

令和 5 年 5 月 29 日現在

機関番号：14401

研究種目：若手研究

研究期間：2020～2022

課題番号：20K19754

研究課題名（和文）ネット生存率に対する多重頑健推測法の開発と応用

研究課題名（英文）Development of multiple robust inference procedures for net survival

研究代表者

小向 翔（Komukai, Sho）

大阪大学・大学院医学系研究科・助教

研究者番号：70794543

交付決定額（研究期間全体）：（直接経費） 2,400,000円

研究成果の概要（和文）：がん登録データ解析で頻繁に使用されるネット生存率の推定において、他死因死亡の影響を調整するために使用される一般集団生存率のバイアス補正法を提案した。これは一般集団生存率にがん患者の影響が含まれている場合に発生するバイアスの補正方法である。また、がん登録データと一般集団間で他死因生存率の分布が異なる場合の補正法を提案した。

過剰ハザードモデルのEMアルゴリズムに基づく推測法の理論的正当化を行い、漸近分散推定量を提案した。これらの研究を踏まえ、ネット生存率の多重頑健推測法の提案を行った。過剰ハザードモデルに関する研究は統計雑誌に掲載された。それ以外の研究成果は現在投稿中1編、投稿準備中2編である。

研究成果の学術的意義や社会的意義

がん登録データはがん対策の評価を行うために重要な役割をになっている。実際のがん登録データ解析の結果からがん対策の見直しが行われる。本研究はがん登録データ解析において使用される一般集団生存率に対する根本的な問題を議論した研究であり、この問題に対する初めての数学的に厳密な方法である。また、過剰ハザードモデルの推測法の理論的正当化は方法論の発展に重要な議論であり、実際には、この理論を応用した多重頑健推測法の理論的正当化が可能となった。多重頑健推測法ではモデルの誤特定リスクを軽減し、より数学的仮定に依存しない推測を可能とし、これは実データ解析と方法論研究の双方に重要な結果である。

研究成果の概要（英文）：In the estimation of net survival used to analyze the cancer registry data, the population survival rates are used to adjust for the impact of other cause death. I proposed bias correction methods for general population survival rates. One is a method for correcting the bias that arises when the general population survival rates are affected the cancer patients. And the other is a correction method for when the distribution of non-cancer survival rates differs between cancer registry data and the general population.

I justified the mathematical theories of an EM-based estimation method for the excess hazard model and proposed an asymptotic variance estimator. Based on these studies, I proposed a multiple robust estimation method for net survival rates. Research on the excess hazard model has been published. Other research results are submitted (1 manuscript) and prepared for submission (2 manuscripts).

研究分野：医学統計学

キーワード：がん登録 ネット生存率 一般集団生存率 多重頑健推測法

1. 研究開始当初の背景

地域がん登録データ解析では、がん患者の生存率を推定することが主要な目的の一つである。しかしながら、地域がん登録データの持つ“死因に関する情報が未観測”であるという特有の特徴から、通常の生存時間解析手法を用いると、がん以外の原因で死亡した症例もがん死亡として扱われ、がんの予後を適切に評価できないという問題が生じる。これにより地域がん登録データ解析では“ネット生存率”と呼ばれる指標が予後評価指標として頻繁に使用されている。ネット生存率の推測法はがん患者集団の生存率を他死因生存率で補正することで構成される。代表的な推測法として、数理モデルを仮定しないノンパラメトリックな推測法が存在するが、この方法は非常に強い数学的な仮定を要求する。その仮定の一つが、打ち切り発生までの期間が共変量に依存しないという仮定である。しかしながら、実際の解析上の問題として、共変量によって打ち切り期間の分布が変化するという問題が頻繁に議論されており (Hakulinen 1982; Kodre and Perme 2013) この場合にはノンパラメトリック法の妥当性が担保できない。この問題は数理モデルを用いた推測を行うことで解決可能である。打ち切り分布に対して数理モデルを仮定する方法や、がん生存期間または任意の理由で死亡するまでの生存期間に数理モデルを仮定する方法である。これらの推測法はいずれもモデリングが適切である場合に妥当な推測となる。二重頑健推測 (DR)法 (Komukai and Hattori 2017) は、打ち切り分布と生存期間にそれぞれモデルを仮定した方法を組み合わせ、二種類のモデルのうち少なくとも一つが適切である場合に妥当となる方法であり、“二種類のモデルのうち少なくとも一つの適格性を要求する” という意味でモデルの不適格性に対して二重頑健性を持つ。

これらの方法で仮定する数理モデルは、変数間の真の関係を正しく特定する必要がある。実際のデータ解析では仮定した数理モデルのデータへの適合度評価 (Lin et al. 1993) により、誤特定のリスクを最小限に抑えることが可能であるが、真の関係を正しく特定した適切な数理モデルを使用しなければ妥当な推定とならないという問題の根本的な解決には至っていない。

2. 研究の目的

本研究は多数の候補モデルの同時使用を許容することで数理モデルの誤特定問題の解消を試みる研究であり、生存時間解析分野においても新規性が高く、この問題に対する世界的に先行した研究である。実際の地域がん登録データ解析においても、より実データに適合する柔軟な新しい統計解析手法となり、地域がん登録に基づくがんの実態把握において大きな貢献となると期待できる。

3. 研究の方法

以下の3つのテーマを行うことで、目的の達成を目指した。

多重頑健推測法の開発 : Komukai and Hattori (2017)では、打ち切り期間に対する数理モデルと任意の生存期間に対する数理モデルを同時に使用する DR 法を構成した。DR 法は二種類の観測期間に関するそれぞれ単一の数理モデルの使用に留まっており、どちらか一方のモデルを正しく特定しなければならぬという制約がある。本研究では、経験尤度の概念を用いてこの制約を解消した推測法を構築する。また、Komukai and Hattori (2017)では利用していないがん生存期間に対する複数の数理モデルも同時に使用し、複数の変数に対する複数の数理モデルに基づいた、よりデータに適合する柔軟な多重頑健推定量を提案する。

多重頑健推測法と他の推定量の数学的性質の評価 : 多重頑健推測法は、複数の数理モデルの同時使用によりモデルの誤特定リスクの改善が期待されるが、一方で、モデルの多数使用により、推定量の分散が大きくなる (有効性を持たない) 可能性がある。因果推論分野の議論では、二つの正しいモデルに基づく DR 法の有効性が示されている。ネット生存率の推定でも同様の性質が成り立つか、多重頑健推測法が有効性を持つか、また、有効性を持つ場合の条件は何かなどを検討することは方法論の精緻化や実データ解析への適用を行う上で重要な課題である。本研究では DR 法や多重頑健推測法の有効性とその条件を検討し、有効性の観点や既存のモデル選択基準などの観点からモデルの最適な個数の決定方法の提案を行う。

多重頑健推測法を容易に実装可能な統計解析ツールの開発 : 本研究における一連の提案手法は複雑な構造を持つことが予想される。実データへの広い適用を可能とするため、無料で使用可能な統計解析ソフト R で実装し、プログラムを公開することで多くの研究者の利用を可能とし、医学研究へ貢献する。

4 . 研究成果

地域がん登録データ解析において他死因による死亡を調整するために使用される一般集団生存率の不完全性に対する議論を行った。この議論により、これまでは生存時間分布に対するモデルと打ち切り分布に対するモデルの二種類のモデルをターゲットとしていたが、さらに一般集団生存率に対するモデルの必要性が明らかとなった。このことを受けて、一部方向を修正し、初めに一般集団生存率に対するモデルを使用した推測法を検討した。

一般集団生存率は、がん以外の他死因死亡の影響を調整するために必要な他死因生存率情報として使用されるにも関わらず、がん患者やがん死亡者も含んだコホートから推定される。このコホートに対する潜在的な生存時間変数を定式化し、ネット生存率に与える影響を評価した。この研究成果は現在、統計学の専門誌に投稿中であり、掲載には至っていない。また、一般集団生存率を抽出するために使用される共変量は、年齢や性別、暦年などの極めて限定された人口動態変数であり、がん登録データで観測されている共変量の一部である。他死因死亡に影響を与える他の共変量がある場合、調整すべき他死因生存率を一般集団生存率で置き換えるとバイアスが生じる。この不完全な一般集団生存率の問題を解決する方法を与えた。この研究に関しては、論文の主要部分は構築できており、数値実験による性能評価及び実データへの適用例まで完了し、投稿準備中である。

これらの一般集団生存率に関する研究により、がん登録データ解析の根本的問題に対する手法の充実化を図ってきた。一方、がん登録データ解析で頻繁に生じるモデルの誤特定の問題に対して頑健な推測法を与えるのが本研究の課題である。これまでは打ち切り分布に対するモデルとがん生存時間分布に対するモデリングのみに着目していたが、他死因生存時間分布や任意の死亡に対する生存時間分布へのモデリングなど、複数種類のモデリングが考えられる。これらのいくつかを含んだ多重頑健推測法への拡張には、がん登録データの持つ死因が未観測という根本的な問題が妨げとなっていることが明らかとなった。

上述の一連の研究から、当初の目的に立ち返り、打ち切り分布、がん生存期間、及び任意の死亡による生存期間の三種類の分布に対するそれぞれ複数のモデリングを許容可能な多重頑健推測法の提案を検討した。多重頑健推測法では、がん生存期間に対するモデリングを同時に考えることを検討しているため、この推測に関する数学的性質を明らかにする必要があった。そこで、ネット生存率に対するCox回帰モデルのEMアルゴリズムを用いた推測法(Perme et al. 2009)に対する識別性や一致性、漸近正規性などの理論的正当化、および漸近分散の導出を行った。本研究成果は統計雑誌に掲載された。

この結果から打ち切り分布、任意の死亡に対する生存時間分布、がん死亡に対する生存時間分布の三種類の生存時間に対するモデリングのみに焦点をあてた多重頑健推測法の提案を目指した。経験尤度法に基づく方法とChen (2013)による回帰に基づく方法を検討し、回帰に基づく方法でより安定した結果が得られた。提案法の一貫性が示されたが、漸近正規性などの理論的正当化や詳細な数値実験による推定値の挙動の調査が今後の課題である。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 Komukai Sho, Hattori Satoshi	4. 巻 -
2. 論文標題 Asymptotic justification of maximum likelihood estimation for the proportional excess hazard model in analysis of cancer registry data	5. 発行年 2023年
3. 雑誌名 Japanese Journal of Statistics and Data Science	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s42081-023-00190-6	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Hattori Satoshi, Komukai Sho, Friede Tim	4. 巻 41
2. 論文標題 Sample size calculation for the augmented logrank test in randomized clinical trials	5. 発行年 2022年
3. 雑誌名 Statistics in Medicine	6. 最初と最後の頁 2627 ~ 2644
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/sim.9374	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Okubo Masashi, Komukai Sho, Izawa Junichi, Aufderheide Tom P., Benoit Justin L., Carlson Jestin N., Daya Mohamud R., Hansen Matthew, Idris Ahamed H., Le Nancy, Lupton Joshua R., Nichol Graham, Wang Henry E., Callaway Clifton W.	4. 巻 79
2. 論文標題 Association of Advanced Airway Insertion Timing and Outcomes After Out-of-Hospital Cardiac Arrest	5. 発行年 2022年
3. 雑誌名 Annals of Emergency Medicine	6. 最初と最後の頁 118 ~ 131
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.annemergmed.2021.07.114	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Izawa Junichi, Komukai Sho, Nishioka Norihiro, Kiguchi Takeyuki, Kitamura Tetsuhisa, Iwami Taku	4. 巻 181
2. 論文標題 Outcomes associated with intra-arrest hyperoxaemia in out-of-hospital cardiac arrest: A registry-based cohort study	5. 発行年 2022年
3. 雑誌名 Resuscitation	6. 最初と最後の頁 173 ~ 181
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.resuscitation.2022.11.008	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Okada Yohei, Komukai Sho, Kitamura Tetsuhisa, Kiguchi Takeyuki, Irisawa Taro et al. and the CRITICAL Research Group Investigators	4. 巻 9
2. 論文標題 Clustering out of hospital cardiac arrest patients with non shockable rhythm by machine learning latent class analysis	5. 発行年 2022年
3. 雑誌名 Acute Medicine & Surgery	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/ams2.760	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Okada Yohei, Komukai Sho, Kitamura Tetsuhisa, Kiguchi Takeyuki, Irisawa Taro, et al.	4. 巻 86
2. 論文標題 Clinical Phenotyping of Out-of-Hospital Cardiac Arrest Patients With Shockable Rhythm Machine Learning-Based Unsupervised Cluster Analysis	5. 発行年 2022年
3. 雑誌名 Circulation Journal	6. 最初と最後の頁 668 ~ 676
掲載論文のDOI (デジタルオブジェクト識別子) 10.1253/circj.CJ-21-0675	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計3件 (うち招待講演 0件 / うち国際学会 1件)

1. 発表者名 小向翔
2. 発表標題 がん死亡を補正した生命表に基づくがん登録データ解析
3. 学会等名 日本計量生物学会年会
4. 発表年 2023年

1. 発表者名 小向翔
2. 発表標題 地域がん登録データ解析における不完全な一般集団生存率に基づくネット生存率の推定
3. 学会等名 2021 年度統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 Sho Komukai
2. 発表標題 Doubly robust inference procedure for relative survival ratio in population-based cancer registry data
3. 学会等名 41st Annual Conference of the International Society for Clinical Biostatistics (国際学会)
4. 発表年 2020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			
英国	LSHTM			
米国	University of Pittsburgh			