の効果指標を用いて比較を行えばよいが、多くの追跡研究では、生存時間型アウトカムの効果指標としてハザード比を用いている。

マッチング後に、マッチされたペア相手の追跡が終わった時点で当該対象者の追跡を打ち切ることがよく行われている。しかし、このような追跡計画、あるいはデータ処理が推定量に与える影響は理論的に調べられておらず、慣例が先行している状況であった。しかしよく調べてみると、傾向スコアマッチングと sibling マッチングとでその影響は異なり、さらに適切な推定量も異なっていることを理論的に導くことができた。特に、sibling マッチングではマッチングされたペアを単位として層別したハザード比を推定する必要がある一方、傾向スコ

アマッチングではそのようなハザード比を推定しても科学的に興味のある効果指標と見なされないことが分かった。そして、sibling マッチングで求められるべきハザード比に関しては、上記のような「ペア内での打ち切り」はほとんどすべての推定量に対して全く影響がなかった。解析的に評価の難しい推定量に関してはシミュレーション実験で数値的な挙動を補い、生物統計の専門誌である International Journal of Biostatistics に採択され、現在オンライン掲載されている。

上記、(3)と(4)の研究成果は、研究課題である「最適な治療レジメを選択する多段階ランダム化試験デザインと推測手法の開発」には直接関与しないが、いずれも統計的因果推論の基礎的な知見として間接的に関連するテーマであったので、本研究課題の一環として行われた。

5. 主な発表論文等

[雑誌論文] (計 4 件)

- ① [Shinozaki, T.](#) and Nojima, M. (2019). Misuse of regression adjustment for additional confounders following insufficient propensity-score balancing. *Epidemiology*, in press. (査読あり)
- ② [Shinozaki, T.](#) and Mansournia, M. A. (2019). Hazard ratio estimators after terminating observation within matched pairs in sibling and propensity score matched designs. *International Journal of Biostatistics*, available online. (査読あり)
- ③ [Shinozaki, T.](#), Mansournia, M. A. and Matsuyama, Y. (2017). On hazard ratio estimators by proportional hazards models in matched-pair cohort studies. *Emerging Themes in Epidemiology*, 14, 6. (査読あり)
- ④ [Shinozaki, T.](#), Hagiwara, Y. and Matsuyama, Y. (2017). Re: Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology*, 28, e40-e41. (査読あり)

[学会発表] (計 11 件)

- ① [篠崎 智大](#) (2018 年 10 月). “観察研究データを用いた統計モデルによる交絡調整の実践：統計モデルによる交絡調整の基礎.” 第 24 回日本薬剤疫学会学術総会, 仙台.
- ② [篠崎 智大](#) (2018 年 10 月). “IPTW 推定量と二重ロバスト推定量.” 第 24 回日本薬剤疫学会学術総会 シンポジウム, 仙台.
- ③ 上村 夕香理, 川原 拓也, [篠崎 智大](#) (2018 年 10 月). “動的治療レジメンの比較方法とその適用場面.” 第 24 回日本薬剤疫学会学術総会, 仙台.
- ④ [篠崎 智大](#) (2018 年 9 月). “治療方針の因果効果と臨床試験.” 日本行動計量学会 第 46 回大会 特別セッション：社会に広がる因果推論の世界, 慶應義塾大学.
- ⑤ [篠崎 智大](#) (2018 年 9 月). “構造ネストモデルによる繰り返し治療の総合効果と直接効果の推定.” 2018 年度統計関連学会連合大会 企画セッション：医療統計学のフロンティア, 中央大学.
- ⑥ Hagiwara, Y., [Shinozaki, T.](#) and Matsuyama, Y. (July 12, 2018). “G-estimation of structural nested restricted mean time lost models to estimate causal effect of treatment on survival outcome.” At: The XXIXth International Biometric Conference, Barcelona, Spain. Abstract CS41.1.
- ⑦ [篠崎 智大](#), 野島 正寛 (2018 年 3 月). “傾向スコアと層別・回帰を組み合わせた交絡調整の誤用について.” 2018 年度日本計量生物学会年会 予稿集, 統計数理研究所, 東京.
- ⑧ [篠崎 智大](#) (2017 年 9 月). “周辺構造モデルによるパネルデータの解析.” 日本臨床疫学会 第 1 回年次学術大会 シンポジウム：観察研究における統計手法, 東京大学.
- ⑨ Nomura, S., [Shinozaki, T.](#) and Hamada, C. (March 12, 2017). “Comparison of several approaches to adjust overall survival (OS) for extensive crossovers in placebo-controlled randomized phase 3 trials.” At: The 2017 ENAR Spring Meeting, Washington, DC. Abstract 6h.
- ⑩ [篠崎 智大](#) (2017 年 3 月). “時間依存性交絡の調整.” 2017 年度日本計量生物学会年会, チュートリアル：観察研究における因果推論, 中央大学.
- ⑪ Hagiwara, Y., [Shinozaki, T.](#) and Matsuyama, Y. (April 13, 2016). “Estimation of modified effectiveness in randomized trials with subsequent treatments using dynamic marginal structural models.” At: UK Causal Inference Meeting 4th Annual Meeting, London, England. Abstract P18.

[図書] (計 1 件)

- ① [Shinozaki, T.](#) and Matsuyama, Y. (2019). Data analysis on real world data: an overview. In: Sudoh, O. (editor). *Epidemiologic Research on Real World Medical Data in Japan*. Switzerland: Springer International Publishing. In press (pages are not yet confirmed).

[その他]

Regulations.gov

篠崎智大らによるコメントページ (“ICH E9 (R1) Statistical Principles for Clinical Trials, Addendum: Estimands and Sensitivity Analysis in Clinical Trials” に対する米 Food and Drug Administration へのパブリックコメント)

<https://www.regulations.gov/document?D=FDA-2017-D-6113-0013>.

6. 研究組織

(1) 研究分担者

なし

(2) 研究協力者

なし

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。