

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 18 日現在

機関番号：22701

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24700278

研究課題名(和文)ボディ・マス・インデックスと死亡の因果的関連の統計学的検討

研究課題名(英文)Research for causal analysis between body mass index and death

研究代表者

田栗 正隆 (Taguri, Masataka)

横浜市立大学・医学部・助教

研究者番号：20587589

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：BMIと死亡の因果的関連を適切に検討することを目的として、時間依存性交絡を伴う曝露変数が存在する場合の統計学的方法論の開発を行った。時間依存性交絡を調整する方法である周辺構造モデルと関連した因果モデルである構造平均モデルのモデル選択手法の開発を行い、提案法は既存の方法よりも高い確率で正しい因果モデルを選択できることをシミュレーション実験により明らかにした。さらに、本質的に時間依存性交絡の問題である直接効果・間接効果の推定に関する方法論の開発を行った。あるコホートデータに関する死亡の解析では、IPTW法の重みのバラつきが非常に小さく、時間依存性交絡の程度があまり大きくないことが明らかになった。

研究成果の概要(英文)：For the purpose of examining a causal association between body mass index and death appropriately, I developed and examined statistical methodologies when there exists time-dependent confounding. I developed model selection criterion for two kinds of causal models, Marginal Structural Models and Structural Mean Models. I showed by simulation experiments that the proposed criterion can choose the right causal model with the probability that is higher than existing methods. I further developed estimation and sensitivity analysis methods for evaluating the direct and indirect effects of an exposure variable, in the presence of time-dependent confounding. By an analysis of a cohort data, I found that there is not large time-dependent confounding.

研究分野：統計科学

キーワード：因果推論 疫学

1. 研究開始当初の背景

これまで、BMI と総死亡の関連については U-shape の関連があることが知られており、実際国内外の多くの疫学研究において、この現象は観察されている (例えば Tsugane et al., 2002; Flegal et al., 2004)。しかし、多くの先行研究において、死亡リスクが最小となる最適なカテゴリーは、BMI が 25 程度の過体重 (overweight) の領域に来ており、BMI が低いところで総死亡との関連が強くなっていることと共に、その因果的解釈に関しては議論を呼んでいる。

ここで、因果的な解釈が難しいとされる 1 つの理由は、「reverse-causation bias」と呼ばれている現象の存在である。これは、BMI の低下による見かけ上の死亡リスクの増加は、異なる疾患の発症によってもたらされている可能性があるということである。例えば癌や糖尿病の患者は、発症後にやせる傾向があり、癌や糖尿病の発症が BMI (体重) と因果的に関連する。また、癌や糖尿病は死亡のリスクを増加させる。そのため、ある時点でデータを測定すると、BMI が低い対象者ほど死亡リスクが高くなっているということが考えられる。

しかしながら、このバイアスを考慮した解析は、全く行われていないといっても過言ではない状況である。癌や糖尿病の対象者を除いたサブグループに対して通常解析を適用した研究は見られるが、対象者を除くこと (対象者を限定してしまうこと) によって癌や糖尿病でない健康な対象者のみが解析に加わることから、選択バイアスが引き起こされてしまう可能性があり、このような解析方法は必ずしも妥当とは言えない。

2. 研究の目的

本研究では、上述のバイアスを補正するために、時間依存性交絡を考慮した推定方法を開発し、大規模前向きコホート研究で評価を行う事を目的とした。

具体的には、Robins(2000)の開発した周辺構造モデルとその推定法である IPTW 法を時間依存性交絡の補正に適用することを考えた。この方法は、対象者を実際に観察された曝露歴 (本研究では BMI のカテゴリーもしくは値) を受ける確率 (あるいは確率密度) の逆数で重み付けた、重み付き解析を行うことにより、興味ある回帰パラメータを推定するというものである。IPTW 法の大きな利点は、時間依存性交絡を調整し、バイアスのない推定値を与えることができる点にある。時間依存性交絡の代表的なケースは、観察研究データから治療の効果を推定する場合に起こる。例えば、日常診療で降圧剤を投与する状況を考える。まず診療時の血圧によって降圧剤を投与するかどうかが決まる。治療を受けることで患者の血圧は変化するが、その結果によって次に行う治療を決める。そしてその結果

によって次の治療を決める、という過程の繰り返しである。このようなデータの場合、血圧治療によって変化する血圧値が、死亡などの最終結果のリスク要因であり、かつ次の治療を規定する要因にもなっているという複雑な状況になっている。このような場合には、その項目 (血圧値) を時間依存性共変量として用いた回帰モデルで調整するという通常の解析方法では、偏りのない正しい結果を導くことができない。一例として、米国の Multicenter AIDS Cohort Study (MACS) における逆転写酵素阻害薬ジドブジンの死亡率に対する効果の検討では、CD4 陽性 T リンパ球数が時間依存性交絡因子になっている。Hernan et al. (2000) の、CD4 数で調整した通常解析の結果は、ハザード比が 2.3 (95%CI : 1.9-2.8) であった。この場合、治療効果が正の値 (死亡リスクを上げる方向) を示しているが、これは先に述べた時間依存性交絡の結果であり、状態の悪い対象者に治療が選択的に行われているためである (treatment by indication)。しかしながら、IPTW 法を用いることにより、全ての交絡因子が測定できているという仮定の下では、治療の正しい因果効果が推定可能である。Hernan et al. の研究では、IPTW 法による解析結果は、ハザード比で 0.6 (95%CI : 0.4 -1.0) となった。

BMI のような連続量の曝露変数に関しては、IPTW 法は統計学的な効率が悪いことが知られているため、その改善も 1 つの目的とした。

3. 研究の方法

時間依存性交絡を伴う曝露変数が存在する場合の統計学的方法論の開発を行った。周辺構造モデルに関しては、当初考えていた新しい推定方法は欠点があることが明らかになったため、方針を一部変更して (1) 周辺構造モデルと関連した因果モデルである構造平均モデルのモデル選択手法の開発、(2) 本質的に時間依存性交絡の問題である直接効果・間接効果の推定に関する方法論の開発を行った。また、(3) 実際のコホート研究のデータを用いた BMI を曝露因子とするデータ解析を行った。

4. 研究成果

上記 (1), (2), (3) の研究課題別に記述する。
(1) 周辺構造モデルと構造平均モデルのモデル選択手法の開発

構造平均モデルのモデル選択基準については、ランダム化臨床試験において治療の不遵守 (ノンコンプライアンス) が存在する場合の因果モデルのモデル選択基準として研究開始当初より開発中であった。本研究期間中に、提案する赤池情報量基準を拡張したモデル選択基準に関する理論的に望ましい性質を明らかにし、実データへの適用を行った

(図1) プラバスタチンの LDL コレステロール低減効果は個人間差が大きく、ベースライン LDL 値が高いほど効果が大きいことが示唆された。

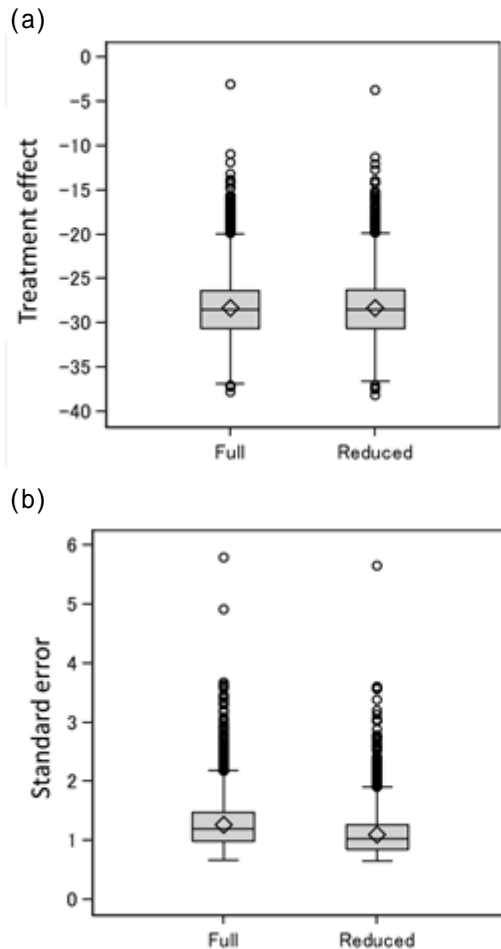


図 1. 大規模臨床試験データにおけるプラバスタチンによる因果治療効果の個人毎の推定値(a)とその標準誤差(b)の箱ヒゲ図(単位は mg/dL)。Full は全ての共変量(LDL 値、年齢、BMI)を推定に用いた場合、右図は提案したモデル選択基準により変数選択を行った場合。

本研究の成果は生物統計学の一流国際誌である *Biometrics* 誌に掲載されている(発表論文)。

さらに、この研究を発展させて、時間依存性交絡の存在下で用いる事のできる周辺構造モデルのモデル選択基準の提案を行った(発表論文)。シミュレーション実験の結果によれば、提案法は Platt et al. (2013) の提案したモデル選択基準よりも遥かに高い確率で正しい因果モデルを選択できることが示された。時間依存性交絡が存在する状況では曝露変数の経時的な繰り返し測定が必要になるが、測定回数が多いほど想定し得る因果モデルは多くなる。また、BMI のような連続量の曝露変数に対しては、様々な形状

の容量反応関係が想定される。したがって、複数のモデルの候補から、できるだけ正しいモデルをデータへの当てはまりの観点から統計学的に選択する手法を開発したことは、非常に意義のある成果だと考えている。

(2) 直接効果・間接効果の推定に関する方法論の開発

疾病発生を前向きに追跡する縦断的な疫学研究の1つの目的は、興味のある曝露の疾病発生に対する因果効果を推定することである。曝露と疾病の間に因果関係が示唆された場合、どのようなメカニズムで効果があるのかについての知見を得ることに興味を持たれる場合がある。例えば、肥満(BMI)の大腸がん発症に対する影響のうち、どの程度の割合を血中CRP濃度を上昇させる効果で説明できるかといった仮説である。この問題については、曝露(BMI)の疾病(大腸がん)に対する影響を、中間変数(CRP)を介した間接効果(indirect effect)と、それを介しない直接効果(direct effect)に分解することが、統計解析の目的となる(図2)。

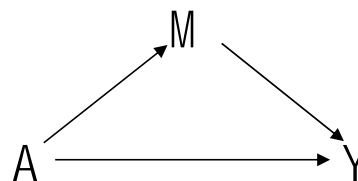


図 1. 直接効果・間接効果の検討の模式図。A は曝露変数、M は中間変数、Y は結果変数を表す。A → Y が直接効果、A → M → Y は間接効果。

曝露変数-中間変数、曝露変数-結果変数、中間変数-結果変数間の交絡を調整した上で直接効果・間接効果を推定する方法は、Robins and Greenland(1992)や Pearl (2001)をはじめ、これまでに多くのものが提案されてきている。しかしながら、中間変数-結果変数間の交絡因子(L)が曝露変数の影響を受け得る場合、ほとんどの既存方法はバイアスをもたらすため用いる事が出来ない。これは、交絡因子Lが曝露の影響を受け、さらに次に興味のある説明変数である中間変数に影響を与える、時間依存性交絡因子であることに起因する問題である。発表論文では、このような状況下で用いる事のできる推定方法を提案し、その方法が妥当であるための仮定を明らかにした。さらに、その仮定が崩れた場合に用いる事のできる感度解析方法を提案し、感度解析に用いる感度パラメータのバウンドを導いた。識別のための仮定が崩れた状況を想定した場合でも、典型的な状況では、提案法は既存の方法よりもバイアスが小さいことが期待できることを理論的に明らかにし、シミュレーション実験によってその通りに

なることの確認を行った。

(3) BMI を曝露因子とするデータ解析

あるコホートデータに関する死亡の解析では、IPTW 法の重みのバラつきが非常に小さく、このデータでは時間依存性交絡の程度があまり大きくないことが明らかになった。また、他のコホートデータに関するがん発症の解析では、BMI の総合効果だけでなく、中間変数に血液検査値等を想定した直接効果・間接効果の検討を行い、このデータでは直接効果が非常に小さいことを確認した。これらの成果は、今後研究論文としてまとめ報告していく予定である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計77件)[内査読付き74件]

Taguri M. Comments on 'A cautionary note concerning the use of stabilized weights in marginal structural models' by D. Talbot, J. Atherton, A. M. Rossi, S. L. Bacon, and G. Lefebvre. *Statistics in Medicine* 2015; 34(8): 1438-1439. (Letter, 査読有) doi: 10.1002/sim.6422.

Taguri M, Chiba Y. A principal stratification approach for evaluating natural direct and indirect effects in the presence of treatment-induced intermediate confounding. *Statistics in Medicine* 2015; 34(1): 131-144. (査読有) doi: 10.1002/sim.6329.

田栗正隆. 直接効果・間接効果の推定および未測定の間絡に対する感度解析. 統計数理 2014; 62(1): 59-75. (査読有) URL: <http://www.ism.ac.jp/editsec/toukei/tokeisuri-62j.html>

Taguri M, Matsuyama Y, Ohashi Y. Model selection criterion for causal parameters in structural mean models based on a quasi-likelihood. *Biometrics* 2014; 70(3): 721-730. (査読有) doi: 10.1111/biom.12165.

Taguri M, Matsuyama Y. Comments on 'An information criterion for marginal structural models' by R. W. Platt, A. M. Brookhart, S. R. Cole, D. Westreich, and E. F. Schisterman. *Statistics in Medicine* 2013; 32(20): 3590-3591. (Letter, 査読有) doi: 10.1002/sim.5810.

Chiba Y, Taguri M. Alternative monotonicity assumptions for improving

bounds on natural direct effects. *The International Journal of Biostatistics* 2013; 9(2): 235-249. (査読有) doi: 10.1515/ijb-2012-0022.

Chiba Y, Taguri M. Conditional and unconditional infectiousness effects in vaccine trials. *Epidemiology* 2013; 24(2): 336-337. (Letter, 査読有) doi: 10.1097/EDE.0b013e31828261f5.

Taguri M, Chiba Y. Instruments and bounds for causal effects under the monotonic selection assumption. *The International Journal of Biostatistics* 2012; 8(1): 24, 1-21. (査読有) doi: 10.1515/1557-4679.1386.

Taguri M, Matsuyama Y, Ohashi Y, Harahada H, Ueshima H. Doubly robust estimation of the generalized impact fraction. *Biostatistics* 2012; 13(3): 455-467. (査読有) doi: 10.1093/biostatistics/kxr038.

[学会発表](計22件)[内招待講演6件]

和泉志津恵, 田栗正隆. 線形構造平均モデルにおける適合性検定についての検討. 2014年度統計関連学会連合大会. 統計数理研究所, 東京都, Sep13-16, 2014. (Oral presentation)

Taguri M, Izumi S. A goodness-of-fit test for linear structural mean models. 27th International Biometric Conference, Florence, Italy, Jul6-11, 2014. (Oral presentation)

Taguri M. Estimation of semiparametric accelerated failure time models with covariate measurement error. 34th Annual Conference of the International Society for Clinical Biostatistics, Munich, Germany, Aug25-29, 2013. (Oral presentation)

Chiba Y, Taguri M. Conditional and unconditional infectiousness effects in vaccine trials: the relation and estimation. 46th Annual Meeting of the Society for Epidemiologic Research, Boston, USA, Jun18-21, 2013. (Poster presentation)

田栗正隆. 共変量の測定誤差を考慮したセミパラメトリック加速モデルの推定. 2013年度日本計量生物学会年会. パルセ飯坂, 福島県, May23-24, 2013. (Poster presentation)

Doi K, Taguri M, Furukawa K, Misumi M, Toshinaga S. Extensions to the regression calibration and substitution approach to mitigating the impact of dosimetric uncertainty on risk regression. 26th International Biometric Conference, 神戸国際会議場, 兵庫県, Aug26-31, 2012. (Invited oral presentation)

Taguri M, Chiba Y. A principal stratification approach for evaluating natural direct and indirect effects in the presence of intermediate confounding. 26th International Biometric Conference, 神戸国際会議場, 兵庫県, Aug26-31, 2012. (Oral presentation)

上村夕香理, 田栗正隆, 千葉康敬. 周辺構造モデルを用いた主要層別効果の推定および感度解析. 2012年度日本計量生物学会年会. 統計数理研究所, 東京都, May25-26, 2012. (Oral presentation)

[図書](計 1 件)

Chiba Y, Taguri M. Assessing the causal infectiousness effect in vaccine trials. In *Vaccines: Benefits and Risks*, iConcept Press, Chapter 3 (pp. 17-37), 2013. (査読有)

6. 研究組織

(1)研究代表者

田栗 正隆 (TAGURI, Masataka)

横浜市立大学・医学部・助教

研究者番号： 20587589