

令和 6 年 6 月 20 日現在

機関番号：32202

研究種目：基盤研究(C)（一般）

研究期間：2018～2023

課題番号：18K10073

研究課題名（和文）アウトカム誤分類を考慮した競合リスク生存時間解析とその感度分析に関する研究

研究課題名（英文）Competing risks analysis with misclassified outcome and sensitivity analysis

研究代表者

三重野 牧子（Mieno, Makiko）

自治医科大学・医学部・准教授

研究者番号：60464707

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：臨床研究や疫学研究において、死亡診断書での原死因を死因データとして用いた分析を行う際、記録された死因の間の誤分類が大きい場合にはリスク因子の推定値へのバイアスが大きくなることが考えられる。本研究では、死因の誤分類における特性を考慮した方法を用いて、リスク推定方法および感度分析方法について検討した。大規模疫学研究では特に誤分類が結論に及ぼす影響は大きい可能性が示唆され、また、リスク推定の際にはアウトカムの分類の精度に影響する要因をよく検討しておく必要があると考えられた。

研究成果の学術的意義や社会的意義

死亡診断書での死因の誤分類を起こすもっとも強い要因は原死因であり、疾患による違いが大きいこと、高齢者であるほど誤分類率は上昇することが示唆されていたが、死亡診断書の死因データを用いたリスク推定研究において全症例の原資料まで精査することは現実的ではない。我々は測定誤差の問題と考え、感度分析を行って結果を解釈する方法について検討したところ、実際の長期追跡コホートデータ解析からも、アウトカム誤分類の影響に留意して結果を解釈する必要性が示唆された。

研究成果の概要（英文）：In clinical and epidemiological studies, when analyzing the original cause of death on the death certificates, large bias in the estimated risk factors should be found if there is a large misclassification between the recorded causes of death. In this study, we investigated risk estimation methods and sensitivity analysis methods using a method that takes into account the characteristics of misclassification of causes of death. It was suggested that misclassification may have a particularly large impact on the conclusions in large-scale epidemiological studies, and that it is necessary to carefully consider factors that affect the accuracy of outcome classification when estimating risks.

研究分野：疫学・生物統計学

キーワード：誤分類 アウトカム 競合リスク 感度分析 死亡診断書

1. 研究開始当初の背景

死亡診断書での原死因を死因データとして用いた分析において、記録された死因の間の誤分類が大きい場合にはリスク因子の推定値へのバイアスが大きくなる。結果の解釈を左右する状況となるにも関わらず、誤分類の影響についてはこれまでにあまり議論されてきていなかった。とくに、アウトカムとして複数の死因が競合する場合の生存時間解析においては、多くの仮定を必要とする既存の誤分類修正方法が適切でない状況も多く、またハザード比で結果を要約する以外の方法についての検討もあまりなされていなかった。

我々の先行研究である、東京都健康長寿医療センターで行われた連続剖検 450 例を対象とした死亡診断書にて報告された主な原死因(悪性腫瘍、心疾患、肺炎)と、剖検データと臨床カルテ記録に基づいて病理学者が判断した参照基準との比較研究では、ICD-10 コードの基準で見たところ、悪性腫瘍についての一致率は 81%と比較的高い結果であったが、心疾患では 55%、肺炎では 9%と低い一致率であった。より広い分類で一致率を確認すると、厚生労働省の死因統計で用いられている分類でも肺炎の一致率は 4 割に満たなかった。さらに、誤分類を起こすもっとも強い要因は原死因であり、疾患による違いが大きいこと、高齢者であるほど誤分類率は上昇することが示唆された。死因の誤分類は実際に起きていて無視できない程度であることが判明したが、死亡診断書の死因データを用いたリスク推定研究において全症例の原資料まで精査することは現実的ではなく、むしろ測定誤差の問題と考え、感度分析を行って結果を解釈するのが妥当と考えた。

2. 研究の目的

死因間に誤分類がある場合の生存時間解析におけるリスク推定への影響について、リスク推定方法および感度分析方法について検討し、シミュレーションデータおよび実際の長期追跡コホート研究データに適用することを目的とした。その際に、死因の誤分類における特性を考慮した方法を特に検討することとし、現実的かつ妥当な方法の提案を行うこととした。あわせて、ハザード比以外の要約指標の適用可能性についても検討した。

3. 研究の方法

(1) 潜在アウトカムの誤分類が及ぼすオッズ比の推定値への影響

競合リスクのある生存時間解析のためのモデル構築を検討する際には生存時間および打ち切りの考慮が必要となるが、モデル構築のために、よりシンプルなロジスティックタイプのモデルでの検討を行った。多項ロジスティックの状況下で誤分類が起きているようなモデルにおける推定方法について検討した。

方法としては、アウトカム誤分類に関して、特に、最も興味のあるアウトカムとそのリファレンスアウトカムの測定には誤分類がないが、第 3 の潜在アウトカムに誤分類が存在した場合、その大きさが興味のあるアウトカムに対するリスク推定値に及ぼす影響を定量的に示すこととした。第 3 のアウトカムとしては例えば、がん研究における循環器疾患死亡や、実は 2 値ではなく中間の値が存在する質問紙スコア、死亡をアウトカムにした臨床試験における脱落中の死亡などが想定された。

リスクを推定したい要因あるいは介入も 2 値であり、真に興味のあるアウトカムを発生した場合には誤差なく測定された(感度 100%である)ような場合に、そのオッズ比の推定値に潜在アウトカムの誤分類が及ぼすバイアスの理論値を、アウトカムの発生率、誤分類率をパラメータとした関数として示した。関数に数値パターンを当てはめて計算し、バイアスの理論的推定値の範囲を求め、どの程度誤分類の影響をうけるか検討した。

(2) 長期追跡コホート研究データへの適用

実データへの適用としては、自治医科大学を中心として行われている多施設コホート研究である JMS コホート研究データを解析対象とした。全国 12 地域の住民 12490 例を対象とした JMS コホート研究データでは追跡期間中央値が約 20 年と長期にわたっており、ベースラインリスク因子のアウトカムへの長期影響を検討できると考えられた。死因情報についてのデータ整理を進めたうえで、死因についての競合アウトカムとして、がん死亡・循環器疾患死亡・肺炎による死亡・その他の原因による死亡の 4 種類に分類し、リスク推定を行った。競合リスク生存時間解析としては、累積死亡率や比例ハザードモデルの枠組みに加え、より柔軟に共変量をモデル化できる分位点回帰生存時間解析あるいは quantile residual life 回帰への適用を検討した。感度分析の方法検討の際には、全体の死亡イベントに対するハザード比推定値が 2 前後であった喫煙の有無(ベースライン情報)にとくに注目して推定値への影響を観察し、性別、年齢、BMI、飲酒状況、血清総コレステロール値、Lp(a)等の解析も行った。

4. 研究成果

(1) 潜在アウトカムの誤分類が及ぼすオッズ比の推定値への影響

真には要因とアウトカムに関連がまったくない場合、潜在アウトカムの誤分類が differential、つまり要因が誤分類に影響する状態であり、例えば、要因によるアウトカム誤分類率が1~3倍ほど差がある状況であれば、潜在アウトカムが45%で、真のアウトカム発生率15%で推定オッズ比に-0.5~+3程度、潜在アウトカムが20%で真のアウトカム発生率40%では-0.2~+1程度のバイアスがみられた。differentialで、ある程度と興味のあるアウトカムへの誤分類、またリファレンスへの誤分類が要因によって方向性が一致しているかいないかによって、バイアスの大きさと方向が大きく影響を受けることが明らかになった。また、潜在アウトカムの存在が5%ほどに抑えられれば、バイアスは-0.2~+0.2の範囲で収まった。潜在的に分類が難しいアウトカムの存在がある状況で、その誤分類が non-differential であれば推定値は影響されないが、differential な場合においては推定したいオッズ比に無視できない程度のバイアスが存在し、特にオッズ比の推定精度が高くなる大規模研究では潜在的誤分類の影響に注意が必要であることが明らかになった。

結果の一例として、第3の変数X2が2値変数で、多い方の分類の割合が0.8であり、曝露変数X1と強い関連(オッズ比が3倍)を仮定した場合の結果を図1(下図左)に示す。横軸は、X2と疾患との関連の強さをオッズ比で推定した場合に生じるバイアスの大きさ、縦軸がX1のオッズ比のバイアスの大きさである。この状況では、X2に負のバイアスが生じるとX1へのバイアスが大きくなることが示されていた。

帰無仮説のもとで、differential な誤分類が潜在カテゴリー分類で生じている状況において曝露の効果をオッズ比として推定した場合に生じるバイアスの大きさを理論的に定量したところ、比較的影響が少ない状況でも0.1程度のバイアスが起これるため、1.1や1.2倍のオッズ比を精度よく推定したい大規模なケース・コントロール研究などを行う場合においては、誤分類が結論に及ぼす影響は大きい可能性が示唆された。また、説明変数自体が誤分類と関連がない場合においても、differential な誤分類の原因となる未測定の変数や疑似変数の存在が疑われる場合においてはバイアスが少なからず生じる可能性も示されたことから、オッズ比を推定する際には、アウトカムの分類の精度に影響する要因をよく検討しておく必要があると考えられた。

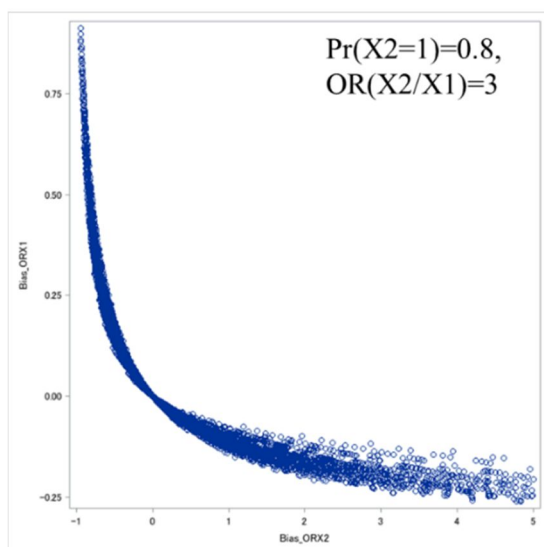


図1: オッズ比推定値への影響例

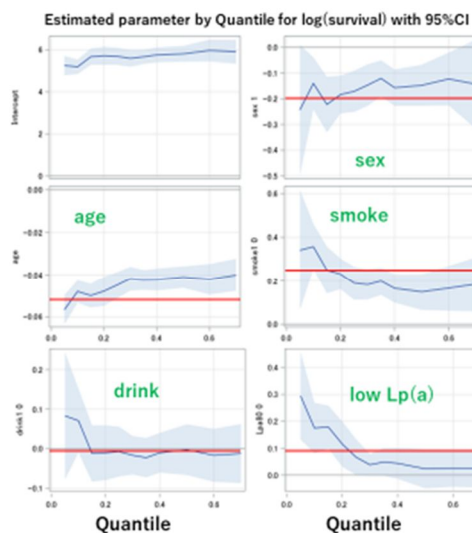


図2: 各リスク因子の生存時間への影響

(2) 長期追跡コホート研究データへの適用

JMS コホート研究のベースライン情報の長期追跡への影響について検討した。全死亡をアウトカムとし場合の性別、年齢、喫煙状況、飲酒状況、低Lp(a)血症について分位点回帰生存時間解析を行った結果を図2(上図右)に示す。赤線は参照線としてワイブル回帰によるパラメータ推定値を示した。その結果、年齢は死亡リスクが増大し続ける一方で、過去の研究において10年程度の追跡でリスク因子と結論づけていた要因の多くについて、長期追跡になると効果の減衰傾向が観察され、長期追跡でとくにハザード比などで要約された結果については慎重に解釈する必要があること、また、より適切な方法論の検討の必要性が示唆された。ただし、生存時間解析における分位点回帰モデルの応用については、アウトカムについて、ある程度のイベント数が見込まれないと推定が不安定となり、応用できる状況が限定される可能性はあると考えられた。実際のコホートデータ解析においても、イベント数の少ないアウトカムについては解釈が難しい可能性が示唆された。

また、コホート研究データが有する死因情報を整理し、がん死亡・循環器疾患死亡・肺炎による死亡・その他の原因による死亡の4種類に分類し、競合リスク生存時間解析を用いたリスク推定を行ってそれぞれの死因への喫煙の影響を確認した。さらに、本研究のモチベーションとなった我々の先行研究ですでに報告していた、疾患による誤分類傾向の違いについても検討モデルに組み込み、さまざまな設定を想定した感度分析を実施した。とくに、高齢者の肺炎死亡についての誤分類の検討においては、先行研究の結果から50%未満の低い感度を想定し、先行研究での信頼区間に相当する20%程度~70%程度まで動かして推定値の変動を確認した。長期観察データにおいて対象者が高齢化していたコホートデータでも、ベースラインの年齢ではなく死亡時年齢に依存する可能性のある同程度の誤分類が起こりえたことが想定でき、より現実的な分析が可能になった。研究方法(1)で検討したオッズ比モデルの3つ以上の競合リスクをもつ生存時間解析への拡張、および感度分析を提案する際の競合リスクの種類や現実的に考慮可能な数についてはさらなる整理が必要であり、今後の検討課題である。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 田中紀子、三重野牧子
2. 発表標題 潜在アウトカムの誤分類が及ぼすオッズ比の推定値への影響
3. 学会等名 第32回日本疫学会学術総会
4. 発表年 2022年

1. 発表者名 Makiko N Mieno, Noriko Tanaka, Motoji Sawabe, Tomio Arai, Shizukiyo Ishikawa
2. 発表標題 Censored quantile regression model for competing events with long follow-up period
3. 学会等名 40th Annual Conference of the International Society for Clinical Biostatistics (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
研究分担者	田中 紀子 (Tanaka Noriko) (10376460)	国立研究開発法人国立国際医療研究センター・その他部局等・医学統計研究室長 (82610)	
研究分担者	新井 富生 (Arai Tomio) (20232019)	地方独立行政法人東京都健康長寿医療センター（東京都健康長寿医療センター研究所）・東京都健康長寿医療センター研究所・研究員 (82674)	

6. 研究組織（つづき）

	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
研究分担者	沢辺 元司 (Sawabe Motoji) (30196331)	東京医科歯科大学・大学院医歯学総合研究科・教授 (12602)	
研究分担者	石川 鎮清 (Ishikawa Shizukiyo) (70306140)	自治医科大学・医学部・教授 (32202)	

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