

令和元年度厚生労働科学研究費補助金  
(政策科学総合研究事業(統計情報総合研究事業))  
分担研究報告書

医学主要誌におけるリンケージデータを用いた研究の動向  
—JAMA・NEJM・BMJ・LANCETの2019年出版分のレビューから—

研究代表者 伊藤 智子 筑波大学医学医療系 助教

研究要旨

本研究では、医学領域におけるリンケージデータ活用の最近の状況について把握するべく、医学領域で広く読まれている雑誌に着目してレビューを行った。その結果、2019年の医学主要4誌(JAMA・NEJM・BMJ・LANCET)において、リンケージデータを用いた研究はメジャーな研究方法ではなかった。しかし各研究においてデータリンケージは、観察研究の弱みを軽減し、また実証可能な研究仮説の幅を拡張することに寄与しているとみられた。

A. 研究目的

本研究では、医学領域におけるリンケージデータ活用の最近の状況について把握するべく、医学領域で広く読まれている雑誌に着目してレビューを行った。

B. 研究方法

レビュー対象は、高いImpact Factorを誇る主要4誌(Journal of American Medical Association: JAMA, The New England of Medicine: NEJM, British Medical Journal: BMJ, The Lancet)の2019年発行分とした。検索語は使用せず、発行されたすべての原著論文を対象とし、研究で使用されているデータがリンケージデータである論文を選出した。

(倫理面への配慮)

本研究で用いるデータは、個人情報を含まない。

C. 研究結果

リンケージデータを用いた研究9文献

を抽出した。9文献のうちわけは、JAMA 5件(2019年原著全148件中3%)、LANCET 2件(全184件中1%)、NEJM 1件(全210件中0.5%)、BMJ 1件(全175件中0.6%)であった。以下、この9件の概要について説明する。

1. 頭部外傷による自殺のリスク<sup>1</sup>

デンマークで行われた研究であり、医療情報と公的死亡データのリンケージを用いて、頭部外傷の受診歴と自殺率との関連を検証した研究である。データリンケージには、個人番号を用いて行われた。700万人以上を対象とした後ろ向きコホート研究であり、その結果、頭部外傷の診断を受けていない人々と比較した頭部外傷患者の自殺リスクの調整後 Incidence rate ratio (IRR) は 1.90 (95% CI, 1.83-1.97) であった。

2. 脳卒中患者の死亡率<sup>2</sup>

オランダで行われた研究であり、全国的

な退院データベースと人口および死因データベースのリンケージを行い、18歳から49歳の若年性脳卒中患者の死亡リスクを明らかにした。データリンケージは、個人番号を用いて行われた。年齢、性別をマッチングさせた一般集団と比較した若年性脳卒中患者の標準化死亡率は 5.6 (95% CI, 5.3-5.9) であった。

### 3. 収入と平均余命<sup>3</sup>

ノルウェー国内の 300 万人以上を対象に、収入と平均余命および死因との関連を検証した研究である。この研究では個人番号を用いて、人口登録簿、個人納税者登録簿、死因登録簿、および国民教育データベースをリンケージし、個人レベルのデータセットを作成している。その結果、収入と余命には比例関係があり、2005年から2015年にかけてその差が拡大していることが明らかになった。

### 4. 貧困状態と医療サービス支出<sup>4</sup>

米国において、貧困状態と医療サービス支出との関連を検証した研究である。対象は社会保障番号、生年月日、性別を使い、メディケア、メディケイドの医療サービス支出情報をリンケージした。貧困状態の基準として、貧困地域に限定された家賃補助バウチャーおよび、米国連邦政府の住宅サポートプログラムである住宅補助バウチャーの受領情報を利用した。分析の結果、成人においてはバウチャー受領者と対照群の医療サービス支出には顕著な差は見られなかった一方で、未成年世代のバウチャー受領者は入院費支出の減少と強い相関関係が見られた。

### 5. 投薬量と脳症リスク<sup>5</sup>

カナダのオンタリオ州の高齢腎不全患者コホートを対象に、筋弛緩薬であるバクロフェンの服用量と脳症リスクとの関連を検

証した研究である。この研究では、オンタリオ州の the Institute for Clinical Evaluative Sciences (ICES) という非営利の研究機関によってデータ提供およびデータリンケージがなされている。ICESでは、研究、データ管理、臨床等の専門家が所属し、オンタリオ州の個人健康情報プライバシー法の目的で設立された組織である。データリンケージは、健康カード番号や姓、名、生年月日、性別、郵便番号といった個人情報の組み合わせを基に一意的機密 ICES 番号 (IKN) を新たに作成し、この IKN を使用してデータ間のリンケージを安全に行っている。この ICES によって提供された医療情報リンケージデータを分析した結果、バクロフェンの低服用服用群との比較における高用量服用群の脳症リスクのリスク比は 3.54 (95% CI, 2.24-5.59) であり、バクロフェンは脳症の発生率を増加させることを指摘しつつ、バクロフェン使用の利点とリスクとのバランスを取る必要があると結論付けている

### 6. アルコール摂取量と疾患リスク<sup>6</sup>

中国で行われた研究であり、アルコールの摂取パターンに関連する 2 種類の遺伝的変異体を持つ集団とそれ以外の集団との間で、アルコール摂取量と疾病の関連性の違いを検証している。この研究は 50 万人以上の医療リンケージデータを内包する China Kadoorie Biobank と死亡票、医療記録とのリンケージデータを使用した前向きコホート研究である。分析の結果として、コホート全体では適度なアルコール摂取が虚血性脳卒中と脳出血、心筋梗塞の発生率をすべて軽減したが、アルコール摂取関連遺伝子を持つ集団においては、アルコール摂取量と心筋梗塞リスクの関連はほとんどみられなかったと報告している。

### 7. 不妊治療によって出生した児の健康状

## 態<sup>7</sup>

フィンランドの調剤記録および出生記録のリンケージデータを用いて行われた研究である。調剤記録からわかる不妊治療に用いられる特定の薬剤の使用を、不妊治療の定義としており、薬剤を使用した対象の妊娠・出産の記録をリンケージしている。出生記録には生まれた児の健康状態も含まれており、これをアウトカムとしている。この研究では、観察期間内に出生した全国の0歳から14歳の65,723人を対象集団としており、対象集団の中で母親の不妊治療を曝露として児の健康状態の比較を行った。結果として、不妊治療そのものによる児の健康状態への影響は小さいことが明らかとなった。

## 8. 医療過誤の経験を持つ医師の診療業務傾向<sup>8</sup>

米国で、医療過誤の経験を持つ医師の診療業務内容の傾向を検証した研究である。この研究ではメディケアのプロバイダデータを用いて医療過誤の有無を特定し、医師の診療業務傾向は医師登録データをリンケージさせて特定している。このデータリンケージではまず医師が有する National Provider Identifier (NPI)で個別の医師のデータリンケージを行い、NPIがない場合はその他の複数の個人識別情報を組み合わせることで医師の診療情報の突合を行っているが、著者の一部がこの医師登録を扱う組織の一員であることが、このデータリンケージを可能にした有力な背景になっていると推察される。分析の結果、医療過誤の経験がある医師は診療業務のサイズダウンを行う傾向にあることが分かった。

## 9. 妊娠中のインフルエンザワクチン接種の児への影響<sup>9</sup>

カナダのオンタリオ州において、母親が妊娠中における H1N1 インフルエンザワ

クチンの接種と出生した児の健康状態との関連を検証した研究である。この研究は記述のカナダ・オンタリオ州で行われた投薬量と脳症リスクに関する研究<sup>5</sup>と同じく ICES によって提供されたリンケージデータを用いており、元のデータは、オンタリオ州の周産期記録、入院記録、救急外来受診記録、がん登録出生調査情報である。結果として、妊娠中のインフルエンザワクチン接種と児の健康アウトカムとの関連はないことが明らかとなった。

## D. 考察

2019年の医学主要4誌において、リンケージデータを用いた研究は発行論文数の中で0.5-3%と決してメジャーな研究方法ではなかった。ただし、そもそも医学主要4誌においては無作為化比較試験(RCT)といった介入試験による研究が多くを占め、観察研究はメジャーではない。観察研究のエビデンスレベルは、RCT(あるいは介入研究)より下位に位置し、その弱点として選択・情報バイアスが生じやすいことや交絡因子を十分に制御できないことが指摘されている。今回選択された9件の研究が、そうしたRCTが大勢を占める主要紙において掲載されている観察研究であることを踏まえると、9件の研究では、そうした観察研究における弱点をできるだけ軽減した研究であるとも考えられ、以下、そうした観察研究の弱点に沿ってリンケージデータを用いた各研究の特徴を述べていく。

まず、対象の選択バイアスについて、中国の例を除く各研究では、ほぼ全国(カナダでは一つの州全域)のデータを扱っており、非常に悉皆性の高い研究対象になっていると言える。そのため選択バイアスは「ほぼ無い」と言える。ただし、この利点はリンケージデータによるものと言うよりは、悉皆性のある既存データの二次利用に

よるものであると言える。同様に情報バイアスについても、各研究が用いたデータは、既存の行政資料等を基にした二次データであり、研究に関連した情報バイアスは生じにくい。ただし、これも観察研究の大きな弱点の一つであるが、研究目的のデータではあにため得られる情報には限りがあり、十分な交絡因子の制御が行われない結果につながる。しかし、データリンケージを行うことで、扱うことのできる情報の量を各段に増やすことができ、できるだけそうした未制御の交絡因子を減らすことに寄与できていると考えられる。

一方で、介入研究には倫理上、実施困難な研究仮説もあり、今回の 9 件はすべてそうした研究仮説であると言える。例えば、妊娠中のインフルエンザワクチン接種の児への影響を検証した研究<sup>9</sup>は、妊娠時にインフルエンザワクチンの接種の有無を研究目的にコントロールできるとは考えにくい。そして、この研究仮説を実証可能にしているのは、インフルエンザ接種歴を含む母親のデータと出生した児の健康状態のデータとのリンケージであると言える。

また、収入と平均余命との関連<sup>3</sup>や貧困状態と医療サービス支出との関連<sup>4</sup>といった、社会経済的地位（Socio-economic Status, SES）を要因として扱う研究は、通常、一つのデータセットになりにくい SES と健康状態や生存の 2 つの要素を、データリンケージによって一つのデータセットにして実施されている。つまり、曝露をデータ A から、アウトカムをデータ B から特定し、そしてデータ A とデータ B とをリンケージすることで、長期縦断的なデータセットを基に因果関係の検証を行うことができている。異なるデータそれぞれの強みを活かしたリンケージデータを作成して、これまで無かったデータセットを生み出せることは、創造的な研究仮説の検証につながっているものと考えられる。

最後に、リンケージ方法については、フィンランド、デンマーク、オランダ、ノルウェーでは国民に付与されている個人番号を使用した **Deterministic linkage** が行われており、従前から示されている北欧諸国の個人番号の強みが改めて示された形である。一方、アメリカ、カナダでは個人識別情報（氏名、性別、生年月日など）の組み合わせを用いた **Probabilistic linkage** を行っていた。こうした違いの背景には、個人情報保護に関する法的問題がある。一方で、個人番号による個人特定が有効であるかというように、個人番号が適切に全国民に付与されているかという個人番号の整備状況にも大きく関係しているとみられた。

#### ・本研究の限界

本研究では、医学主要 4 誌の 2019 年発行分のみを検索対象としており、その結果は極めて限局的である。しかし、一定のレベルを誇る医学主要誌に掲載されたリンケージデータ研究は、質の高いものであると言え、そうした研究のレビューはリンケージデータの活用上の意義について、一定の見解を導き出せるものと考えられる。

#### E. 結論

2019 年の医学主要 4 誌において、リンケージデータを用いた研究はメジャーな研究方法ではなかった。しかし各研究においてデータリンケージは、観察研究の弱みを軽減し、また実証可能な研究仮説の幅を拡張することに寄与しているとみられた。

#### F. 研究発表

1. 論文発表 なし
2. 学会発表 なし

#### G. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得 なし

2. 実用新案登録 なし
3. その他 なし

参考文献

1. Trine Madsen, et al. Association Between Traumatic Brain Injury and Risk of Suicide. *JAMA*. 2018 Aug 14;320(6):580-588
2. Merel Sanne Ekker, et al. Association of Stroke Among Adults Aged 18 to 49 Years With Long-term Mortality. *JAMA*. 2019 Jun 4;321(21):2113-2123
3. Jonas Minet Kinge, et al. Association of Household Income With Life Expectancy and Cause-Specific Mortality in Norway, 2005-2015. *JAMA*. 2019 May 21;321(19):1916-1925
4. Craig Evan Pollack et al. Association of Receipt of a Housing Voucher With Subsequent Hospital Utilization and Spending. *JAMA*. 2019 Dec 3;322(21):2115-2124
5. Flory T Muanda, et al. Association of Baclofen With Encephalopathy in Patients With Chronic Kidney Disease. *JAMA*. 2019 Nov 9;322(20):1987-1995
6. Iona Y Millwood, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet*. 2019 May 4;393(10183):1831-1842
7. Alice Goisis, et al. Medically assisted reproduction and birth outcomes: a within-family analysis using Finnish population registers. *Lancet*. 2019 Mar 23;393(10177):1225-1232
8. David M Studdert, et al. Changes in Practice among Physicians with Malpractice Claims. *N Engl J Med*. 2019 Mar 28;380(13):1247-1255
9. Laura K Walsh, et al. Health outcomes of young children born to mothers who received 2009 pandemic H1N1 influenza vaccination during pregnancy: retrospective cohort study. *BMJ*. 2019 Jul 10;366:l4151

表 レビュー結果の概要

雑誌		研究内容・結果	国	使用データ	リンケージ方法
JAMA	1	頭部外傷による自殺のリスクがある。	デンマーク	住民登録、労働記録、入院記録、患者記録、死亡票	個人番号
	2	若年性脳卒中の長期予後(死亡)は悪い。	オランダ	住民登録、入院記録、死亡票	個人番号
	3	収入による平均余命の差は2005年から2015年にかけて開いた。	ノルウェー	住民登録、納税記録、死亡票、教育記録	個人番号
	4	居住区域の貧困率によって医療サービス利用は変わらなかった。	アメリカ	医療保険請求記録	個人識別情報の組み合わせ
	5	バクロフェンの使用によって、高齢腎不全患者の脳症リスクが高かった。	カナダ オンタリオ州	退院記録、医師登録、救急外来受診記録、調剤記録など8種	個人識別情報の組み合わせ
LANCET	6	遺伝的に予測されたアルコール摂取は血管障害にほとんど関連していなかった。	中国	バイオバンク、死亡票、医療記録	個人番号
	7	不妊治療による児への影響は小さい。	フィンランド	調剤記録、出生記録	個人番号
NEJM	8	医療過誤の経験のある医師では診療業務のサイズダウンがみられた。	アメリカ	メディケアのプロバイダ調査、医師登録	個人識別情報の組み合わせ
BMJ	9	妊娠中のインフルエンザワクチン接種による児への影響はない。	カナダ オンタリオ州	周産期記録、入院記録、救急外来受診記録、がん登録	個人識別情報の組み合わせ