

リアルワールドデータを利用した観察研究の利点と欠点

山本 陵平

Advantages and disadvantages of observational study using real world data

Ryohei Yamamoto, MD, PhD

抄 録

1990 年代以降に普及した evidence-based medicine においては、ランダム化比較試験 (RCT) によるエビデンスが重要視される。しかしながら、RCT で確認された効能 (efficacy) が、リアルワールドにおいても同様に確認されるとは限らないため、大規模なリアルワールドデータ (RWD) を利用した観察研究において、その効果 (effectiveness) を確認しなければならない。RWD を利用した観察研究は、さまざまなバイアスの影響を受けやすいため、傾向スコア法や操作変数法等の技術を活用して、適切にバイアスに対処しなければならない。

キーワード

リアルワールドデータ、バイアス、交絡、傾向スコア、操作変数

ABSTRACT

In evidence-based medicine since 1990's, randomized controlled trials (RCTs) have been regarded as the most valuable clinical studies. However, treatment efficacy in RCTs is not always observed in clinical settings in the real world. Large-scale real-world data (RWD) is very useful to confirm clinical effectiveness of treatments in the real world, which RCTs have suggested. Observational studies using RWD are very susceptible to various biases. To control these biases, scientists should use appropriate methods, including propensity score method and instrumental variable method.

Key words:

Real world data, Bias, Confounding, Propensity score, Instrumental variable

I. はじめに

近年、デジタル技術の進歩に伴い、医療領域においてもビッグデータの活用が著しく促進されている。電子カルテ等に日々蓄積されていく大規模なリアルワールドデータ (real world data: RWD) を活用することによって、実際の医療現場を研究領域とする観察研究

を実施することが比較的容易になってきた。RWD を利用した観察研究は、大規模ではあるが、さまざまなバイアスの影響を受けやすい。1990 年代以降に普及した evidence-based medicine において重要視されるエビデンスレベル¹⁾からすれば、ランダム化比較試験 (randomized controlled trial: RCT) には及ばない (図 1)。しかしながら、RCT のみでは EBM の実践は不可能であり、RWD を利用した観察研究と RCT が

大阪大学キャンパスライフ健康支援・相談センター保健管理部門

大阪大学医学系研究科健康増進医学講座行動健康医学

大阪大学医学系研究科腎臓内科学

Health and Counseling Center, Osaka University

Laboratory of Behavioral Health Promotion, Department of Health Promotion Medicine, Osaka University Graduate School of Medicine

Department of Nephrology, Osaka University Graduate School of Medicine

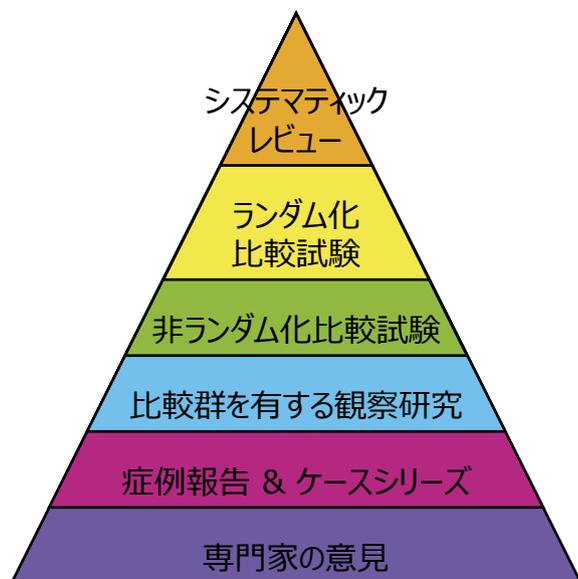


図 1 Evidence-based medicine ピラミッド¹⁾

相互に補いながらエビデンスを蓄積する必要がある。

本稿では、RCT によってもたらされたエビデンスがリアルワールドにもたらした弊害の一例を紹介することによって、RWD を利用した観察研究の重要性を示す。さらに、RWD を利用した観察研究において対処しなければならないバイアスのうち、特に交絡に注目し、代表的な交絡の対処法の特徴を概略する。

II. ランダム化比較試験の欠点と観察研究の利点

ミネラルコルチコイド受容体拮抗薬 (mineral corticoid receptor antagonist: MRA) は、アンジオテンシン変換酵素 (ACE) 阻害薬・アンジオテンシン II 受容体拮抗薬、 β 遮断薬とともに、心不全治療の主要薬剤である。MRA は、腎臓の遠位尿細管・集合管のミネラルコルチコイド受容体へのアルドステロンの結合を阻害することによって、尿中ナトリウムの排泄を促進すると同時に尿中カリウムの排泄を抑制する。したがって、MRA の副作用として、致死性不整脈の誘因となる高カリウム血症に注意が必要である。

Randomized Aldactone Evaluation Study (RALES) は、心不全患者において第一世代 MRA であるスピロラクトンの効能 (efficacy) を評価した無作為化二重盲検並行群間比較試験である²⁾。中等症以上 (New York Heart Association 分類 III ~ IV 度) の心不全患者 1,663 人を無作為にスピロラクトン (25 ~ 50 mg) 群とプラセボ群に割り付け、両群の死亡率が比較された。平均追跡期間 24 か月時点の中間解析に

おいて、スピロラクトンの死亡抑制効果 (相対リスク 0.70 [95% 信頼区間 0.60, 0.82]) が確認されたため、RALES 試験は早期終了になり、その研究結果が 1999 年 New England Journal of Medicine 誌に掲載された。治療継続困難な重症高カリウム血症は、スピロラクトン群 1.2%、プラセボ群 1.7% に認められ、両群でほぼ同等であった。以上の結果から、スピロラクトンは、中等症以上の心不全患者において、高カリウム血症のリスクを上昇させることなく、死亡リスクを低減させることが示された。

RALES の対象者は、その半数を中年層が占める比較的若い心不全患者 (65±12 歳) であり、高カリウム血症のリスクが高い腎不全患者 (血清クレアチニン >2.5 mg/dL) は除外されていた。RALES が示したのは、有害事象が発生しにくいと考えられる臨床背景を有する限定的な集団において、綿密に計画された治療計画に従った特殊な状況におけるスピロラクトンの効能である。リアルワールドにおいて、腎不全等のさまざまな臨床的背景を有する高齢心不全患者におけるスピロラクトンの効果 (effectiveness) は示されていない。一般的に、ランダム化比較試験では、有害事象が発生しやすい高齢者や腎不全患者は研究対象から除外される傾向が強い。したがって、そのような集団における介入の有効性および安全性は、リアルワールドを用いた観察研究が最もエビデンスレベルが高い研究デザインになる。

1999 年に RALES 試験の結果が公表された前後に、リアルワールドにおいて心不全患者におけるスピロラクトンの処方率と高カリウム血症の発症率を評価した大規模観察研究が、2004 年 New England Journal of Medicine 誌に発表された³⁾。1994 ~ 2001 年にカナダのオンタリオ州の 66 歳以上の高齢者約 130 万人の医療データベースを用いて、過去 3 年間に心不全の入院歴があり、ACE 阻害薬が処方されている高齢心不全患者を対象にした時系列分析 (1994 年対象者数 20,820 人、2001 年 32,283 人) を行い、スピロラクトンの処方率と高カリウム血症による入院率および入院死亡率の変遷を auto-regressive integrated moving average model (ARIMA) を用いて評価した。RALES の結果公表後に、スピロラクトンの処方率が大幅に上昇すると同時に (1994 年 3.4% → 2001 年 14.9%)、高カリウム血症による入院率 (1994 年 0.24% → 2001 年 1.10%) と死亡率の著明な上昇 (1994 年 0.03% → 2001 年 0.20%) が認められた。本研究結果は、スピロラクトンの処方によって致死的な高カリウム血症が誘発された可

能性を示唆しており、RALESとは異なる結果である。本研究においてスピロラクトンが投与された心不全患者の年齢分布は、RALES結果発表前で78.0±7.4歳、発表後で78.6±7.2歳であり、RALES試験よりもかなり高齢である。また、RALES試験では除外された腎不全患者も対象に含んでいる。このような対象者の違いが異なる結果に至った理由の一つかもしれない。

綿密に計画されたランダム化試験であるRALES試験と比較すると、本研究はさまざまなバイアスの影響を受けやすい観察研究であり、その内的妥当性は低い。一方、ランダム化比較試験の対象とはなりにくい対象を含む研究であり、その外的妥当性は高い。内的妥当性が高い研究は外的妥当性が低くなり、外的妥当性が高い研究は内的妥当性が低くならざるを得ない。1990年代以降に普及したevidence-based medicineにおいて、エビデンスレベルが高いとされるランダム化比較試験の結果が重要視されているが、その適応には限界がある。RWDを用いた観察研究は、ランダム化比較試験で得られた知見を補完する役割が期待される。しかしながら、観察研究はさまざまなバイアスの影響を受けやすいため、それらのバイアスの影響を小さくする工夫が必要である。

III. 交絡の対処法

バイアスは、選択バイアス (selection bias)、情報バイアス (information bias)、交絡 (confounding) の3種類に大別される⁴⁾。選択バイアスは、研究対象が母集団を代表していない場合に生じるバイアスである。情報バイアスは、データ収集時に生じるバイアスである。交絡は、アウトカムリスク因子であると同時に暴露因子に関連する交絡因子によって生じるバイアスである。統計解析によって対処することが可能なバイアスは交絡のみである。選択バイアスと情報バイアスは統計解析では対処することができず、研究デザインを立案する時点で対処しなければならない。

交絡への対処法は、統計解析による方法 (サブグループ、マッチング、回帰モデル、傾向スコア法、操作変数法等) と研究デザインによる方法 (ランダム化比較試験) に大別される。サブグループとマッチングは最も古くから行われてきた交絡の統計学的な対処法である。性別が交絡因子であるならば、男性あるいは女性のサブグループで解析を行う、あるいは性別によるマッチングで抽出した対象で解析を行えば、性別の交絡に対処することが可能である。サブグループ解析とマッチング解析による交絡の対処は非常に強力であ

るが、一度に多くの交絡因子を同時に対処することは困難であるため、活用できる場面は限られる。

ロジスティック回帰モデルやCox比例ハザードモデル等による回帰モデルは、観察研究において頻用されている交絡の統計学的な対処方法である。多数の交絡因子に対して一度に簡単に対処できるという利点がある。独立変数と従属変数の関連が直線的であると仮定する直線回帰モデルが主に用いられてきたが、近年ではrestricted cubic spline regression model等の非線形回帰モデルもしばしば利用される。それぞれの回帰モデルにはさまざまな前提条件があり、その前提条件が満たされない場合には十分に対処できない可能性があることに留意する必要がある。

傾向スコア (propensity score) 法は近年著しく普及した統計学的な対処法である。対象者が暴露因子を有する予測確率 (= 傾向スコア) を、暴露前の背景因子 (交絡因子を含む) を独立変数として組み込んだロジスティック回帰モデルを利用して算出する。その結果、ほぼ同レベルの傾向スコアを有する暴露群の対象者と非暴露群の対象者は、ほぼ同じ背景因子を有することになる。算出された傾向スコアは、マッチングや逆確率重み付け (inverse probability weighting) 等の手法を用いて、交絡因子の影響を調整することが可能である⁵⁾。傾向スコアでマッチングされた暴露群と非暴露群の背景因子を比較すると、まるでランダム化比較試験を実施したかのように両群の背景因子が同等になることから、準ランダム化 (quasi-randomization) あるいは偽ランダム化 (psuedo-randomization) とも呼ばれることもある。しかし、傾向スコア法は測定された背景因子のみを考慮した交絡の対処法であるため、未測定の交絡因子には全く対処していない点が、操作変数法やランダム化比較試験と大きく異なる (図2)。

傾向スコアに続いて、近年急速に普及している統計学的方法が操作変数 (instrumental variable) 法である。操作変数は、交絡因子と関連がなく、暴露因子を介してのみアウトカムに影響を与える変数である⁶⁾。治療介入の効果を評価する疫学研究では、これまで患者の居住地と医療機関の距離や医療機関の治療介入の実施率等が操作変数として取り扱われてきた。近年では、遺伝子工学の進歩によって、UK Biobank⁷⁾ に代表される大規模なゲノムワイド関連研究 (genome-wide association study: GWAS) が実現可能になった。GWASで同定された遺伝子多形群の中から操作変数として機能し得る遺伝子多形を抽出し、操作変数法を実施するMendelian randomization⁸⁾の報告が急増

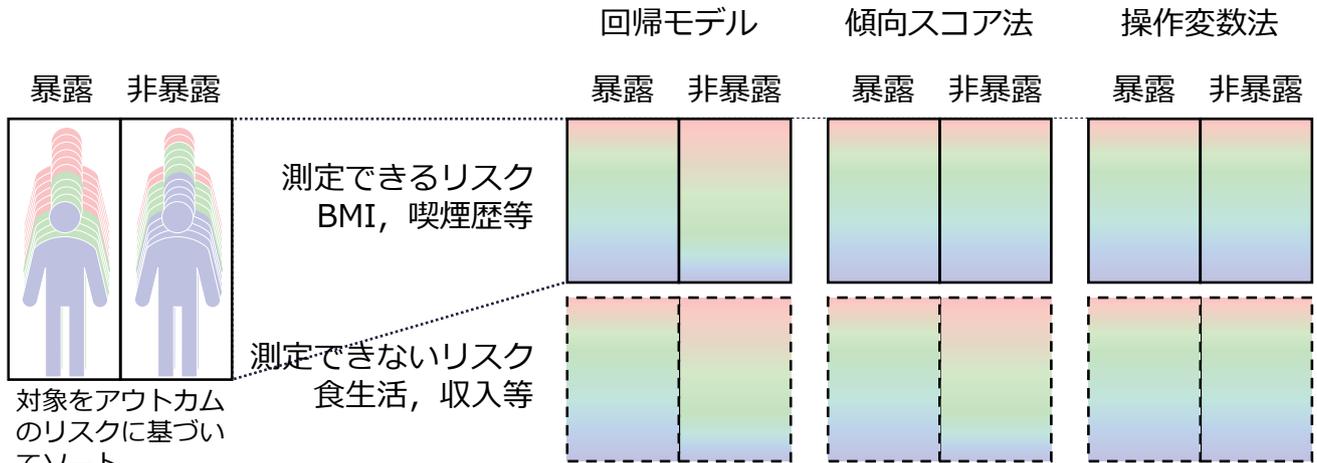


図2 回帰モデル, 傾向スコア法, 操作変数法による交絡の対処
BMI, body mass index

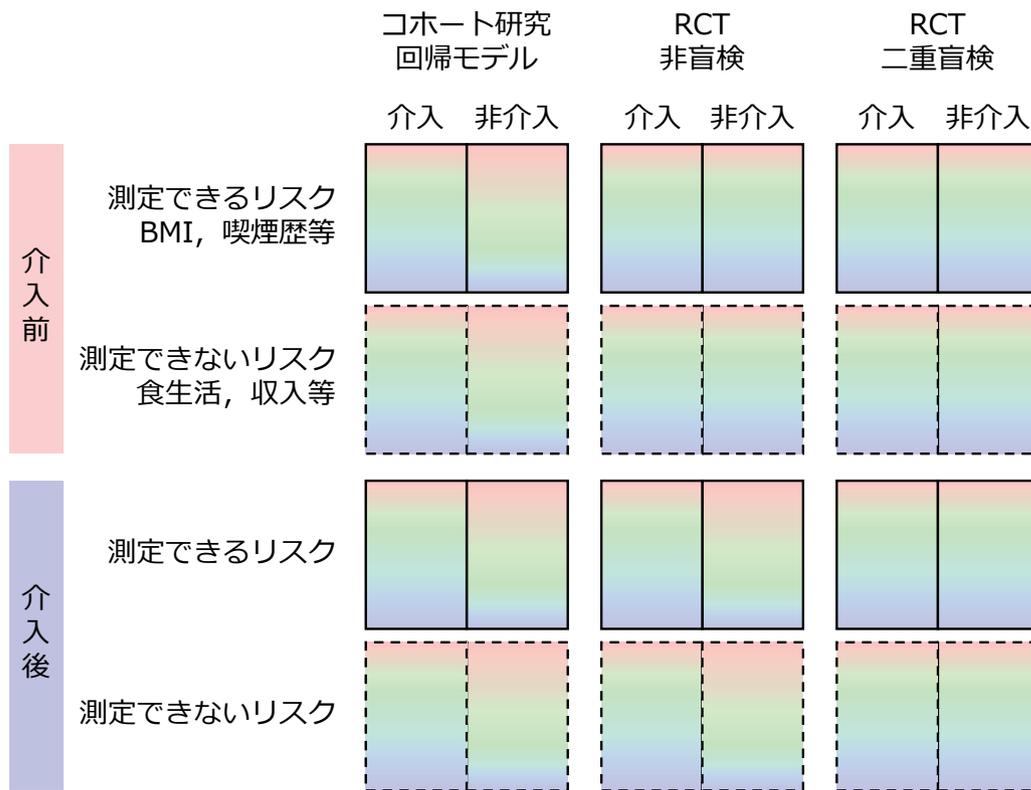


図3 コホート研究 (回帰モデル) とランダム化試験 (盲検, 二重盲検) による交絡の対処

している。操作変数法は、測定された交絡因子のみならず、未測定の交絡因子にも対処している点が、傾向スコア法と大きく異なる (図2)。しかし、操作変数に設定した変数が、本当に操作変数として機能しているかを証明する方法がないため、あくまでも理論上は未測定の交絡にも対処しているはずであると言わざるを得ない。

ランダム化比較試験は、研究デザインレベルで交絡に対処する方法である。研究デザイン的にはコホート研究の一種であり、対象を介入 (暴露) 群と非介入 (非暴露) 群に無作為に割り付ける過程が追加されている。ランダム割り付けによって、介入前の背景因子 (交絡因子を含む) が介入群と非介入群でほぼ同等になる。ランダム割り付けだけでは、介入後の追跡期間中にお

ける両群の背景因子が均等にならない可能性があるため、盲検化が実施される。ランダム化二重盲検試験は、ランダム割り付けの結果(介入 vs. 非介入(プラセボ))を対象者と医療従事者の両者に知らせないため、バイアスの影響が小さく、エビデンスレベルが高い研究デザインである(図3)。一方、非常に綿密に計画され、前向きに実施される研究であるため、多大な労力と費用が必要である。

IV. 最後に

RCTは、最もバイアスの影響を受けにくく、内的妥当性が高い研究デザインであるが、その外的妥当性は低く、その結果の適応範囲は限定的であることに注意が必要である。近年大規模なRWDを用いた観察研究が可能になり、RCTの対象にされにくい高齢者や合併症を有する患者などを対象にした外的妥当性の高い研究を比較的容易に実施することができるようになった。しかしながら、RCTと比較して、観察研究はさまざまなバイアスの影響を受けやすく、内的妥当性は低いため、さまざまなバイアスに適切に対処しなければならない。交絡は、選択バイアスや情報バイアスと異なり、統計学手法によって対処できるバイアスであり、さまざまな対処法が活用されている。大規模なRWDを利用して、適切に実施された観察研究は、RCTでは検証が困難な医学領域における貴重なエビデンスをもたらす。今後RWDを利用した大規模観察研究とRCTが相互連携することによって、人類の健康増進に有用な医学研究がさらに発展していくことを期待したい。

文 献

- 1) Golden SH, Bass EB. Validity of meta-analysis in diabetes: meta-analysis is an indispensable tool in evidence synthesis. *Diabetes Care* 2013; 36: 3368-73.
- 2) Pitt B, Zannad F, Remme WJ et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* 1999; 341: 709-17.
- 3) Juurlink DN, Mamdani MM, Lee DS et al. Rates of hyperkalemia after publication of the Randomized Aldactone Evaluation Study. *N Engl J Med* 2004; 351: 543-51.
- 4) Delgado-Rodríguez M, Llorca J. Bias. *J Epidemiol Community Health* 2004; 58: 635-41.
- 5) Shiba K, Kawahara T. Using propensity scores for causal inference: Pitfalls and tips. *J Epidemiol* 2021; 31: 457-63.
- 6) Aso S, Yasunaga H. Introduction to instrumental variable analysis. *Annals of Clinical Epidemiology* 2020; 2: 69-74.
- 7) Bycroft C, Freeman C, Petkova D et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018; 562(7726): 203-9.
- 8) Sekula P, Del Greco FM, Pattaro C, Köttgen A. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol* 2016; 27: 3253-65.

著者連絡先：山本 陵平

〒560-0043 大阪府豊中市待兼山町1-17

Tel: 06-6850-6015

Fax: 06-6850-6040

E-mail: yamamoto.ryohei.ras@osaka-u.ac.jp