

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 10 日現在

機関番号：82602

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500365

研究課題名(和文) イベント発現時間が正確に特定できない場合の臨床試験のデザインに関する統計学的研究

研究課題名(英文) Statistical research on design of clinical trials where the exact time to event can not be observed

研究代表者

西川 正子 (Nishikawa, Masako)

国立保健医療科学院・研究情報支援研究センター・上席主任研究官

研究者番号：50373395

交付決定額(研究期間全体)：(直接経費) 3,900,000円、(間接経費) 0円

研究成果の概要(和文)：イベント発現までの時間が正確に特定できず区間打ち切りデータとして得られる場合に従来用いられている解析方法、即ち観察打ち切り区間に対してある1つの(代表となる)時点をあたかもその時点にイベントが起こったかのように扱う1点代入法や、予備研究においてわかったZaoの方法が誤った評価結果を出す現象について詳細な分析を行った結果をまとめ、国際学会で発表した。Zaoの検定方法について、不合理な点の改良を検討し、改良した検定統計量の合理性を評価した。その過程で汎用ソフトRも利用し、パッケージ(IntCox)自体の挙動の不安定さが発覚した。これを救済する方法を検討し結果をまとめ、国際学会で発表する予定である。

研究成果の概要(英文)：As common practice, if the exact time of occurrence is not known due to interval-censoring, a single point imputation (e.g. right-point imputation) for the censoring interval is most often done as though it was the exact time of event and the standard survival analysis are performed. On the other hand, there are several corresponding methods to interval-censored data. Zao's method is one of them. It was found that Zao's method does not maintain type I error probability in our preliminary research. We explored characteristics of single point imputations and Zao's method deeply, and reported the results at an international conference. We tried to improve Zao's method. While using R to evaluate an improved statistics, we found unstable performance of a package (IntCox). We consider some remedies for 'IntCox' and will report the results at an international conference.

研究分野：総合領域

科研費の分科・細目：情報学・統計科学

キーワード：医薬生物 ゲノム統計解析 医療 臨床試験

1. 研究開始当初の背景

根拠に基づく医療 (Evidence based medicine) の大切さが認識され、特に、承認申請のための治験は臨床試験の実施の基準 (GCP、厚生労働省令) という法的な規制があり、臨床試験の計画・実施・管理・評価も厳格になされつつある。科学面では、統計ガイドラインに臨床試験のための統計的原則が示されている。そこでは、偏りを極力抑えること、および第1種の過誤を増大させないことを求めている。第1種の過誤を増大させる原因の1つとして検定の多重性があり、多重性の調整が必要となる。そのため、統計ガイドラインでは「臨床試験の主要な評価項目は1つに絞る」ことを薦めている。そして多くの場合、治療の有効性に関する評価項目が一義的に採用され、安全性は相対的に重視されない傾向にある。その結果、例えば、新型インフルエンザの治療薬タミフルのように承認後に重大な安全性の問題が出現する場合も見られる。本来は、有効性と安全性のバランスを医学的に勘案し、これらを定量的に統合評価する必要があると指摘されていた。しかし、約10年前はそれに答える有効な方法がほとんどなかった。

この状況を改良するために、研究代表者らは平成18年度より「医学的な重要性を加味した臨床試験のデザインと評価方法に関する統計学的研究」を実施し、複数の評価項目から合成されるエンドポイントについてそれぞれの項目の医学的重大さを反映した検定方法を考案し、国際学会等で発表した。さらに平成20年度からは、その研究成果を発展させて、治療方法の有効性のみならず安全性をも同時に精度良く検定し得る評価方法を考案し、国際学会等で発表した。

以上開発した評価方法は、従来の臨床試験における評価方法の問題点を解決するものであるが、そこで用いられるイベント発現までの時間は、実際の臨床試験では正確に把握できないことも少なくない。例えば、胃がんの臨床試験では、従来、主要評価項目は生存時間であったが、近年は、主要評価項目は生存時間ではなく無病生存時間(がんの再発又は死亡のいずれか早い方をイベントと定める)や無増悪生存時間(がん細胞の増悪又は死亡のいずれか早い方をイベントと定める)が多用されるようになった。患者の死亡よりも再発や増悪が先に起こる場合が多く、がんが再発しているか否か、または患者のがん細胞が増悪しているか否かは検査して初めてわかる。したがって、がん細胞の増悪時間は増悪していない最後の検査時点から増悪が認められた最初の検査時点までの区間のある時点ということしか分からない。再発時間も同様である。このようなイベント発現までの時

間区間で与えられるようなデータを区間打ち切りデータと呼び、その区間を観察打ち切り区間と呼ぶ。

区間打ち切りデータに基づいて、がん治療における増悪時間を評価する場合、各被験者の増悪発現時間として観察打ち切り区間の1点(例えば左端、中点、右端の値)を代入して既存の解析手法により治療効果を評価する簡便法と、観察打ち切り区間をそのまま用いて評価する Zhao らの方法がある。研究代表者らの予備研究では、Zhao らの方法は正攻法ではあるが、重大な欠陥もあり信頼できる結果が必ずしも得られないことが判明した。一般には、1点代入法が用いられ、中でも右端代入が慣例として用いられることが多い。右端代入法では、増悪時間の評価においては効果が過大評価になるが、研究代表者らの研究により、1点代入により Kaplan-Meier 推定量を用いた集団としての生存率の推定では結果の挙動に不合理な点が見られることが判明している。1点代入法では生存率の標準誤差が過小評価になり、観察打ち切り区間が広い場合や被験者によって異なる場合には推定値にバイアスが入ること、さらに研究代表者らの予備研究において、右端代入を用いた Log-rank 検定では結果に不合理な挙動が見られることなどが判明している。

上のような区間打ち切りデータに関して、臨床試験の検査時点が定期的に計画されている場合でも、各検査時点で前後に何日かの受診許容幅を持つのが普通であり、指定された検査日に来院しなかったりして、被験者により実際の検査日が異なる場合や、抗がん剤治療などでは治療方法により治療サイクルが異なるため比較する治療の間で検査時点が異なる場合がある。このような場合、観察打ち切り区間も被験者により異なることになり、評価結果に重要な影響を及ぼす。

2. 研究の目的

有効で安全な最適な治療方法を確立するためには、その有効性・安全性を確認する信頼性の高い臨床試験の開発が必須である。イベント発現までの時間の多くの解析方法においては、検査によってイベント発現時間が特定できると仮定された。しかし通常の臨床試験では、イベント発現時間は正確に特定できず、区間打ち切りデータとして与えられる場合も少なくない。そこで本研究では、イベント発現の区間打ち切りデータに基づいて、有効性と安全性のバランスを勘案した信頼性の高い臨床試験の評価方法を確立する。

3. 研究の方法

臨床試験結果が区間打ち切りデータとして与えられる場合の治療法の有効性・安全性の評

備手法について予備研究を行った結果、区間打ち切りデータの観察打ち切り区間をそのまま用いる Zao の方法では評価結果の信頼性に疑問点があることが判明し、他方、一般的に用いられている区間打ち切りデータの 1 点代入法についても、取り扱いは簡便ではあるが、誤った評価結果の出る可能性があることが判明した。そこで、本研究では評価結果の信頼性を保証するため、以下の手順で区間打ち切りデータの合理的な解析方法を検討した。

(1) . 区間打ち切りデータを取り扱う正攻法として、区間打ち切りデータを観察打ち切り区間として以後の解析を実施する方法で、Zao 以外の検定方法について調査した。

(2) . がんの治療において、抗がん剤の投与サイクル、放射線治療サイクルや検査時点などについての最新情報を、文献や医学系の国内学会等で調査した。

(3) . これらの調査結果をもとに 1 点代入法および Zao の検定方法、比例ハザードモデルを用いた方法についての simulation デザインおよび解析プログラムの検討を行った。

(4) . 本研究の研究期間内に、本研究と類似の研究の成果が成書 (Interval-censored time-to-event data: methods and applications, 2013, edited by DG Chen, J Sun, KE Peace.) として出版された。1 点代入法および Finkelstein の検定方法の simulation による検討がされていたので、本研究と成書の simulation デザインの相違点を検討した。また、成書では汎用ソフト R のあるパッケージ (IntCox) の利用を推薦していた。本研究での simulation デザインのもとでも R のこのパッケージを利用し結果の検討を行った。

(5) . 臨床試験の医療経済的側面からの評価を考え、汎用されている「看護必要度」を利用した重症期間を取り入れることに着目した。重症期間は区間打ち切りデータとなり疾患によって重症か否かの観測間隔を変えても精度を保ち記録を簡素化できる可能性があると考えられるので、重症期間の現状について調査検討を行った。

4 . 研究成果

(1) . 従来用いられている、観察打ち切り区間に対してある 1 つの (代表となる) 時点をあたかもその時点にイベントが起こったかのように扱う 1 点代入法や、予備研究においてわかった Zao の方法が誤った評価結果を出す現象について詳細な分析を行った結果をまとめ、国際学会で発表した。

(2) . 観察打ち切り区間を用いて解析を行う Zao 以外の検定方法として、比例ハザードモデルおよびそれに関連するモデルを用いた方法 (Finkelstein 法、Kim 法、Betensky 法、

Goetghebeur 法等) について調査を行った。がんの治療において、抗がん剤の投与サイクル、放射線治療サイクルや検査時点などについての最新情報を、文献や医学系の国内学会などで調査した。これらの調査結果をもとに 1 点代入法および Zao の検定方法、比例ハザードモデルを用いた方法についての simulation デザインおよび解析プログラムの検討を行った。ログランク検定及び比例ハザードモデルを用いた解析方法の考え方を考慮にいれて、区間打ち切りデータを用いる Zao の検定方法について、不合理な点の改良を検討し、改良した検定統計量の合理性を評価した。

(3) . 本研究の研究期間内に出版された、本研究と類似の研究の成果をまとめた成書 (Interval-censored time-to-event data: methods and applications, 2013, edited by DG Chen, J Sun, KE Peace.) について検討と考察を行った。1 点代入法および Finkelstein の検定方法の simulation による比較がされていたので、本研究と成書の simulation デザインの相違点を検討した。また、成書では汎用ソフト R のあるパッケージ (IntCox) の利用を推薦していた。本研究での simulation デザインのもとでも R のこのパッケージを利用し結果の検討を行った。R のパッケージ自体の挙動の不安定さが発覚し、これを救済する方法を検討した。結果をまとめ、国際学会で発表する予定である。

(4) . 臨床試験の医療経済的側面からの評価を考え、汎用されている「看護必要度」を利用した重症期間を取り入れることに着目した。症状が「重症」から「非重症」に変わること、又はその逆をイベントと定義すれば、重症期間は区間打ち切りデータとなる。疾患によって重症か否かの観測間隔を変えても精度を保ち記録を簡素化することが可能であると考えられるので、重症期間の現状について調査検討を行った。

(5) . 本研究に関連する方法を用いた疫学研究や臨床試験についての発表を行った。

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 2 件)

(1) Tetsuya Kawahara, Masako Nishikawa, Tetsuya Furusawa, Tetsuya Inazu, Gen Suzuki. Atorvastatin, Etidronate, or Both in Patients at High Risk for Atherosclerotic Aortic Plaques: a Randomized Controlled Trial. [Circulation](https://doi.org/10.1161/CIRCULATIONAHA.113). 2013,11;127(23):2327-35. doi: 10.1161/CIRCULATIONAHA.113.

(2) Misa Adachi, Kazue Yamaoka, Mariko Watanabe, Masako Nishikawa, Itsuro Kobayashi, Eisuke Hida and Toshiro Tango. Effects of lifestyle education program for type 2 diabetes patients in clinics: a cluster randomized controlled trial. BMC Public Health 2013, 13:467.

[学会発表](計 10件)

- (1) Kazue Yamaoka, Misa Adachi, Mariko Watanabe, Masako Nishikawa, and Toshiro Tango. Sensitivity analysis for a cluster randomized trial on effects of lifestyle education program for type 2 diabetes in clinics. 32nd Annual conference of the International Society for Clinical Biostatistics, Ottawa, Canada. 2011.8.
- (2) 安達美佐, 山岡和枝, 渡辺満利子, 西川正子, 飛田英佑, 丹後俊郎. 糖尿病のための個別生活習慣改善プログラムによる教育効果の評価. 第70回日本公衆衛生学会, 秋田市 2011.10.19-21. 日本公衆衛生雑誌 58,10:p480.
- (3) 安達美佐, 山岡和枝, 渡辺満利子, 小林逸郎, 西川正子, 飛田英佑, 丹後俊郎. クラスター無作為化比較試験による2型糖尿病のための生活習慣改善プログラムの介入効果の評価. 第22回日本疫学会, 2012.
- (4) Kawaguchi, Hida, Nishikawa, Yamaoka, and Tango. Comparison of Two-Sample Tests for Partly Interval-Censored Survival Data. XXVIth International Biometric Conference. Kobe, Japan. 2012.8.
- (5) Adachi, Yamaoka, Watanabe, Nishikawa, Hida, and Tango. Effects of Lifestyle Education Program for Type 2. XXVIth International Biometric Conference. Kobe, Japan. 2012.8.
- (6) Nishikawa, Taguchi, Akazawa, and Takahashi. On the Adjustment of the cause-Specific Log-Rank Test in Competing Risks Data. XXVIth International Biometric Conference. Kobe, Japan. 2012.8.
- (7) Uehara, Nishikawa, Takahashi, Hida, Yamaoka and Tango. Mixture Models for Epileptic Seizure Counts. XXVIth International Biometric Conference. Kobe, Japan. 2012.8.
- (8) 西川正子, 田口和伸. Time to event data を評価尺度とする臨床試験の競合リスクに関わる最近の話題. 第40回日本行動計量学会, 新潟市 2012.9.13-15. 第40回日本行動計量学会特別セッション講演集別冊.

(9) 西川正子. 筒井孝子. 東野定律. 大冨賀政昭. 入院医療機関における処置と患者の状況の検討(1)-在院日数別看護必要度得点の推移-. 第72回日本公衆衛生学会, 津市 2013.10.23-25. 日本公衆衛生雑誌 60,10:p264.

(10) 筒井孝子. 東野定律. 大冨賀政昭. 西川正子. 入院医療機関における処置と患者の状況の検討(2)-認知症有無別看護必要度得点の比較-. 第72回日本公衆衛生学会, 津市 2013.10.23-25. 日本公衆衛生雑誌 60,10:p264.

[図書](計 2件)

(1) 西川正子. 臨床試験. In: 統計応用の百科事典(松原望, 美添泰人, 編集委員長). 東京: 丸善出版; 2011. p.550-551.

(2) 西川正子. 臨床試験における無作為割付け. In: 統計応用の百科事典(松原望, 美添泰人, 編集委員長). 東京: 丸善出版; 2011. p.552-553.

6. 研究組織

(1) 研究代表者

西川 正子 (NISHIKAWA MASAKO)
国立保健医療科学院・研究情報支援研究センター・ 上席主任研究官
研究者番号: 50373395

(2) 研究分担者

該当なし

(3) 連携研究者

丹後 俊郎 (TANGO TOSHIRO)
昭和女子大学
研究者番号: 70124477