

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 28 日現在

機関番号：14202

研究種目：基盤研究(C) (一般)

研究期間：2012～2015

課題番号：24500341

研究課題名(和文) 栄養疫学データにおける測定誤差モデルの推測方法の提案

研究課題名(英文) Measurement error model for the analysis of nutrition epidemiology

研究代表者

田中 佐智子 (Tanaka-mizuno, Sachiko)

滋賀医科大学・医学部・准教授

研究者番号：50453824

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：本研究では、調査票に得られる食事摂取量にともなう測定誤差データを用いた解析における統計学的手法の提案をすることを目的とした。妥当性研究データの解析について、セミパラメトリックな分布を仮定した新しい測定誤差モデルを用いた解析手法を提案した。コホート研究において、測定誤差を伴う共変量がある場合のスプライン回帰モデルを提案し、シミュレーション研究とデータへの応用により提案法の有用性の検討を行った。

研究成果の概要(英文)：We proposed the hierarchical measurement error models for the analysis of the validation study in the nutrition epidemiology. Then, we also proposed the semiparametric spline regression model in which the measurement error distribution is not parametrised, and applied to the data of prospective cohort study.

研究分野：統計科学

キーワード：測定誤差

1. 研究開始当初の背景

近年、がんなどの生活習慣病の増加が大きな健康問題となっている。これらの疾患の発症は生活習慣と密接な関連があり、とりわけ食習慣との関連が深いことが知られており、予防医学の観点から、食事と疾患の発症との関連を明らかにすることは重要な問題である。

食事摂取量の評価には通常調査票が用いられる。調査票を用いた測定方法としては24時間食事思い出し法(24-hour dietary recalls; 24HR)がある。24HRは、一日の飲食物を記録する方法であり、一定期間の摂取量を評価するには測定日を複数回設ける必要があり、調査に負担とコストがかかるため、数万人規模の調査は現実的には困難である。そこで代替りの方法として、食物頻度摂取法(food frequency questionnaires; FFQ)が用いられる。FFQは過去の特定の期間の食事内容を100項目以上の質問表を用いて調査する方法である。この方法は、一定期間の食生活を思い出して測定する為、測定回数は少ないが、大きな測定誤差が含まれる可能性が示唆されておる。この測定誤差を無視した解析を行うと、食事摂取量そのものの過小評価や、がん発症との関連におけるバイアスを伴う結果をもたらす可能性がある。従って、FFQによる食事摂取量を用いてデータを解析する際には、測定誤差を考慮した統計モデルによる解析・検討を行う必要がある。

FFQの摂取量が「真の摂取量」に比べ妥当であるを評価する方法として、一部の対象者のみにFFQと24HRを複数回測定を行う妥当性研究がある。このような妥当性研究の解析方法としては、2つの測定方法の平均値の相関を評価することが多い(Kipnisら; AJE, 2001)。しかし、FFQの測定値に対しては「真の摂取量」からの乖離として、ランダム誤差だけでなく系統的誤差も存在すると考えられ、相関係数の評価のみではこれらの誤差により生じるバイアスが評価できない。この問題に対し、Kipnisら(AJE, 2004)は24HR

とFFQの食事摂取量データの同時モデルに基づく測定誤差モデルを提案した。ここで、Kipnisらは食事摂取量の分布に正規分布を仮定した。しかし、食事摂取量は正規分布に従わない報告も多く、より分布の仮定に対して頑健な方法が望まれる。

2. 研究の目的

本研究では、調査票により得られる食事摂取量に伴う測定誤差に対する統計学的解析手法の提案を目的とする。まず、妥当性研究データの解析において、「真の食事摂取量」の分布により頑健なセミパラメトリックな分布を仮定した新しい測定誤差モデルを用いることを提案する。次に、「真の食事摂取量」の分布に正規分布を仮定している先行研究の方法と提案法をシミュレーション研究により比較し、提案法が他の方法に比べて測定誤差の推定バイアスと精度の評価を行う。更に、実際のがん疫学研究のデータへの適用を通じて、提案する統計手法の評価を行う。

3. 研究の方法

本研究では、調査票により得られる食事摂取量に伴う測定誤差に対する統計学的手法の提案を目的とした。そこで、研究を3段階に分け、初年度は主に提案する測定誤差モデルのシミュレーション実験について、2年度は提案したモデルの実際の妥当性研究データへの適用について研究を行う。最後に、最終年度は前年度までに行われた食事摂取量の解析結果に基づき、食事摂取量の測定誤差を考慮した生存時間データ解析により、疾病と食事の関連評価まで行う。3年度を通して、提案する方法の理論的根拠の提示、実行するためのプログラムの作成、シミュレーション研究、臨床研究実データへの応用を行う。最後に結果をまとめ、論文投稿まで行うこととする。詳細な方法を以下に記す。

(1) 初年度の計画

まず、提案する方法の理論的な性質をまとめる。食事摂取量の測定誤差の評価においては、複数回測定された 24HR を基本として、FFQ の測定誤差による系統的バイアスとランダム誤差がどの程度かを検討する統計解析が望ましい。広く用いられている方法は、複数回測定される FFQ と 24HR の平均値の相関係数をとることであった (Kipnis ら; AJE 2001)。これに対し、Kipnis ら (AJE, 2004) は、FFQ と 24HR の同時モデルを提案し (以下、Kipnis の方法) FFQ による総エネルギー量は過小方向のバイアスがあることを示した。この方法では、「真の食事摂取量」の分布に正規分布を仮定している。しかし、食事摂取量は一般に正規分布から外れた分布を取る傾向があり、正規分布の仮定は現実的ではなく、測定誤差構造を誤って評価している可能性がある。そこで本研究では、Kipnis の方法に対して「真の食事摂取量」の分布により頑健なセミパラメトリックな分布を仮定した新しい統計手法を提案する (以下、提案法)。更に、提案法が Kipnis の方法と比べて測定誤差のバイアスが小さくなることを理論的に示す。

次に、統計ソフトウェア SAS を用いて、先行研究 (Kipnis の方法) と提案法を実行する。ここで、プログラムが正確に作成されていることを確認するために、プログラミング専門家に依頼し、ダブルプログラミングを行う。ダブルプログラミングを行うことは、プログラムの問題点を明らかにし、エラーを見つけるだけでなく速度の向上にもつながることが期待される。

また、シミュレーション研究 (計算によるアプローチ) も実施する。作成したプログラムを用いて、シミュレーション研究を行う。実際のがん疫学データを参考にして、食事調査を行う大規模疫学研究を設定し、2 つの方法 (Kipnis の方法、提案法) を応用し、バイアスと分散の大きさを評価する。期待される傾向は、提案法により、測定誤差バイアスが

正しく評価されることである。更に、他の設定に対しても応用し、性能評価を行う。

(2) 2 年度の計画

妥当性研究データへの応用 (実際のデータへの応用) を行う。提案した方法を妥当性研究データに応用し、FFQ による食事摂取量の測定誤差バイアスを推定する。先行研究では、同じデータに対して相関係数により評価を行った結果が提示されており、今回得られた結果と先行研究による結果が乖離していないかの確認を行う。

統計ソフトウェア SAS を用いて、データに応用する場合のプログラムを作成する。ここでも、プログラムが正確に作成されていることを確認するために、プログラミング専門家に依頼し、ダブルプログラミングを行う。また、実用化を目指して、発表論文の添付文書とすることでプログラムの公開を行う。

(3) 最終年度の計画

前年度までに検討された測定誤差モデルを考慮した下で、一般集団を対象とする大規模疫学研究のデータに応用し、がんの発症に対する食事摂取量のハザード比を推定する。先行研究では、同じデータに対する単純なコックス回帰による結果が提示されており、今回得られた結果と先行研究による結果が乖離していないかの確認を行う。

ホームページを作成し、一般住民の関心が深いであろう、食事摂取量とがん発症の関連の解析結果内容をまとめ提示する。

4 . 研究成果

(1) 提案する手法の理論的な検討

提案する重み付き推定の方法の理論的性質をまとめた。食事摂取量の測定誤差の評価においては、複数回測定された 24HR を基本として、FFQ の測定誤差による系統的バイアスとランダム誤差がどの程度かを検討する統計解析について検討を行った。先行研究で

は、「真の食事摂取量」の分布に正規分布を仮定している。しかし、食事摂取量は一般に正規分布から外れた分布を取る傾向があり、正規分布の仮定は現実的ではなく、測定誤差構造を誤って評価している可能性がある。そこで本研究では、先行研究の方法に対して「真の食事摂取量」の分布により頑健なセミパラメトリックな分布を仮定した新しい統計手法を提案した(以下、提案法)。そして、結果変数が連続変数である食事摂取量の場合、本研究の回帰モデルの推定は最尤法によって推定することが可能であることも示した。

(2) 統計ソフトウェアによるプログラム作成

統計ソフトウェア SAS を用いて、先行研究(Kipnisの方法)と提案法を実行した。ここで、プログラムが正確に作成されていることを確認するために、プログラミング専門家に依頼し、ダブルプログラミングを行った。ダブルプログラミングを行うことで、プログラムの問題点を明らかにし、エラーを見つけるだけでなく速度の向上にもつながったと考えられる。

(3) シミュレーション研究(計算によるアプローチ)

作成したプログラムを用いて、シミュレーション研究を行う。実際のがん疫学データを参考にして、食事調査を行う大規模疫学研究を設定し、2つの方法(Kipnisの方法、提案法)を応用し、バイアスと分散の大きさを評価する。この検討により、提案法により、測定誤差バイアスが正しく評価されることが示された。

(4) 妥当性研究データへの応用(実際のデータへの応用)

提案した方法を妥当性研究データに応用し、FFQによる食事摂取量の測定誤差バイアスを推定する。先行研究では、同じデータに対し

て相関係数により評価を行った結果が提示されており、今回得られた結果と先行研究による結果が乖離していないかの確認を行った。前年度までに検討された測定誤差モデルを考慮した下で、一般集団を対象とする大規模疫学研究のデータに応用し、がんの発症に対する食事摂取量のハザード比の推定を推定した。また、同様に糖尿病の臨床研究のデータにも応用を行った。

5. 主な発表論文等

[雑誌論文](計 3 件)

Horiakwa C, Yoshimura Y, Kamada C, Tanaka S, Tanaka S, Hanyu O, Araki A, Ito H, Tanaka A, Ohashi Y, Akanuma Y, Yamada N, Sone H; Japan Diabetes Complications Study Group. Dietary Sodium Intake and Incidence of Diabetes Complications in Japanese Patients with Type 2 Diabetes - Analysis of the Japan Diabetes Complications Study (JDCS). J Diabetes Investig. 2014 Mar 23;5(2):176-87
査読あり

Tanaka S, Yoshimura Y, Kamada C, Tanaka S, Horikawa C, Okumura R, Ito H, Ohashi Y, Akanuma Y, Yamada N, Sone H; Japan Diabetes Complications Study Group. Intakes of dietary fiber, vegetables, and fruits and incidence of cardiovascular disease in Japanese patients with type 2 diabetes. Diabetes Care; 36(12):3916-3922. 2013
査読あり

Tanaka S, Yoshimura Y, Kawasaki R, Kamada C, Tanaka S, Horikawa C, Ohashi Y, Araki A, Ito H, Akanuma Y,

Yamada N, Yamashita H, Sone H; for the Japan Diabetes Complications Study Group. Fruit intake and incident diabetic retinopathy in Japanese patients with type 2 diabetes. Epidemiology; 24(2):204-11. 2013
査読あり

(3)連携研究者
なし

〔学会発表〕(計 2 件)

Tanaka S, Tanaka S. Semiparametric Regression Analysis for Interval Censored Time-to-event Data with Covariate Measured with Error. International Biometric Conference, Kobe, Japan, 2012

Tanaka S. Comparison of the inverse probability weighted estimators in the analysis of the nested case-control study data. East Asia Regional Biometric Conference 2012, Seoul, Korea, 2012

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

記載事項なし

取得状況(計 0 件)

記載事項なし

〔その他〕

記載事項なし

6. 研究組織

(1)研究代表者

田中 佐智子 (TANAKA, Sachiko)

滋賀医科大学・医学部・准教授

研究者番号: 50453824

(2)研究分担者

なし