

厚生労働行政推進調査事業費補助金
政策科学総合研究事業（政策科学推進研究事業）

診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる
人材育成促進に資するための研究

令和4年度 総括・分担研究報告書

研究代表者 康永 秀生

令和4（2022）年 5月

目 次

I. 総括研究報告	
診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究-----	1
康永秀生、中山 健夫、田宮菜奈子、笹渕裕介	
II. 分担研究報告	
1. ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践-----	15
康永秀生	
2. NDBデータ研究の実践と人材育成プログラム開発・検証-----	232
中山健夫	
3. 医療・介護データ活用による研究・人材育成-----	236
田宮菜奈子	
4. 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究-----	241
笹渕裕介	
III. 研究成果の刊行に関する一覧表	246

令和4年度厚生労働行政推進調査事業費補助金
政策科学総合研究事業（政策科学推進研究事業）
総括研究報告書

診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究

研究代表者	康永秀生	東京大学大学院医学系研究科臨床疫学・経済学	教授
研究分担者	中山健夫	京都大学医学研究科健康情報学分野	教授
研究分担者	田宮菜奈子	筑波大学医学医療系ヘルスサービスリサーチ	教授
研究分担者	笹渕裕介	自治医科大学データサイエンスセンター	講師

NDB・DPC等の医療ビッグデータ解析に精通した研究者を育成することは急務である。令和4年度において、研究代表者・康永秀生の分担研究「ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践」では、令和3年度に開発したコンテンツ（NDBのeラーニング型教育コンテンツやオンデマンド型教育用ビデオコンテンツから構成される20コンテンツ）を用いた人材育成プログラムを実践した。また、大規模データ（NDB, DPC, JMDCデータ等）を用いた研究を継続的に実践した。

研究分担者・中山健夫らの分担研究「NDBデータ研究の実践と人材育成プログラム開発・検証」では、令和3年度に引き続き、アカデミアにおける研究者養成とは異なる形式での保健医療介護ビッグデータ研究人材育成について検討した。

研究分担者・田宮菜奈子らの分担研究「医療・介護データ活用による研究・人材育成」では、令和3年度に引き続き、医療・介護保険レセプトデータやDiagnosis Procedure Combination (DPC)データなどの医療ビッグデータを用いた研究を推し進め、特に博士課程・研修医・若手教員を対象にした人材育成に一層力を入れた。

研究分担者・笹渕裕介らの分担研究「診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究」では、令和3年度研究結果を元に、令和4年度は研究計画立案、統計解析用ソフトウェアRについて双方向型オンラインプログラムの開発を行い、研究者へ提供した。

今回我々が実践した研究は、NDBばかりでなくあらゆる保健・医療・介護ビッグデータに対応できる人材育成に関連するものである。開発されたプログラムやそれに基づく研究を継続的に実践することにより、ビッグデータのデータハンドリング、データベースマネージメントに関する総合的な技術を持つ研究者を多数輩出できる。また、日常臨床のクリニカル・クエスチョンを既存のビッグデータを用いて解明する研究実践能力を持つ研究者を多数輩出できる。さらに、データハンドリング技術と臨床研究実践能力の両方に長けた人材を多数育成でき、それによりわが国の医療ビッグデータ研究の進歩を加速できる。

A. 研究目的

近年、NDB・DPCなどの保健医療ビッグデータや介護ビッグデータを用いた大規模データベース研究が拡大している。これらの研究をさらに発展させるためには、臨床医学・疫学・統計学の知識に加えて、データベースに関連する知識や技能に精通した人材の育成が必須である。

研究代表者の康永秀生、研究分担者の中山健夫、田宮菜奈子、笹渕裕介は、長年にわたるビッグデータ研究の実績があり、各機関にお

いてビッグデータ研究の人材育成も担ってきた。康永は東京大学の、中山は京都大学のNDBオンサイトセンターの運営にもあたっている。NDBデータを研究用のデータセットに転換する際に存在する数多くの落とし穴の洗い出しやNDBデータの処理方法などは、オンサイトセンターの研究プロジェクトで既に明らかにしてきた。本研究では、康永・笹渕はNDB・DPC等、中山はNDBデータ等、田宮は医療・介護データ等の人材育

成プログラム開発・実践、およびビッグデータ研究の実施・指導を担当した。

先行の研究班のもとに、ビッグデータ研究利活用のための若手研究者コンソーシアムを立ち上げた。疫学・統計学・医療情報学等の社会医学系研究者に加えて、臨床医学、健康科学・看護学などの若手研究者も多数参画した。研究班メンバーが若手研究者を直接指導し、そこで日夜培われ開発される個別技術（データハンドリング技術、観察研究における統計解析技術など）を体系化・一般化し、既存の知識と合わせた、種々のビッグデータに応用可能な教育プログラムを開発してきており、これをさらに継続した。

研究代表者・康永秀生の研究「ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践」では、令和3年度に開発したコンテンツ（NDBのeラーニング型教育コンテンツやオンデマンド型教育用ビデオコンテンツから構成される20コンテンツ）を用いた人材育成プログラムを実践した。また、大規模データ(NDB,DPC,JMDCデータ等)を用いた研究を継続的に実践した。研究分担者・中山健夫らの分担研究「NDBデータ研究の実践と人材育成プログラム開発・検証」では、令和3年度に引き続き、アカデミアにおける研究者養成とは異なる形式での保健医療介護ビッグデータ研究人材育成について検討した。

研究分担者・田宮菜奈子らの分担研究「医療・介護データ活用による研究・人材育成」では、令和3年度に引き続き、医療・介護保険レセプトデータやDiagnosis Procedure Combination (DPC)データなどの医療ビッグデータを用いた研究を推し進め、特に博士課程・研修医・若手教員を対象にした人材育成に一層力を入れた。

研究分担者・笹渕裕介らの分担研究「診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究」では、令和3年度研究結果を元に、令和4年度は研究計画立案、統計解析用ソフトウェアRについて双方向型オンラインプログラムの開発を行い、研究者へ提供した。

B. 研究方法

1. ビッグデータ研究実践能力およびデータ

ハンドリング技術養成プログラムの実践

研究代表者・康永秀生らの分担研究では、令和4年度に以下を実施した。

(1) 令和3年度に開発したプログラムを用いた人材育成の実践

先行研究において、ビッグデータ研究実践能力養成プログラムおよびビッグデータハンドリング技術養成プログラムを開発した。前者は、日常臨床からクリニカル・クエスチョンを紡ぎ出し、検証可能なりサーチ・クエスチョンに構造化し、既存のビッグデータを用いてリサーチ・クエスチョンを解き明かす臨床研究・疫学研究・ヘルスサービスリサーチの実践能力を養成するプログラムである。後者は、テラバイト級のデータベースから個別の研究目的に沿うデータセットを抽出するデータハンドリング技術、膨大なテキストデータを含む大規模かつ複雑な構造のデータを研究用の扱いやすいデータベースに再構築するなどのデータベースマネジメント技術を養成するプログラムである。上記について、現下のコロナ禍の状況においても教育実践が可能となるように、令和3年度にオンラインによるプログラムに改良し、令和4年度にプログラムを実践した。

具体的には、「NDB・DPCデータベース研究人材育成 Webinar」を2023年1月10日(火)～1月20日(金)に開催した。Webによる動画配信（講義とeLearning）の形式とした。申請者に事前に登録してもらったアドレスに動画のURLを送信し、オンデマンドで視聴してもらった。配信期間中は全ての講義をいつでも視聴することを可能とした。

20講義のタイトルは下記のとおりである。

1. NDB・DPC等の公的データベースの動向
2. NDBガイドラインの理解度テスト
3. SQL入門
4. SQLを用いたレセプトデータのハンドリング
5. NDBを用いた研究
6. DPCデータを用いた研究1
7. DPCデータを用いた研究2
8. DeSCデータを用いた研究
9. JMDCデータを用いた研究
10. 傾向スコア
11. 高次元傾向スコア

12. 操作変数法
13. 不連続回帰デザイン・差の差分析
14. 時間依存性交絡と周辺構造モデル
15. 生存時間分析における競合リスクモデル
16. 欠測データの取り扱いと多重代入法
17. マルチレベル分析
18. 自己対照ケースシリーズ
19. 機械学習
20. データベースにおけるバリデーション研究

(2) 大規模データベースを用いた研究の実践

これまでの人材育成プログラムの内容に沿って、日常臨床からクリニカル・クエスチョンを紡ぎ出し、検証可能なリサーチ・クエスチョンに構造化し、NDB, DPC, JMDC データ等を用いた臨床研究・疫学研究・ヘルスサービスリサーチを、令和3年度に引き続き多数実践した。

◆倫理面への配慮について

NDB、DPC、介護データなどあらゆるデータは、すべての個人情報削除されている。NDB データは東大オンサイトセンターでの利用に限られた。その他のデータは東京大学のサーバー室内で厳重に管理される。データベースから研究プロジェクトごとに切り出されたデータセットは内部での利用に限られ、データセットのコピーの持ち出しは禁止とされている。本研究は東京大学の倫理委員会に倫理申請を行い承認を得た。

2. NDBデータ研究の実践と人材育成プログラム開発・検証

研究分担者・中山健夫らの分担研究では、令和4年度に以下を実施した。

京都大学は2019年度より文部科学省の助成を受け、関西広域の計13大学とコンソーシアムを構築して「関西広域・医療データ人材育成拠点形成事業 (Kansai Union / Kyoto University Education Program for Digital Health Innovation: KUEP-DHI. 責任者：黒田知宏・京都大学医学部附属病院医療情報企画部教授)」に取り組んでいる。本事業は医療現場から大

規模に収集される多様なデータの利活用を推進し、質の高い医療を実現するため、医療データの活用基盤を構築・運営する人材、医療データを利活用できる人材を育成することを目指すものである。KUEP-DHIの根幹は、大学院生（主に修士学生）を対象とした「医療データ取扱専門家育成コース」であり、京都大学大学院医学研究科と同大学院情報学研究科に追加履修コースとして設置されている。次世代医療基盤法や保健医療データプラットフォームに蓄積されるデータを活用し、医療データサイエンスの国際的牽引役を担えるよう本邦の医療データサイエンスを発展させるために、医療データが生まれてから活用されるまでの情報流の始点から終点までを確実に支え、正しく統制できる人材を育成することを目指している。

本研究は同事業の一環として実施されている2つの社会人向けインテンシブコースの運営経験に基づく事例検討である。

3. 医療・介護データ活用による研究・人材育成

研究分担者・田宮菜奈子らの分担研究では、令和4年度は若手研究者による医療・介護ビッグデータ活用による論文を2本、および学会発表を1回行った。解析及び論文執筆は、研究分担者（田宮）および研究協力者（岩上）の指導のもと、筑波大学医学医療系ヘルスサービスリサーチ分野に所属する博士課程学生および研修医が行った。

4. 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究

研究分担者・笹渕裕介らの分担研究では、①これまで自治医科大学データサイエンスセンターにおいて医療ビッグデータを利用した研究を行っている研究者、大学院生に対して必要なプログラムについて意見交換を行った。

②研究者・大学院生の意見を元にこれまでに作成した医療ビッグデータハンドリング教育プログラムから双方向型オンラインプログラムを開発するコンテンツの選別を行った。

- ③選別したコンテンツに関する双方向型のオンラインプログラムを開発した。
- ④実際にプログラムを提供し医療ビッグデータ研究実施へとつなげた。

C. 研究結果

1. ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践

(1) 令和3年度に開発したプログラムを用いた人材育成の実践

期間中、参加登録者は1,158名であった。内訳は、医療従事者717人（医師411人、薬剤師87人、看護師60人、歯科医師10人、その他コメディカル149人）、製薬会社187人、医療機器メーカー10人、保険会社6人、コンサルティング会社13人、医療IT企業26人、生物統計家6人、データサイエンティスト48人、診療情報管理士3人、審査支払機関5人、CRO18人、学生54人、教員18人、研究員21人、その他26人であった。アンケート回答はのべ7411人から得た。受講した一人当たり平均のアンケート回答回数は6.4回であった。

Q1.本講義の難易度について、「とてもわかりやすい」と「わかりやすい」を合計した割合は、講義番号3,5,6,7,8,9,10,16,17,18,20は90%を超え、課題番号1,4,11,12,13,14,15,19は80%を超えていたのに対し、講義番号2は49%と低かった。

Q2.本講義の進行速度について、いずれの課題番号も「ちょうどよい」が80%を超えていた。

Q3.本講義の満足度について、「とても満足」と「満足」を合計した割合は、講義番号1,3,4,5,6,7,8,9,10,11,12,13,14,16,17,18,20は90%を超え、講義番号15,19は80%を超えていたのに対し、講義番号2は59%と低かった。

なお、参加者の属性に関わらず、これらの傾向は一定であった。

(2) 大規模データベースを用いた研究の実践 大規模データベース(NDB,DPC,JMDC データ等)を用いたを用いた研究を行い、2022年

中に116編の英文原著研究論文を出版した。その成果は「G. 研究発表」の一覧に示す。

2. NDBデータ研究の実践と人材育成プログラム開発・検証

KUEP-DHI では大学院生という立場に依らず、企業の勤務者等、社会人を想定して次の2種類のインテンシブコースを提供している。

(1)DHIEP Program(ディープ プログラム、責任者: 奥野恭史・京都大学大学院医学研究科ビッグデータ医科学教授) Data-driven Healthcare Innovation Evangelist Promotion Program 分析だけを請け負う従来型データサイエンティストではなく、ライフサイエンス・ヘルスケア領域に深い造詣を有し、社会変革を実現しうる人材育成を目指すプログラム。対象は医療データの利活用を志す個人であり、個人参加、登録等も個人で行う。出願資格は大卒資格を有する者で、ライフサイエンス・IT 関連企業に所属する者及び今後当該領域に参入を検討している者、医療従事者、地方自治体職員、規制当局関係者、大学院生など定員は20名。カリキュラムの修業期間は半年間であり、大きく経営・社会変革、医学・生命科学、情報・データサイエンスの3領域で基礎知識をe-learning で学び、必修科目45時間(プレゼンテーション・コミュニケーション、環境・市場理解、ビジネスモデリング、デザインシンキング、チェンジマネジメント、医療情報倫理学、バイオインフォマティクス、可視化、AI・機械学習など)、選択科目21時間(ゲノム医療、IT 創薬、ライフサイエンスデータベース、時系列分析、テキストマイニング、ベイジアンネットワーク等)、実習(必修)20時間(医療データから患者を層別化し、診断アプローチを提案、既存データベースを活用して分子標的探索を行い現状プロセスの改善点を提案、電子カルテの副作用情報を用いたドラッグリポジショニングの探索、ライフログと医療データを用いた生活の中からの医療シーズの発見等)で構成されている。修了要件は、合計60時間以上の履修と各科目評価での合格(修了要件を満たせば、履修証明書が授与される)。受講料は社会人が50万円(税抜)、大学官公庁等の職員と大学院生は無料である。

(2)医療データ人材育成拠点形成事業ビジネス特化型インテンシブコース Kansai Union /

Kyoto University Education Program for Digital Health Innovation directly on themes of business (KUEP-DHI dot.b. 運営委員長: 中山健夫・京都大学大学院医学研究科社会健康医学系専攻健康情報学教授): 民間(企業)のニーズを踏まえつつ、医療データの活用に造詣のある人材を育成していくことを目的とするプログラム。対象は医療データ利活用に造詣のあるスタッフを確保したい企業であり、個人参加ではあるが、契約は企業—大学間で締結する。

育成する人材像は参加者が所定のプログラムを終えられた後には、各企業に戻り、医療データを企業内で有効かつ安全に利用する方針を立案したり、使用できるデータを適切に活用したりすることで有用な商品やビジネス等の開発に繋げ、社会に貢献できる人材を想定。到達目標は次の通りである。

- ・医療データを活用するための知識を系統立てて理解し、応用できる。
- ・医療データの活用に必要な現場感覚をもとに、医療データを分析して結果を解釈できる。

プログラムの内容は、dot.b 独自の講義、大学院の既存科目については科目等履修制度の活用、文献抄読、分析実習、臨床講義・実習、グループ演習で構成されている。分析実習では NDB オープンデータ等を用いた R、SQL 等の利用、グループ演習では、NDB 申出の練習や実際の NDB サンプリングデータセットを用いた分析を実施している。また臨床講義・実習では、電子カルテ入力体験により、電子カルテの自習環境の操作、電子カルテのインターフェイスの理解、予め提示された模擬患者に対し、医師になったつもりでのカルテ記入等を通して、入力されるデータの背景とリアリティを実感できる機会を提供している。グループ演習では、EBM の PECO (patient, exposure, comparison, outcome) 形式を学び、構造化された研究計画を立ててデータを抽出し分析するという、データベース研究の一連のプロセスを提供している。dot.b プログラムは、参加者個人と契約を行うのではなく、まず本プログラムの趣旨に賛同する企業を募り、大学と有償の共同研究契約を締結する。そして各企業内で参加の該当者を選定し、プログラムに派遣する。参加者は京都大学と各企業の共同研究のプロジェクト研究員として、プログラム参加期間中は、京都大学での立場を持つことができ、京都大学の各種リ

ソースを一定の条件で利用できる。

コースは 2020 年度から開始され、2020 年度には 6 社・9 名、2021 年度は 3 社・4 名、2022 年度は 7 社・8 名が参加した。また、各年度の参加者が一同に会して NDB を分析する演習を行い、コース修了者との関係を維持し、参加者間の繋がりを深化させる機会を提供している。

3. 医療・介護データ活用による研究・人材育成

1. 論文タイトル: Evaluation of enhanced home care support clinics regarding emergency home visits, hospitalization, and end-of-life care: a retrospective cohort study in a city of Japan (BMC Health Serv Res. 2023 Feb 3;23(1):115.)

【発表者】孫瑜 (博士課程 4 年)

2. 論文タイトル: Pulse Methylprednisolone versus Dexamethasone in COVID-19: A Multicenter Cohort Study (Crit Care Explor. 2023;5(4):e0886.)

【発表者】渡邊淳之 (筑波大学病院初期研修医)

3. 講演タイトル: 高齢 COVID-19 入院患者における入院早期リハビリテーションと在院死亡との関連 (2022 年度第 5 回日本臨床疫学会 YIA 候補者発表会)

【発表者】宇田和晃 (助教)

以上のように、医療・介護ビッグデータを有効に活用し新たな医学的知見を得ることができた。また、研究計画立案から、データクリーニング、統計解析、論文執筆、論文投稿、レビューワーとのやり取り、までの一連の流れについて、次世代を担う若手に経験してもらおうと共に、明確な成果物として世に発信することができた。今回の論文執筆をきっかけに、今後の活躍が期待できる。

4. 診療現場の実態に即した医療ビッグデータ (NDB 等) を利活用できる人材育成促進に資するための研究

①②特に統計解析のプログラムは双方向型であることが習得の近道であるとの意見が多かった。また、ほぼすべての研究者、大学院生から研究計画立案について知識が足りないことについて教育を受ける機会が不十分であるとの意見があった。そこで研究者にとって必要性の高い研究計画および統計ソフトウェア R についての双方向型オンライ

ンプログラムの開発を行うこととした。
双方向型教育プログラムのコンテンツ

(1) 研究計画立案

研究計画の立案について教育を受ける機会が少ない、書籍等での学習と自身での計画立案にはギャップがあり、自力でそのギャップを埋めることが難しい等の意見を得た。そこで、研究計画立案に際して必須の以下の項目について双方向型オンラインプログラムの開発を行った。具体的には講義に加えて、実践的な作業を行うことができるような研究計画立案に関する課題を設定し、それを解くことで能動的に講義の内容を理解できるようなコンテンツとした。テーマは以下のとおりである。

- ・クリニカルクエスチョン (CQ) からリサーチクエスチョン (RQ) への作り変え
- ・FINER による RQ の評価
- ・文献検索
- ・研究デザイン
- ・抽出変数の定義

(2) R による統計解析

R に関して令和 2 年度までに開発した e-learning プログラムを改良し、以下の項目について双方向型オンライン教育プログラムの開発を行った。具体的には講義中に実際に自分でスクリプトを書くことで R の基本的な使用方法を理解することに加えて、講義の内容を反映した課題を設定し、それを解くことで能動的にスクリプトを書くことができるようなコンテンツとした。テーマは以下のとおりである。

- ・データの整形
- ・各変数の集計及び可視化
- ・検定
- ・回帰分析

③研究計画立案、R による統計解析に関して動画撮影が終了し、オンラインプログラムとして使用可能なプログラムの開発が完了した。

④プログラムを自治医科大学データサイエンスセンターにおいて大学院生、客員研究員、共同研究者に提供した。提供を受けた複数の研究者により研究計画の立案から論文発表まで行われた。

D. 考察

本研究は、わが国のビッグデータ研究において不足している「人材育成」を最重視し、ビッグデータ研究のための種々の技術を一般化し体系的なプログラムを構築した上で、それらを多数の研究者等に利活用してもらえようように社会実装を試みる点が、既存研究にない独創的な点である。

1. ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践

本プログラムの内容の多くは実質的に既に研究代表者を中心とする研究チームで実践されてきたものを踏襲しており、それらを用いて多くの論文投稿・出版の実績を挙げてきた。その実績を以て、すでにその効果は実証済みである。本研究は、それを体系化・一般化する試みである。

多数の大規模データ研究・論文執筆を通して用いられてきた個別技術(データハンドリング技術、観察研究における統計解析技術、など)を体系化・一般化し、既存の知識と合わせて、種々のビッグデータに応用可能な人材育成プログラム version 1 を 2017 年度に最初に開発した。これを 2018 年度に実施し、その効果を検証した。2019 年度研究においては、2018 年度の実践結果を踏まえて、プログラムを改良し、人材育成プログラム ver.2 を作成し、同年度に実施した対面型セミナーの評価結果は概ね良好であった。さらにプログラムの利用者拡大のため、2020 年度にはオンライン教育プログラムの作成を行った。評価結果は極めて良好であった。

さらに 2021 年度は上記の開発を行い、2022 年度にリリースした。

上記のすべては、厚生労働省が進める NDB 高度利活用 に直接反映される研究である。NDB ばかりでなく、あらゆる保健・医療・介護ビッグデータに対応できる人材育成を図ることにより、大規模データのデータハンドリング、データベースマネージメント等に関する総合的な技術を持つ人材を多数育成できる。

データハンドリング技術と臨床研究実践能力の両方に長けた人材を多数育成することにより、わが国の大規模医療データベース研究の進歩を加速でき、それによってわが国発

のエビデンスを量産できることが期待される。

2. NDBデータ研究の実践と人材育成プログラム開発・検証

研究分担者・中山健夫らの分担研究では、「関西広域・医療データ人材育成拠点形成事業」における2種類の社会人向けインテンシブコースの経験から、データサイエンスの人材育成に関するいくつかの知見を得た。下記に抜粋して記述する。

- ・医療者でない参加者の場合、医療データ以前の「医療そのもの」への距離感が大きいいため、プログラム提供に当たっては、そのギャップを十分意識することが必要。
- ・NDB等の実際のデータに触れることは学習者のモチベーションを大きく高め得る。
- ・「公益性のある研究」としてNDB 申出書を書くことの意義を体感することは大いに重要。
- ・医師の診断プロセスや、電子カルテ入力体験を通じた、「データが絶対とは限らない」「バリデーションが重要」という感覚の共有はきわめて重要。
- ・レセプトでの傷病名同定における投薬や処置の情報を理解することの重要性。
- ・医療の場で入力されている情報が必ずしも正確かつ網羅的とは言えないことの実感。
- ・模擬的なデータ分析体験を通じた、臨床研究への意識の高まり。
- ・「企業が求める人材育成」と「大学教育が想定する人材育成」の違いの認識。
- ・受講者ニーズと提供するコンテンツとのマッチングの重要性。

本プログラムは2022年度も実施しており、引き続き、より良いデータサイエンス人材育成の在り方、方向性を探っていきたい。

3. 医療・介護データ活用による研究・人材育成

医療・介護ビッグデータを有効に活用し新たな医学的知見を得ることができた。また、研究計画立案から、データクリーニング、統計

解析、論文執筆、論文投稿、レビューワーとのやり取り、までの一連の流れについて、次世代を担う若手に経験してもらおうと共に、明確な成果物として世に発信することができた。今回の論文執筆をきっかけに、今後の活躍が期待できる。

4. 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究

医療ビッグデータを用いた研究のための双方向型教育プログラムの作成にあたり、研究者・大学院生と意見交換を行い、研究者はデータハンドリングや統計解析に加えて研究デザインの教育も必要であることが明らかとなった。

当センターで行う研究についての研究計画の立案は分担者が個別に各研究者や大学院生に対して直接指導を行っている。令和2年度までにこれらの指導内容を体系化し、直接指導を行わなくとも保健医療ビッグデータ研究を行うためのデータハンドリングから解析までを習得可能な教育プログラムを開発し、研究者や大学院生に提供してきた。これらの教育プログラムによってビッグデータを用いた研究の実践を通して、学会発表・論文執筆につながった。しかしながらこれらの教育プログラムは一方向性であり、一部十分な習得に至らなかった部分に関して改良の余地が残されていた。本研究ではこれらの習得が不十分な点について、プログラム受講者から意見を集めることで特に研究者にとってデータハンドリングや統計解析に加えて研究計画の立案についても教育プログラムが必要であることが明らかとなった。ずさんな研究計画のまま研究を開始すると、研究途中で様々な問題が生じ、研究自体が頓挫してしまうこともしばしば起こる。そのため、綿密な研究計画を立案する必要がある。本研究で開発した研究計画立案のための双方向型オンライン教育プログラムはこれまで分担者が行ってきた指導をオンライン用に構築したものである。また、令和2年度までに開発したR習得e-learningプログラムを、動画視聴の合間に複数の課題を設定するこ

とにより能動的に R を習得することのできる双方向型のオンライン教育プログラムとして改良した。プログラム受講者からの評価は良好であったが、更に発展的な内容のプログラムを望む声も寄せられており、今後の課題としたい。

本研究事業により開発した教育プログラムは医療ビッグデータを利活用できる人材育成につながったと考えられる。実際にプログラム受講者が研究デザインやデータハンドリングを行うことができるようになっており、研究成果として論文発表につながっている (G. 研究発表参照) ことから有用なプログラムであったことが示された。今後このプログラムをより多くの研究者へ提供していく予定である。

4つの研究を総括すると、下記のような貢献が期待できる。

近年、保健・医療・介護データのインフラ整備が進められているが、そのインフラを活用できる人材の育成に貢献し、今後待たれる医療・介護統合データベースの稼働の際には全省的な政策課題に関する研究・知見を提供することにも貢献できる。さらに、日常臨床のクリニカル・クエスチョンを既存のビッグデータを用いて解明する研究実践能力を持つ研究者を多数育成できる。さらに、データハンドリング技術と臨床研究実践能力の両方に長けた人材を多数育成することにより、わが国の医療ビッグデータ研究の進歩を加速できる。わが国の医療ビッグデータ研究の技術水準を世界トップレベルに向上させ、それによってわが国発のエビデンスを量産できることが期待される。

E. 結論

今回我々が実践した研究は、NDBばかりでなくあらゆる保健・医療・介護ビッグデータに対応できる人材育成に関連するものである。開発されたプログラムやそれに基づく研究を継続的に実践することにより、ビッグデータのデータハンドリング、データベースマネージメントに関する総合的な技術を持つ研究者を多数輩出できる。また、日常臨床のクリニカル・クエスチョンを既存のビッグデー

タを用いて解明する研究実践能力を持つ研究者を多数輩出できる。さらに、データハンドリング技術と臨床研究実践能力の両方に長けた人材を多数育成でき、それによりわが国の医療ビッグデータ研究の進歩を加速できる。

わが国の医療ビッグデータ研究の技術水準を世界トップレベルに向上させ、それによってわが国発のエビデンスを量産できることが期待される。様々なビッグデータ解析の結果は厚生労働行政に資する基礎資料として間接的に活用されることが期待される。さらに、英文論文化されたエビデンスは全世界の臨床家・公衆衛生実務家・民間企業のプラクティスに間接的に活用されるだろう。

また本研究を通じて、医療・介護ビッグデータ研究の後進の育成実績およびそのノウハウが順調に蓄積されている。これは、今後の医療人材のあるべき方向にも沿ったものであり、かつ学生や若手自身が自ら意欲的に取り組める内容でもあり、適切な指導により高い成果につながれることが示唆された。今後、以上の経験を生かし、それぞれの分野やレベルの違いに応じた医療・介護ビッグデータ人材教育の最適な方法論についてまとめていくことが必要であると考えられる。

F. 健康危険情報 なし

G. 研究発表

<康永秀生>

論文発表

1. Abe H, Sumitani M, Matsui H, Inoue R, Fushimi K, Uchida K, Yasunaga H. Use of naldemedine is associated with reduced incidence of hyperactive delirium in cancer patients with opioid-induced constipation: a nationwide retrospective cohort study in Japan. *Pharmacotherapy*. 2022;42(3):241-249.
2. Abe H, Sumitani M, Matsui H, Inoue R, Konishi M, Fushimi K, Uchida K, Yasunaga H. Gabapentinoid use is associated with reduced occurrence of hyperactive delirium in older cancer patients undergoing chemotherapy: a nationwide retrospective cohort study in Japan. *Anesthesia & Analgesia*. 2022;135(2):362-369.
3. Aso S, Matsui H, Fushimi K, Yasunaga H. Effect of Dobutamine on Patients with Septic Shock: A Retrospective Nationwide Study. *Annals of Clinical Epidemiology*. 2022;4(2): 41-48.
4. Awano N, Jo T, Izumo T, Inomata M, Morita K, Matsui H, Fushimi K, Urushiyama H, Nagase T, Yasunaga H. Recombinant human soluble thrombomodulin for acute exacerbation of idiopathic pulmonary fibrosis: A nationwide observational study. *Journal of Intensive Care*.

- 2022;10(1):14.
5. Chinen T, Sasabuchi Y, Matsui H, Yamaguchi H, Yasunaga H. Oxaliplatin- versus cisplatin-based regimens for elderly individuals with advanced gastric cancer: A retrospective cohort study. *BMC Cancer*. 2022;22(1):460.
 6. Fujita A, Hashimoto Y, Matsui H, Yasunaga H, Aihara M. Recent trends in glaucoma surgery: a nationwide database study in Japan, 2011–2019. *Japanese Journal of Ophthalmology*. 2022;66(2):183-192.
 7. Fujita A, Hashimoto Y, Okada A, Obata R, Aihara M, Matsui H, Yasunaga H. Association Between Proteinuria and Retinal Vein Occlusion in Individuals with Preserved Renal Function: A Retrospective Cohort Study. *Acta Ophthalmologica*. 2022;100(7):e1510-e1517
 8. Fukui R, Kaneko H, Okada A, Yano Y, Itoh H, Matsuoka S, Morita K, Kiriyama H, Kamon T, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Nishiyama A, Node K, Yokoo T, Nangaku M, Yasunaga H, Komuro I. Semiquantitative assessed proteinuria and risk of heart failure: Analysis of a nationwide epidemiological database. *Nephrology Dialysis Transplantation*. 2022;37(9):1691-1699.
 9. Furuya H, Nakajima M, Ikeda K, Nakamura K, Ohbe H, Aso S, Kumazawa R, Iwamoto T, Iwata A, Furuta S, Matsui H, Fushimi K, Yasunaga H, Nakajima H. Prognosis and treatment of myositis-associated severe interstitial lung disease. *Arthritis Care & Research*. 2022;74(3):478-483.
 10. Fuseda Y, Michihata N, Kumazawa R, Fushimi K, Yasunaga H. Annual Trend in Lumbar Puncture for Infants Younger Than 3 Months Hospitalized with Suspected Serious Bacterial Infection: A Nationwide Inpatient Database Study. *Pediatric Infectious Disease Journal*. 2022;41(8):631-635.
 11. Hamada S, Nakajima M, Kaszynski RH, Otaka S, Goto H, Matui H, Fushimi K, Yamaguchi Y, Yasunaga H. Association between adjunct clindamycin and in-hospital mortality in patients with necrotizing soft tissue infection due to group A *Streptococcus*: A nationwide cohort study. *European Journal of Clinical Microbiology & Infectious Diseases*. 2022;41(2):263-270.
 12. Hashimoto Y, Kaneko H, Okada A, Matsui H, Yasunaga H, Aihara M, Obata R. Association between retinal vein occlusion and Life's Simple 7 cardiovascular health metrics: A large claims database study. *Ophthalmology Retina*. 2022;6(8):684-692.
 13. Hashimoto Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, Kaburaki T. Incidence of sympathetic ophthalmia after inciting events: a national database study in Japan. *Ophthalmology*. 2022;129(3):344-352.
 14. Hashimoto Y, Michihata N, Yamana H, Shigemi D, Morita K, Matsui H, Yasunaga H, Aihara M. Safety of topical ophthalmic antibiotics in pregnant women with hordeola, chalazia, blepharitis, or bacterial conjunctivitis: propensity score analyses. *EYE*. 2022 ;36(5):1066-1073.
 15. Hashimoto Y, Yamana H, Michihata N, Shigemi D, Ishimaru M, Matsui H, Yasunaga H, Aihara M. Eye drops for dry eye disease during pregnancy and adverse neonatal outcomes: high-dimensional propensity score analyses. *Ophthalmic Epidemiology*. 2022;29(4):384-393.
 16. Hirano T, Nakajima M, Ohbe H, Kaszynski RH, Iwasaki Y, Arakawa Y, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H. Corticosteroid use with extracorporeal cardiopulmonary resuscitation for out-of-hospital cardiac arrest: A nationwide observational study. *Resuscitation Plus*. 2022;12:100308.
 17. Hirano Y, Kaneko H, Konishi T, Itoh H, Matsuda S, Kawakubo H, Uda K, Matsui H, Fushimi K, Daiko H, Itano O, Yasunaga H, Kitagawa Y. Short-Term Outcomes of Epidural Analgesia in Minimally Invasive Esophagectomy for Esophageal Cancer: Nationwide Inpatient Data Study in Japan. *Annals of Surgical Oncology*. 2022;29(13):8225-8234.
 18. Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H. Association between nurse aide staffing and patient mortality after major cancer surgeries in acute care settings: A retrospective cohort study using a national inpatient database. *Nursing and Health Sciences*. 2022;24(1):283-292.
 19. Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H. Dose-response association between nurse staffing and patient outcomes following major cancer surgeries using a nationwide inpatient database in Japan. *Journal of Clinical Nursing*. 2022;31(17-18):2562-2573.
 20. Honda A, Michihata N, Iizuka Y, Uda K, Morita K, Mieda T, Takasawa E, Ishiwata S, Tajika T, Matsui H, Fushimi K, Yasunaga H, Chikuda H. Risk factors for severe lower extremity ischemia following venoarterial extracorporeal membrane oxygenation: An analysis using a nationwide inpatient database. *Trauma Surgery & Acute Care Open*. 2022;7:e000776.
 21. Hosoi T, Yamana H, Tamiya H, Matsui H, Fushimi K, Akishita M, Yasunaga H, Ogawa S. Association between comprehensive geriatric assessment and polypharmacy at discharge in patients with ischaemic stroke: a nationwide retrospective cohort study. *eClinicalMedicine*. 2022;50:101528.
 22. Inooka Y, Yamana H, Shinoda Y, Inokuchi H, Matsui H, Fushimi K, Yasunaga H, Haga N. Predictive factors for oral intake recovery after acute stroke: analysis of a Japanese nationwide inpatient database. *Dysphagia*. 2022;37(6):1623-1632.
 23. Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H. Prevalence, incidence rate and risk factors of medication-related osteonecrosis of the jaw (MRONJ) in patients with osteoporosis and cancer: a nationwide population-based study in Japan. *Journal of Oral and Maxillofacial Surgery*. 2022;80(4):714-727.
 24. Ishimaru T, Fujiogi M, Michihata N, Konishi T, Morita K, Matsui H, Uda K, Fushimi K, Kawashima H, Fujishiro J, Yasunaga H. Perioperative Outcomes of Laparoscopically Assisted Anorectoplasty versus Conventional Procedures for Anorectal Malformation: A Retrospective Nationwide Database Study. *Pediatric Surgery International*. 2022;38(12):1785-1791.
 25. Itoh H, Kaneko H, Okada A, Suzuki Y, Fujiu K, Matsuoka S, Michihata N, Jo T, Nakanishi K, Takeda N, Morita H, Node K, Di Tullio MR, Homma S, Yasunaga H, Komuro I. Age-Specific Relation of Cardiovascular Health Metrics With Incident Cardiovascular Disease. *American Journal of Cardiology*. 2022;177:34-39.
 26. Iwasaki Y, Ohbe H, Shigemi D, Fushimi K, Yasunaga H. Effect of antithrombin III among patients with disseminated intravascular coagulation in obstetrics: a nationwide observational study in Japan. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2022;129(5):805-811.
 27. Kamijo K, Nakajima M, Shigemi D, Kaszynski RH, Ohbe H, Goto T, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H. Resuscitative endovascular balloon occlusion of the aorta for life-threatening postpartum hemorrhage: A nationwide observational study in Japan. *The Journal of Trauma and Acute Care Surgery*. 2022;93(3):418-423.
 28. Kamon T, Kaneko H, Itoh H, Okada A, Matsuoka S, Kiriyama H, Fujiu K, Morita K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I. Sex difference in the association between lipid profile and incident cardiovascular disease among young adults. *Journal of Atherosclerosis and Thrombosis*. 2022;29(10):1475-1486.

29. Kaneko H, Suzuki Y, Ueno K, Okada A, Fujiu K, Matsuoka S, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Node K, Yasunaga H, Komuro I. Association of Life's Simple 7 with incident cardiovascular disease in 53,974 patients with cancer. *European Journal of Preventive Cardiology*. 2022;29(18):2324-2332.
30. Kaneko H, Yano Y, Lee HH, Lee H, Okada A, Itoh H, Morita K, Fukui A, Fujiu K, Suzuki Y, Matsuoka S, Nakamura S, Michihata N, Jo T, Takeda N, Morita H, Yokoo T, Nishiyama A, Node K, Viera AJ, Muntner P, Oparil S, Kim HC, Yasunaga H, Komuro I. Medication-Naïve Blood Pressure and Incident Cancers: Analysis of Two Nationwide Population-Based Databases. *American Journal of Hypertension*. 2022;35(8):731-739.
31. Kaneko H, Yano Y, Suzuki Y, Okada A, Itoh H, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Node K, Viera AJ, Lima JAC, Oparil S, Lam CSP, Carey RM, Yasunaga H, Komuro I. Reduction in blood pressure for elevated blood pressure/ stage I hypertension according to the ACC/AHA guideline and cardiovascular outcomes. *European Journal of Preventive Cardiology*. 2022;29(14):1921-1929.
32. Karakawa R, Konishi T, Yoshimatsu H, Fuse Y, Hashimoto Y, Matsui H, Fushimi K, Yano T, Yasunaga H. Comparison of short-term outcomes between pedicled- and free-flap autologous breast reconstruction. *Breast Cancer*. 2022;29(6):1067-1075.
33. Kodama S, Jo T, Yasunaga H, Michihata N, Matsui H, Kumazawa R, Shirota Y, Fushimi K, Toda T, Hamada M. Outcomes of gastrointestinal cancer surgeries in Parkinson's disease patients: a nationwide study. *Parkinsonism and Related Disorders*. 2022;96:45-49.
34. Konishi T, Fujiogi M, Michihata N, Kumazawa R, Ohbe H, Matsui H, Fushimi K, Ogita M, Tanabe M, Seto Y, Yasunaga H. Interstitial lung disorders following postoperative radiotherapy with concurrent or sequential hormonal therapy for breast cancer: a nationwide database study in Japan. *Breast Cancer*. 2022;29(4):688-697.
35. Konishi T, Fujiogi M, Michihata N, Matsui H, Tanabe M, Seto Y, Yasunaga H. Association between body mass index and incidence of breast cancer in premenopausal women: A Japanese nationwide database study. *Breast Cancer Research and Treatment*. 2022;194(2):315-325.
36. Konishi T, Fujiogi M, Michihata N, Morita K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Association of operative day of the week with length of stay and total hospitalization costs in patients with partial mastectomy: a nationwide database study in Japan. *JMA Journal*. 2022;5(3):319-327.
37. Konishi T, Fujiogi M, Michihata N, Ohbe H, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Cost-effectiveness analysis of trastuzumab monotherapy versus adjuvant chemotherapy plus trastuzumab in elderly patients with HER2-positive early breast cancer. *Japanese Journal of Clinical Oncology*. 2022;52(10):1115-1123.
38. Konishi T, Fujiogi M, Sato A, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Short-term outcomes following breast cancer surgery with and without neoadjuvant chemotherapy: a nationwide administrative database study in Japan. *Annals of Surgical Oncology*. 2022;29(5):3000-3010.
39. Konishi T, Fujiogi M, Shigemi D, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Risk factors for postoperative bleeding following breast cancer surgery: a nationwide database study of 477,108 cases in Japan. *World Journal of Surgery*. 2022;46(12):3062-3071.
40. Konishi T, Fujiogi M, Shigemi D, Nishioka K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Surgical and obstetric outcomes of breast cancer surgery during pregnancy: A nationwide database study in Japan. *Breast Cancer Research and Treatment*. 2022;195(3):289-299.
41. Konishi T, Goto T, Fujiogi M, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using a Japanese nationwide inpatient database. *Surgery*. 2022;171(4):1036-1042.
42. Konishi T, Sakata A, Inokuchi H, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Treatments and outcomes of adult parapharyngeal and retropharyngeal abscess: 1,882 cases from a Japanese nationwide database. *American Journal of Otolaryngology*. 2022;44(2):103770.
43. Konishi T, Takamoto T, Fujiogi M, Hashimoto Y, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Laparoscopic versus open distal pancreatectomy with or without splenectomy: A propensity score analysis in Japan. *International Journal of Surgery*. 2022;104:106765.
44. Konishi T, Yoshimoto T, Fujiogi M, Yamana H, Tanabe M, Seto Y, Yasunaga H. Validity of operative information in Japanese administrative data: a chart review-based analysis of 1221 cases at a single institution. *Surgery Today*. 2022;52(10):1484-1490.
45. Kumazawa R, Jo T, Matsui H, Fushimi K, Yasunaga H. Direct oral anticoagulants versus warfarin for secondary prevention of cerebral infarction and bleeding in older adults with atrial fibrillation. *Journal of the American Geriatrics Society*. 2022;70(7):2029-2039.
46. Kurakawa KI, Okada A, Bessho K, Jo T, Ono S, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yamaguchi S, Yamauchi T, Nangaku M, Kadowaki T, Yasunaga H. Major complications after percutaneous biopsy of native or transplanted liver in pediatric patients: A nationwide inpatient database study in Japan. *BMC Gastroenterology*. 2022;22(1):395.
47. Kutsuna S, Ohbe H, Matsui H, Yasunaga H. Effectiveness of fluoroquinolone antimicrobials in addition to tetracyclines for Japanese Spotted Fever: A retrospective analysis using a national inpatient database. *International Journal of Infectious Diseases*. 2022;123:70-75.
48. Kutsuna S, Ohbe H, Matsui H, Yasunaga H. Delayed Tetracycline Initiation Increases Mortality Risk in Patients With Japanese Spotted Fever: Retrospective Analysis Using a National Inpatient Database. *Open Forum Infectious Diseases*. 2022;9(11):ofac573.
49. Matsuda K, Jo T, Matsui H, Fushimi K, Yasunaga H, Sugimoto K. Institutional factors associated with early mortality of newly diagnosed acute promyelocytic leukemia. *Blood Cancer Journal*. 2022;12(12):167.
50. Matsuda K, Jo T, Shimura A, Honda A, Taoka K, Masamoto Y, Matsui H, Fushimi K, Yasunaga H, Kurokawa M. Risk of febrile neutropenia in very elderly patients aged ≥ 80 years receiving their first cycle of R-CHOP regimen: a nationwide real-world study in Japan. *British Journal of Haematology*. 2022;197(3):e37-e41.
51. Matsui H, Yamana H, Fushimi K, Yasunaga H. Development of deep learning models for predicting in-hospital mortality using an administrative claims database. *JMIR Medical Informatics*. 2022;10(2):e27936.
52. Matsuoka S, Kaneko H, Kamon T, Suzuki Y, Yano Y, Okada A, Itoh H, Morita K, Fukui A, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Yokoo T, Nishiyama A, Node K, Yasunaga H, Komuro I. Incorporation of Retinal Arteriosclerosis into Risk Stratification of Blood Pressure Category According to the 2017 ACC/AHA Blood Pressure Guideline. *Journal of Atherosclerosis and Thrombosis*. 2022;29(10):1487-1498.
53. Matsuoka S, Kaneko H, Okada A, Fukui A, Yano Y, Itoh H, Morita K, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Nakamura S, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I.

Association between Proteinuria and Incident Colorectal Cancer: Analysis of a Nationwide Population-Based Database. *BMJ Open*. 2022;12(4):e056250.

54. Matsuoka S, Kaneko H, Okada A, Itoh H, Suzuki Y, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I. Association of retinal atherosclerosis assessed using Keith-Wagener-Barker system with incident heart failure and other atherosclerotic cardiovascular disease: Analysis of 319,501 individuals from the general population. *Atherosclerosis*. 2022;348:68-74.
55. Matsuoka S, Kaneko H, Okada A, Morita K, Itoh H, Michihata N, Jo T, Takeda N, Morita H, Fujiu K, Nakamura S, Node K, Yasunaga H, Komuro I. Age Modified Relationship between Modifiable Risk Factors and the Risk of Atrial Fibrillation. *Circulation: Arrhythmia and Electrophysiology*. 2022;15(1):e010409.
56. Michihata N, Shigemi D, Yamana H, Matsui H, Jo T, Yasunaga H. Safety of Japanese Herbal Kampo Medicines for the Treatment of Depression During Pregnancy. *International Journal of Gynecology and Obstetrics*. 2022 159(3):865-869.
57. Michihata N, Suzuki T, Honda A, Oka I, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Clinical features of Kawasaki disease with atlantoaxial rotatory fixation. *Pediatric Infectious Disease Journal*. 2022;41(8):626-630.
58. Michihata N, Suzuki T, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Association between intravenous immunoglobulin dose and outcomes in patients with acute Kawasaki disease. *European Journal of Pediatrics*. 2022;181(10):3607-3615.
59. Miyachi H, Konishi T, Kumazawa R, Matsui H, Shimizu S, Fushimi K, Matsui H, Yasunaga H. Treatments and outcomes of generalized pustular psoriasis: a cohort of 1516 patients in a nationwide inpatient database in Japan. *Journal of the American Academy of Dermatology*. 2022;86(6):1266-1274.
60. Miyachi H, Konishi T, Shigemi D, Matsui H, Shimizu S, Fushimi K, Matsue H, Yasunaga H. Association between psoriasis and short-term outcomes of acute myocardial infarction: a matched-pair cohort study using a nationwide inpatient database in Japan. *JAAD International*. 2022;8:21-30.
61. Miyakawa T, Michihata N, Morita K, Matsui H, Honda M, Yasunaga H. Ambient temperature and hospital admissions for acute cholecystitis: A nationwide inpatient database study in Japan. *HPB*. 2022;24(3):398-403.
62. Miyamoto Y, Iwagami M, Aso S, Uda K, Fushimi K, Hamasaki Y, Nangaku M, Yasunaga H, Doi K. Postoperative outcomes of cancer surgery in patients with and without kidney failure with dialysis therapy: A matched-pair cohort study. *Clinical Kidney Journal* 2022;15(6):1137-1143.
63. Miyamoto Y, Matsuyama T, Goto T, Ohbe H, Kitamura T, Yasunaga H, Ohta B. Association between age and neurological outcomes in out-of-hospital cardiac arrest patients resuscitated with extracorporeal cardiopulmonary resuscitation: a nationwide multicentre observational study. *European Heart Journal: Acute Cardiovascular Care*. 2022;11(1):35-42.
64. Mizuno S, Ono S, Takano A, Yasunaga H, Iwase H. Dental characteristics associated with methamphetamine use: Analysis using forensic autopsy data. *BMC Oral Health*. 2022;22(1):141.
65. Nagata S, Ohbe H, Jo T, Matsui H, Fushimi K, Yasunaga H. Glucocorticoids and rates of biphasic reactions in patients with adrenaline-treated anaphylaxis: A propensity score matching analysis. *International Archives of Allergy and Immunology*. 2022;183(9):939-945.
66. Nishimoto Y, Ohbe H, Matsui H, Nakajima M, Sasabuchi Y, Goto T, Morita K, Fushimi K, Sato Y, Yasunaga H. Predictive ability of the sequential organ failure assessment score for in-hospital mortality in patients with cardiac critical illness: a nationwide observational study. *EHJ Acute Cardiovascular Care*. 2022;11(4):312-321.
67. Ogasawara H, Obe H, Yasunaga H. Oral food challenge management in Japan: A retrospective analysis of health insurance claims data. *Clinical & Experimental Allergy*. 2022;52(7):898-900.
68. Ogawa Y, Yamana H, Noda T, Kishimoto M, Yoshihara S, Kanaoka K, Matsui H, Fushimi K, Yasunaga H, Kawaguchi M, Imamura T. Severe complications after general anesthesia versus sedation during pediatric diagnostic cardiac catheterization for ventricular septal defect. *Journal of Clinical Medicine*. 2022;11(17):5165.
69. Ohbe H, Goto T, Matsui H, Fushimi K, Yasunaga H. Associations of government-issued ICU admission criteria with clinical practices, outcomes, and ICU bed occupancy. *Annals of the American Thoracic Society*. 2022;19(6):1013-1021.
70. Ohbe H, Goto T, Nakamura K, Matsui H, Yasunaga H. Development and validation of early prediction models for new-onset functional impairment at hospital discharge of ICU admission. *Intensive Care Medicine*. 2022;48(6):679-689.
71. Ohbe H, Matsui H, Kumazawa R, Yasunaga H. Intensive care unit versus high dependency care unit admission after emergency surgery: a nationwide in-patient registry study. *British Journal of Anaesthesia*. 2022;129(4):527-535.
72. Ohbe H, Matsui H, Kumazawa R, Yasunaga H. Postoperative intensive care unit admission following major elective surgery: a nationwide inpatient database study. *European Journal of Anaesthesiology*. 2022;39(5):436-444.
73. Ohbe H, Matsui H, Yasunaga H. ICU versus high-dependency care unit for patients with acute myocardial infarction: a nationwide propensity score-matched cohort study. *Critical Care Medicine*. 2022;50(6):977-985.
74. Ohbe H, Sasabuchi Y, Kumazawa R, Matsui H, Yasunaga H. Intensive care unit occupancy in Japan, 2015–2018: a nationwide inpatient database study. *Journal of Epidemiology*. 2022;32(12):535-542.
75. Ohbe H, Sasabuchi Y, Matsui H, Yasunaga H. Impact of the COVID-19 pandemic on critical care utilization in Japan: a nationwide inpatient database study. *Journal of Intensive Care*. 2022;10(1):51.
76. Ohbe H, Tagami T, Ogura T, Matsui H, Yasunaga H. Low-flow duration and outcomes of extracorporeal cardiopulmonary resuscitation in adults with in-hospital cardiac arrest: a nationwide inpatient database study. *Critical Care Medicine*. 2022;50(12):1768-1777.
77. Ohbe H, Tagami T, Uda K, Matsui H, Yasunaga H. Incidence and outcomes of in-hospital cardiac arrest in Japan 2011–2017: a nationwide inpatient database study. *Journal of Intensive Care*. 2022;10(1):10.
78. Ohisa K, Yamana H, Morita K, Matsui H, Fushimi K, Yasunaga H. Association between subcutaneous steroid injection for extravasation of vesicant anticancer drugs and skin ulcers requiring surgery. *European Journal of Oncology Nursing*. 2022;58:102119.
79. Okada A, Hashimoto Y, Goto T, Yamaguchi S, Ono S, Ikeda Kurakawa K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T. A machine learning-based predictive model to identify patients who failed to attend a follow-up visit for diabetes care after recommendations from a national screening program. *Diabetes Care*. 2022;45(6):1346-1354.
80. Okada A, Yamana H, Yamaguchi S, Kurakawa K, Michihata N, Matsui H, Fushimi K, Nangaku M, Yamauchi

- T, Yasunaga H, Kadowaki T. Effect of branched-chain amino acid infusion on in-hospital mortality of patients with hepatic encephalopathy and endstage kidney disease: a retrospective cohort study using a national inpatient database. *Journal of Renal Nutrition*. 2022;32(4):432-440.
81. Okada A, Yasunaga H. Prevalence of non-communicable diseases in Japan using a newly developed administrative claims database covering young, middle-aged, and elderly people. *JMA Journal*. 2022;5(2):190-198.
 82. Okubo Y, Nishi A, Uda K, Miyairi I, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yasunaga H. Financial Incentives for Infection Prevention and Antimicrobial Stewardship to Reduce Antibiotic Use: A Japan's Nationwide Observational Study. *Journal of Hospital Infection*. 2022;131:89-98.
 83. Ono S, Ishimaru M, Yokota I, Konishi T, Okada A, Ono Y, Matsui H, Itai S, Yonenaga K, Tonosaki K, Watanabe R, Hoshi K, Yasunaga H. Risk of post-extraction bleeding with direct oral anticoagulant compared with warfarin: Retrospective cohort study using large scale claims data in Japan. *Thrombosis Research*. 2022;222:24-30.
 84. Sakamoto T, Fujiogi M, Ishimaru M, Matsui H, Fushimi K, Yasunaga H. Comparison of postoperative infection after emergency inguinal hernia surgery with enterectomy between mesh repair and non-mesh repair: a national database analysis. *Hernia*. 2022;26(1):217-223.
 85. Sato M, Yamana H, Ono S, Ishimaru M, Matsui H, Yasunaga H. Trends in prophylactic antibiotic use for tooth extraction from 2015 to 2018 in Japan: an analysis using a health insurance claims database. *Journal of Infection and Chemotherapy*. 2022;28(4):504-509.
 86. Shibahashi K, Konishi T, Ohbe H, Yasunaga H. Cost-effectiveness analysis of termination-of-resuscitation rules for patients with out-of-hospital cardiac arrest. *Resuscitation*. 2022;180:45-51.
 87. Shibahashi K, Ohbe H, Yasunaga H. Association between intensive care unit admission practices and outcomes in patients with isolated traumatic subarachnoid hemorrhage: A nationwide inpatient database analysis in Japan. *Neurocritical Care*. 2022;37(2):497-505.
 88. Shibahashi K, Ohbe H, Yasunaga H. Urban-rural inequalities in care and outcomes of severe traumatic brain injury: A nationwide inpatient database analysis in Japan. *World Neurosurgery*. 2022;163:e628-e634.
 89. Shibahashi K, Ohbe H, Yasunaga H. Adjuvant oral tranexamic acid and reoperation after burr-hole surgery in patients with chronic subdural hematoma: Propensity score-matched analysis using a nationwide inpatient database. *Journal of Neurosurgery*. 2022;138(2):430-436.
 90. Shigemi D, Tabuchi T, Okawa S, Yasunaga H. Association between health literacy and COVID-19 prevention behaviors among pregnant and postpartum women. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2022;35(25):9971-9977.
 91. Sugawara C, Yamana H, Sasagawa E, Yonezawa K, Hikita N, Morita K, Matsui H, Yasunaga H, Haruna M. Factors associated with surgical treatment in postpartum women with mastitis or breast abscess: A retrospective cohort study. *Breastfeeding Medicine*. 2022;17(3):233-238.
 92. Sugihara T, Kamei J, Yasunaga H, Sasabuchi Y, Fujimura T. Prescription of Choreito, a Japanese Kampo medicine, with antimicrobials for treatment of acute cystitis: a retrospective cohort study. *Antibiotics*. 2022;11:1840.
 93. Suzuki J, Sasabuchi Y, Hatakeyama S, Matsui H, Sasahara T, Morisawa Y, Yamada T, Fushimi K, Yasunaga H. Additional effect of azithromycin over β -lactam alone for severe community-acquired pneumonia-associated acute respiratory distress syndrome: a retrospective cohort study. *Pneumonia*. 2022 ;14(1):1.
 94. Suzuki T, Michihata N, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Low- versus high-concentration intravenous immunoglobulin for children with Kawasaki disease in the acute phase. *International Journal of Rheumatic Diseases*. 2022;25(5):576-583.
 95. Suzuki Y, Kaneko H, Okada A, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Impact of Glucose Tolerance on Incident Proteinuria in Prediabetes: Analysis of a Nationwide Population-Based Dataset. *American Journal of Nephrology*. 2022;53(4):307-315.
 96. Suzuki Y, Kaneko H, Okada A, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Yano Y, Kamiya K, Matsunaga A, Ako J, Fukui A, Yokoo T, Nishiyama A, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Risk for Proteinuria in Newly Defined Hypertensive People Based on the 2017 American College of Cardiology/American Heart Association Blood Pressure Guideline. *American Journal of Cardiology*. 2022;168:83-89.
 97. Suzuki Y, Kaneko H, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I. Comparison of cardiovascular outcomes between SGLT2 inhibitors in diabetes mellitus. *Cardiovascular Diabetology*. 2022;21(1):67.
 98. Suzuki Y, Kaneko H, Okada A, Itoh H, Morita K, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Fukui A, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Change in Cardiovascular Health Metrics and Risk for Proteinuria Development: Analysis of a Nationwide Population-Based Database. *American Journal of Nephrology*. 2022;53(2-3):240-248.
 99. Suzuki Y, Kaneko H, Okada A, Matsuoka S, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Node K, Yamauchi T, Yasunaga H, Komuro I. Prediabetes in Young Adults and Its Association with Cardiovascular Health Metrics in the Progression to Diabetes. *Journal of Clinical Endocrinology & Metabolism*. 2022;107(7):1843-1853.
 100. Suzuki Y, Kaneko H, Okada A, Nangaku M, Yasunaga H, Komuro I. Kidney outcomes in patients with diabetes mellitus did not differ between individual sodium-glucose cotransporter-2 inhibitors. *Kidney International*. 2022;102(5):1147-1153.
 101. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I. Glycemic status and the association of change in blood pressure with incident cardiovascular disease. *American Heart Journal*. 2022;254:48-56.
 102. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Matsunaga A, Node K, McEvoy JW, Lam CSP, Oparil S, Yasunaga H, Komuro I. Association of Cardiovascular Health Metrics with Risk of Transition to Hypertension in Non-Hypertensive Young Adults. *American Journal of Hypertension*. 2022;35(10):858-866.
 103. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Yamaguchi S, Michihata N, Jo T, Takeda N, Morita H, Node K, Kim HC, Viera AJ, Oparil S, Yasunaga H, Komuro I. Age-Dependent Relationship of Hypertension Subtypes with Incident Heart Failure. *JAHA*. 2022;11(9):e025406.
 104. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Ueno K, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yokota I, Node K, Yasunaga H, Komuro I. Dose-dependent relationship of blood pressure and glycemic status with risk of aortic dissection and aneurysm.

European Journal of Preventive Cardiology. 2022;29(18):2338-2346.

105. Tagami T, Yamakawa K, Endo A, Hayakawa M, Ogura T, Hirayama A, Yasunaga H. Japanese Multicenter Research of COVID-19 by Assembling Real-world Data: A Study Protocol. *Annals of Clinical Epidemiology*. 2022;4(3):92-100.
106. Taguchi R, Shigemi D, Yasunaga H. Trends in antihypertensive prescription for pregnant women with hypertension and their peripartum outcomes before and after label and guideline revisions in Japan. *Hypertension Research*. 2022;45(11):1823-1831.
107. Tahara S, Hattori Y, Aso S, Uda K, Kumazawa R, Matsui H, Fushimi K, Yasunaga H, Morita A. Trends in Surgical Procedures for Spontaneous Intracerebral Hemorrhage in Japan. *Journal of Stroke and Cerebrovascular Diseases*. 2022;31(9):106664.
108. Ueno K, Kamiya K, Kaneko H, Okada A, Itoh H, Fujiu K, Takeda N, Morita H, Michihata N, Jo T, Yasunaga H, Komuro I. Acute-phase initiation of cardiac rehabilitation for improvement in activities of daily living in patients hospitalized for acute heart failure. *Journal of Cardiovascular Development and Disease*. 2022;9(4):97.
109. Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Node K, Yamauchi T, Yasunaga H, Komuro I. Relationship of normal-weight central obesity with the risk for heart failure and atrial fibrillation: Analysis of a nationwide health check-up and claims database. *European Heart Journal Open*. 2022;2(3):oeac026.
110. Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Yasunaga H, Komuro I. Clinical utility of simple subjective gait speed for the risk stratification of heart failure in a primary prevention setting. *Scientific Reports*. 2022;12(1):11641.
111. Urushiyama H, Jo T, Hasegawa W, Yokoyama A, Ando T, Sakamoto Y, Kumazawa R, Uda K, Michihata N, Awano N, Hiroki M, Fushimi K, Yasunaga H, Nagase T. Effect of nintedanib on acute exacerbations of fibrosing interstitial lung diseases: A national database study in Japan. *ERJ Open Research*. 2022;8(4):00209-2022.
112. Wada Y, Takahashi H, Matsui H, Yasunaga H, Fujiwara H, Sasabuchi Y. Adhesion barriers and intraperitoneal or uterine infections after cesarean section: A retrospective cohort study. *Surgery*. 2022;172(6):1722-1727.
113. Yamana H, Sasabuchi Y, Matsui H, Michihata N, Jo T, Yasunaga H. Maoto plus neuraminidase inhibitor versus neuraminidase inhibitor alone for reducing hospitalisation in older adults with seasonal influenza. *Annals of Clinical Epidemiology*. 2022;4:120-128
114. Yamana H, Tsuchiya A, Horiguchi H, Morita S, Kuroki T, Nakai K, Nishimura H, Jo T, Fushimi K, Yasunaga H. Validity of a model using routinely collected data for identifying infections following gastric, colorectal, and liver cancer surgeries. *Pharmacoepidemiology and Drug Safety*. 2022;31(4):452-460.
115. Yamazaki R, Ohbe H, Matsuda Y, Kito S, Shigeta M, Morita K, Matsui H, Fushimi K, Yasunaga H. Early electroconvulsive therapy in patients with bipolar depression: a propensity score-matched analysis using a nationwide inpatient database. *Journal of Affective Disorders*. 2022;312:245-251.
116. Zhang L, Michihata N, Matsui H, Fushimi K, Yasunaga H, Tsuda Y, Tanaka S, Kobayashi H. Preoperative Arterial Embolization and Wound Complications After Resection of Malignant Bone Tumor in the Pelvis: A Nationwide Database Study. *Japanese Journal of Clinical Oncology*. 2022;52(10):1176-1182.

<中山健夫>

論文発表

1. Tsuji N, Takahashi Y, Sakai M, Ohtera S, Kaneyama J, Cho K, Kato G, Ohtsuru S, Nakayama T. Trend of anticoagulant therapy in elderly patients with atrial fibrillation considering risks of cerebral infarction and bleeding. *Sci Rep*. 2023 Jan 5;13(1):192.
2. Ishizaki M, Toyama M, Imura H, Takahashi Y, Nakayama T. Tracheostomy decannulation rates in Japan: a retrospective cohort study using a claims database. *Sci Rep*. 2022 Nov 17;12(1):19801.
3. Sakai F, Hirata K, Igarashi H, Takeshima T, Nakayama T, Sano H, Kondo H, Shibasaki Y, Koga N. A study to investigate the prevalence of headache disorders and migraine among people registered in a health insurance association in Japan. *J Headache Pain*. 2022 Jun 23;23(1):70. doi: 10.1186/s10194-022-01439-3. Erratum in: *J Headache Pain*. 2022 Dec 28;23(1):164.
4. Hoshino N, Hida K, Fukui Y, Takahashi Y, Nakayama T, Obama K. Relationship between diverting stoma and adjuvant chemotherapy in patients with rectal cancer: a nationwide study using the National Database of Health Insurance Claims and Specific Health Checkups of Japan. *Int J Clin Oncol*. 2022 Mar;27(3):545-552.
5. Sengoku T, Ishizaki T, Goto Y, Iwao T, Ohtera S, Sakai M, Kato G, Nakayama T, Takahashi Y. Prevalence of type 2 diabetes by age, sex and geographical area among two million public assistance recipients in Japan: a cross-sectional study using a nationally representative claims database. *J Epidemiol Community Health*. 2022 Apr;76(4):391-397.

学会発表

1. 中山健夫. 京都大学における医療データ活用人材育成プログラム. 日本臨床疫学会第5回年次学術大会康永秀生理事企画シンポジウム「保健医療介護ビッグデータ研究. 人材育成の取り組み」早稲田大学 2022.11.13

<田宮菜奈子>

論文発表

1. Yu Sun, Masao Iwagami, Nobuo Sakata, Tomoko Ito, Ryota Inokuchi, Jun Komiyama, Naoaki Kuroda, Nanako Tamiya. Evaluation of enhanced home care support clinics regarding emergency home visits, hospitalization, and end-of-life care: a retrospective cohort study in a city of Japan. *BMC Health Serv Res* 2023 Feb 3;23(1):115.
2. Atsuyuki Watanabe, Ryota Inokuchi, Toshiki Kuno, Kazuaki Uda, Jun Komiyama, Motohiko Adomi, Yoshiko Ishisaka, Toshikazu Abe, Nanako Tamiya, Masao Iwagami. Pulse Methylprednisolone versus Dexamethasone in COVID-19: A Multicenter Cohort Study. *Crit Care Explore*. 2023;5(4):e0886.

学会発表

1. 宇田和晃、井口竜太、小宮山潤、田宮菜奈子、岩上将夫. 高齢 COVID-19 入院患者における入院早期リハビリテーションと在院死亡との関連、日本臨床疫学会第 5 回年次学術大会、2022 年
2. 田宮菜奈子、シンポジウム：リアルワールドデータ研究の人材育成、日本臨床疫学会第 5 回年次学術大会、2022 年
3. 岩上将夫、シンポジウム：臨床疫学研究に使われるデータベース：電子カルテデータベース

<笹渕裕介>

論文発表

1. Ono S, Sasabuchi Y, Ishimaru M, Ono Y, Matsui H, Yasunaga H. Short-term effects of reduced cost sharing on childhood dental care utilization and dental caries prevention in Japan. *Community Dent Oral Epidemiol.* 2023;51:228-235.
2. Yoshida T, Uchino S, Sasabuchi Y. Epidemiology of constipation in critically ill patients and its impact on in-hospital mortality: a retrospective observational study. *J Anesth.* 2022;36:349-358.
3. Suzuki R, Uchino S, Sasabuchi Y, Kawarai Lefor A, Sanui M. Dopamine use and its consequences in the intensive care unit: a cohort study utilizing the Japanese Intensive Care PATient Database. *Crit Care.* 2022;26:90.
4. Chinen T, Sasabuchi Y, Matsui H, Yamaguchi H, Yasunaga H. Oxaliplatin- versus cisplatin-based regimens for elderly individuals with advanced gastric cancer: a retrospective cohort study. *BMC Cancer.* 2022;22:460.
5. Ookawara S, Ito K, Sasabuchi Y, Miyahara M, Miyashita T, Takemi N, Nagamine C, Nakahara S, Horiuchi Y, Inose N, Shiina M, Murakoshi M, Sanayama H, Hirai K, Morishita Y. Cerebral oxygenation and body mass index association with cognitive function in chronic kidney disease patients without dialysis: a longitudinal study. *Sci Rep.* 2022;12:10809.
6. Iwasaki Y, Sasabuchi Y, Horikita S, Furukawa T, Shiotsuka J, Lefor AK, Sanui M. The effect of preoperative sodium-glucose cotransporter 2 inhibitors on the incidence of perioperative metabolic acidosis: A retrospective cohort study. *BMC Endocr Disord.* 2022;22:209.
7. Wada Y, Takahashi H, Matsui H, Yasunaga H, Fujiwara H, Sasabuchi Y. Adhesion barriers and intraperitoneal or uterine infections after cesarean section: A retrospective cohort study. *Surgery.* 2022;172:1722-1727.
8. Hayashi K, Sasabuchi Y, Matsui H, Nakajima M, Ohbe H, Fushimi K, Ono K, Yasunaga H. Additive effectiveness of acrylonitrile-co-methallyl sulfonate surface-treated membranes in the treatment of pneumonia: A propensity score-matched retrospective cohort study. *Artif Organs.* 2023;47:408-416.
9. Shiotsuka J, Uchino S, Sasabuchi Y, Masuyama T, Lefor AK, Sanui M. Reducing the effect of immortal time bias affects the analysis of prevention of delirium by suvorexant in critically ill patients: A retrospective cohort study. *PLoS One.* 2022;17:e0277916.
10. Hayashi K, Sasabuchi Y, Matsui H, Nakajima M, Otawara M, Ohbe H, Fushimi K, Ono K, Yasunaga H. Does early excision or skin grafting of severe burns improve prognosis? A retrospective cohort study. *Burns.* 2023;49:554-561.
11. Ohbe H, Ouchi K, Miyamoto Y, Ishigami Y, Matsui H, Yasunaga H, Sasabuchi Y. One-Year Functional Outcomes After Invasive Mechanical Ventilation for Older Adults With Preexisting Long-Term Care-Needs. *Crit Care Med.* 2023:e005822. Epub ahead of print.
12. Murooka Y, Sasabuchi Y, Takazawa T, Matsui H, Yasunaga H, Saito S. Long-Term Prognosis Following Early Rehabilitation in the ICU: A Retrospective Cohort Study. *Crit Care Med.* 2023. Epub ahead of print.
13. Wada Y, Takahashi H, Sasabuchi Y, Usui R, Ogoyama M, Suzuki H, Ohkuchi A, Fujiwara H. Maternal outcomes of placental abruption with intrauterine fetal death and delivery routes: A nationwide observational study. *Acta Obstet Gynecol Scand.* 2023. Epub ahead of print.

学会発表

1. 笹渕裕介. 自治医科大学における人材育成の取り組み. 第5回日本臨床疫学会年次学術大会. 2022/11/13 (東京).

H. 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

厚生労働行政推進調査事業補助金

政策科学総合研究事業（政策科学推進研究事業）

診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資
するための研究

ビッグデータ研究実践能力およびデータハンドリング技術養成プログラムの実践

研究代表者 康永秀生 東京大学大学院医学系研究科臨床疫学・経済学 教授

研究要旨： NDB・DPC 等の大規模データ解析に精通した研究者を育成することは急務である。令和4年度研究では、令和3年度に開発したコンテンツ（NDBのeラーニング型教育コンテンツやオンデマンド型教育用ビデオコンテンツから構成される20コンテンツ）を用いた人材育成プログラムを実践した。医療従事者のみならず多くの職種から1158人が参加登録し、のべ7411人から受講後アンケートの回答を得た。理解度・満足度ともに概ね高い結果となった。NDB、DPCなどの大規模データベースを用いた研究を引き続き実施し、2022年は英文原著論文が117編出版された。

日常臨床のクリニカル・クエスチョンをデータベース活用により解明する研究実践能力を持つ研究者を多数育成し、データハンドリング技術と臨床研究実践能力の両方に長けた人材も多数育成することにより、わが国の大規模データベース研究の進歩を加速できる。

A. 研究目的

近年、保健医療介護のリアルワールドデータを収集した大規模データベース研究が拡大している。平成23年からNDBデータの第三者提供が開始され、提供件数は年々増加している。NDBと介護DBとの連結した「医療・介護データ等の解析基盤（HIC）」の運用が予定されており、さらに今後NDBはDPCデータベースと連結も予定されており、保健医療介護データベ

ス研究の素地は整いつつある。

一方で、データベース研究に精通した研究者は非常に限られている。保健医療介護の実態を理解し、なおかつデータ解析に精通した研究者を育成することは喫緊の課題である。そのために、臨床医学・疫学・統計学の知識に加えて、データベースに関連する知識や技能の教育が必須である。

また、令和2年度には国内3カ所のNDBオ

ンサイトリサーチセンターが一般開放されたものの、コロナ禍において稼働は制限されたままである。同センターの運用方法については前例がなく、各センターが模索している状態にある。

研究代表者らは、平成29-30年度および平成31-令和2年度厚生労働科学研究において、保健医療介護データベース研究者を育成する短期的な教育プログラムを開発した。これらを用いて、2019年と2020年にNDB・DPCデータベース研究人材育成<短期集中セミナー>を実践し、参加者によるプログラム評価も実施した。また、大規模データベース研究を行い、2019年に約50編、2020年に約70編、2021年に約100編の英文原著論文を出版した。

本研究では、上記研究をさらに発展させ、令和3年度は先行研究にて開発された人材育成プログラムを改良し、NDBのeLearningプログラムを開発した。さらに、オンサイトリサーチセンター利用者の増加を想定し、その運用方法を具体化した。

令和4年度は、令和3年度に開発したプログラムを用いた人材育成を実践するとともに、オンサイトリサーチセンターの利用を推進するとともに、大規模データベース研究を継続的に実践した。

本研究は、わが国の大規模データベース研究において不足している「人材育成」を最重視し、大規模データベース研究のための種々の技術を一般化し体系的なプログラムを構築した上で、それらを多数の研究者

等に利活用してもらえるように社会実装を試みる点が、既存研究にない独創的な点である。

B. 研究方法

◆研究計画を遂行するための研究体制

先行の研究班のもとに、大規模データベース利活用のための若手研究者コンソーシアムを立ち上げた。疫学・統計学・医療情報学等の社会医学系研究者に加えて、臨床医学、健康科学・看護学などの若手研究者も多数参画した。研究班メンバーが若手研究者を直接指導し、そこで日夜培われ開発される個別技術（データハンドリング技術、観察研究における統計解析技術など）を体系化・一般化し、既存の知識と合わせた、種々のビッグデータに適用可能な教育プログラムを開発した。

研究代表者はこれまで多くの臨床家との共同研究を行っており、研究協力者として総勢約200名の臨床家を動員し、ビッグデータからエビデンスを産み出す恒常的なサイクルを回し、英文原著論文を量産する体制を築いている。

1. 令和3年度に開発したプログラムを用いた人材育成の実践

先行研究において、ビッグデータ研究実践能力養成プログラムおよびビッグデータハンドリング技術養成プログラムを開発した。前者は、日常臨床から臨床的・クエスチョンを紡ぎ出し、検証可能な

リサーチ・クエスチョンに構造化し、既存のビッグデータを用いてリサーチ・クエスチョンを解き明かす臨床研究・疫学研究・ヘルスサービスリサーチの実践能力を養成するプログラムである。後者は、テラバイト級のデータベースから個別の研究目的に沿うデータセットを抽出するデータハンドリング技術、膨大なテキストデータを含む大規模かつ複雑な構造のデータを研究用の扱いやすいデータベースに再構築するなどのデータベースマネジメント技術を養成するプログラムである。上記について、現下のコロナ禍の状況においても教育実践が可能となるように、令和3年度にオンラインによるプログラムに改良し、令和4年度にプログラムを実践した。

具体的には、「NDB・DPC データベース研究人材育成 Webinar」を2023年1月10日(火)～1月20日(金)に開催した。Webによる動画配信(講義とeLearning)の形式とした。申請者に事前に登録してもらったアドレスに動画のURLを送信し、オンデマンドで視聴してもらった。配信期間中は全ての講義をいつでも視聴することを可能とした。

主催は厚生労働省科学研究・保健医療介護ビッグデータ人材育成研究班(研究代表者:東京大学 康永秀生)とし、日本臨床疫学会が共催となった。事前登録制であり、無料とした。

Webinar の概要と詳細については別紙1・

2を参照。

【本 Webinar の特徴】

- ・保健医療介護ビッグデータ研究で実績のある講師陣による講義・演習を受けることができる
- ・短期集中で大規模データベース研究の計画立案から統計解析まで学習できる

【受講対象】

保健医療介護ビッグデータ研究に興味のある方。

- ・医療・介護従事者
 - ・大学関係者
 - ・各種研究機関に勤務する研究者
 - ・企業に所属する者
 - ・医療政策の担当者
- など

【コンテンツ】

1. NDB・DPC 等の公的データベースの動向
2. NDB ガイドラインの理解度テスト
3. SQL 入門
4. SQL を用いたレセプトデータのハンドリング
5. NDB を用いた研究
6. DPC データを用いた研究 1
7. DPC データを用いた研究 2
8. DeSC データを用いた研究
9. JMDC データを用いた研究
10. 傾向スコア
11. 高次元傾向スコア
12. 操作変数法

- 13. 不連続回帰デザイン・差の差分析
- 14. 時間依存性交絡と周辺構造モデル
- 15. 生存時間分析における競合リスクモデル
- 16. 欠測データの取り扱いと多重代入法
- 17. マルチレベル分析
- 18. 自己対照ケースシリーズ
- 19. 機械学習
- 20. データベースにおけるバリデーション研究

【アンケートによる評価】

各講義の後に無記名アンケートを実施した。各講義のアンケートに回答すると、各講義資料の pdf がダウンロード可能になる、というインセンティブを与えた。

Q1. 本講義の難易度はいかがでしたか？
(入力必須)

- とてもわかりやすい
- わかりやすい
- どちらともいえない
- わかりにくい
- とてもわかりにくい

Q2. 本講義の進行速度はいかがでしたか？
(入力必須)

- とても速い
- 速い
- ちょうどよい
- 遅い
- とても遅い

Q3. 本講義の満足度はいかがでしたか？
(入力必須)

- とても満足
- 満足
- どちらともいえない
- 不満
- とても不満

Q4. 本講義で学べて良かったことを教えてください
(入力必須、自由記載)

Q5. 本講義で疑問に感じたこと、不明点を教えてください
(入力必須、自由記載)

2. 大規模データベース研究の実践

本プログラム受講者をはじめ、多施設から多くの研究者と共同研究体制を構築し、NDB, DPC 等のデータベースを用いた臨床研究、疫学研究、医療経済・政策研究、ヘルスサービスリサーチを実践し、原著論文出版、学会発表を実践した。

C. 研究結果

1. 令和3年度に開発したプログラムを用いた人材育成の実践

コンテンツ 1-20 の具体的な内容は巻末資料に提示する。

期間中、参加登録者は 1,158 名であった。内訳は、医療従事者 717 人 (医師 411 人、薬剤師 87 人、看護師 60 人、歯科医師 10 人、その他コメディカル 149 人)、製薬会社

187人、医療機器メーカー10人、保険会社6人、コンサルティング会社13人、医療IT企業26人、生物統計家6人、データサイエンティスト48人、診療情報管理士3人、審査支払機関5人、CRO18人、学生54人、教員18人、研究員21人、その他26人であった。

アンケート回答はのべ7411人から得た。受講した一人当たり平均のアンケート回答回数は6.4回であった。

各コンテンツ（講義番号1-20）に対するQ1,Q2,Q3に対する回答結果の分布をそれぞれ図1,2,3に示す。

Q1.本講義の難易度について、「とてもわかりやすい」と「わかりやすい」を合計した割合は、講義番号3,5,6,7,8,9,10,16,17,18,20は90%を超え、講義番号1,4,11,12,13,14,15,19は80%を超えていたのに対し、講義番号2は49%と低かった。

Q2.本講義の進行速度について、いずれの講義番号も「ちょうどよい」が80%を超えていた。

Q3.本講義の満足度について、「とても満足」と「満足」を合計した割合は、課題番号1,3,4,5,6,7,8,9,10,11,12,13,14,16,17,18,20は90%を超え、講義番号15,19は80%を超えていたのに対し、講義番号2は59%と低かった。

なお、参加者の属性に関わらず、これらの傾向は一定であった。

Q4.「本講義で学べて良かったこと」（入力必須、自由記載）の結果の抜粋を表1に示

す。

Q5.「本講義で疑問に感じたこと、不明点」（入力必須、自由記載）結果の抜粋を表2に示す。

2. 大規模データベース研究の実践

116編の英文原著論文を出版した。その一覧は下記の「G.研究発表」に示す。

D. 考察

1. 令和3年度に開発したプログラムを用いた人材育成の実践

対面型およびオンライン教育プログラムを改良し、コロナ禍においても双方向型の教育が可能なオンラインプログラムを実践運用した。

参加登録者は千人を超え、そのうち医療従事者が約62%であり、製薬会社・医療機器メーカーなど一般企業の従事者は21%であった。その他、生物統計家・データサイエンティスト等が約6%、学生が約5%であり、参加者の所属は多岐にわたった。

各コンテンツ（講義番号1-20）に対する講義の難易度について、「とてもわかりやすい」と「わかりやすい」を合計した割合は、ほぼすべて80%を超えており、概ね良好であった。講義の進行速度について、いずれの課題番号も「ちょうどよい」が80%を超えていた。講義の満足度について、「とても満足」と「満足」を合計した割合は、ほぼ80%を超えていた。

課題番号2は他と異なり、設問に回答していく形であり、正誤を確認した後に解説を

加えるという形式であるため、受講者が難しさを感じたかもしれない。

Q4. 「本講義で学べて良かったこと」において、非常に多くのコメントを受け取った。内容はほぼ本講義に対する肯定的なものである。Q5. 「本講義で疑問に感じたこと、不明点」に関するコメントは、Q4 に比べて圧倒的に低く、内容的には本講義で紹介されなかった内容の追加を求める声が圧倒的に多かった。コメントを参考に今後のコンテンツの改良を行う予定である。

2. 大規模データベース研究の実践

上記1を通じて、大規模データベース研究の実践をさらに進めた。2019年の約50編、2020年の約70編、2021年の約100編にひきつづき、2022年には117編の原著論文を出版した。コンスタントに増加しており、本研究の人材育成プログラムの効果を示唆するものである。

E. 結論

本研究は、厚生労働省が進める NDB 等データベースの高度利活用に直接反映される研究である。NDB ばかりでなく、あらゆる保健・医療・介護ビッグデータに対応できる人材育成を図ることにより、ビッグデータのデータハンドリング、データベースマネジメント等に関する総合的な技術を持つ人材を多数育成してきた。

近年、保健医療介護データベースのインフラ整備が進められている。それらのイ

ンフラを活用できる人材の育成にさらに貢献し、今後整備される種々の統合データベースの稼働の際には全省的な政策課題に関する研究・知見の提供にも貢献できる。さらに、日常臨床のクリニカル・クエスチョンをデータベース活用により解明する研究実践能力を持つ研究者を多数育成し、データハンドリング技術と臨床研究実践能力の両方に長けた人材も多数育成することにより、わが国の大規模データベース研究の進歩を加速できる。わが国の大規模データベース研究の技術水準を世界トップレベルに向上させ、わが国発のエビデンスを量産できることが期待される。

F. 健康危険情報

なし

G. 研究発表

論文発表

1. Abe H, Sumitani M, Matsui H, Inoue R, Fushimi K, Uchida K, Yasunaga H. Use of naldemedine is associated with reduced incidence of hyperactive delirium in cancer patients with opioid-induced constipation: a nationwide retrospective cohort study in Japan. *Pharmacotherapy*. 2022;42(3):241-249.
2. Abe H, Sumitani M, Matsui H, Inoue R, Konishi M, Fushimi K, Uchida K, Yasunaga H. Gabapentinoid use is associated with reduced occurrence of hyperactive delirium in older cancer patients undergoing chemotherapy: a nationwide retrospective cohort study in Japan. *Anesthesia & Analgesia*. 2022;135(2):362-369.
3. Aso S, Matsui H, Fushimi K, Yasunaga H. Effect of Dobutamine on Patients with Septic Shock: A Retrospective Nationwide Study. *Annals of Clinical Epidemiology*. 2022;4(2): 41-48.
4. Awano N, Jo T, Izumo T, Inomata M, Morita K, Matsui H, Fushimi K, Urushiyama H, Nagase T, Yasunaga H. Recombinant human soluble thrombomodulin for acute exacerbation of idiopathic pulmonary fibrosis: A nationwide observational study. *Journal of Intensive Care*. 2022;10(1):14.

5. Chinen T, Sasabuchi Y, Matsui H, Yamaguchi H, Yasunaga H. Oxaliplatin- versus cisplatin-based regimens for elderly individuals with advanced gastric cancer: A retrospective cohort study. *BMC Cancer*. 2022;22(1):460.
6. Fujita A, Hashimoto Y, Matsui H, Yasunaga H, Aihara M. Recent trends in glaucoma surgery: a nationwide database study in Japan, 2011–2019. *Japanese Journal of Ophthalmology*. 2022;66(2):183-192.
7. Fujita A, Hashimoto Y, Okada A, Obata R, Aihara M, Matsui H, Yasunaga H. Association Between Proteinuria and Retinal Vein Occlusion in Individuals with Preserved Renal Function: A Retrospective Cohort Study. *Acta Ophthalmologica*. 2022;100(7):e1510-e1517
8. Fukui R, Kaneko H, Okada A, Yano Y, Itoh H, Matsuoka S, Morita K, Kiriyama H, Kamon T, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Nishiyama A, Node K, Yokoo T, Nangaku M, Yasunaga H, Komuro I. Semiquantitative assessed proteinuria and risk of heart failure: Analysis of a nationwide epidemiological database. *Nephrology Dialysis Transplantation*. 2022;37(9):1691-1699.
9. Furuya H, Nakajima M, Ikeda K, Nakamura K, Ohbe H, Aso S, Kumazawa R, Iwamoto T, Iwata A, Furuta S, Matsui H, Fushimi K, Yasunaga H, Nakajima H. Prognosis and treatment of myositis-associated severe interstitial lung disease. *Arthritis Care & Research*. 2022;74(3):478-483.
10. Fuseda Y, Michihata N, Kumazawa R, Fushimi K, Yasunaga H. Annual Trend in Lumbar Puncture for Infants Younger Than 3 Months Hospitalized with Suspected Serious Bacterial Infection: A Nationwide Inpatient Database Study. *Pediatric Infectious Disease Journal*. 2022;41(8):631-635.
11. Hamada S, Nakajima M, Kaszynski RH, Otaka S, Goto H, Matui H, Fushimi K, Yamaguchi Y, Yasunaga H. Association between adjunct clindamycin and in-hospital mortality in patients with necrotizing soft tissue infection due to group A Streptococcus: A nationwide cohort study. *European Journal of Clinical Microbiology & Infectious Diseases*. 2022;41(2):263-270.
12. Hashimoto Y, Kaneko H, Okada A, Matsui H, Yasunaga H, Aihara M, Obata R. Association between retinal vein occlusion and Life's Simple 7 cardiovascular health metrics: A large claims database study. *Ophthalmology Retina*. 2022;6(8):684-692.
13. Hashimoto Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, Kaburaki T. Incidence of sympathetic ophthalmia after inciting events: a national database study in Japan. *Ophthalmology*. 2022;129(3):344-352.
14. Hashimoto Y, Michihata N, Yamana H, Shigemi D, Morita K, Matsui H, Yasunaga H, Aihara M. Safety of topical ophthalmic antibiotics in pregnant women with hordeola, chalazia, blepharitis, or bacterial conjunctivitis: propensity score analyses. *EYE*. 2022 ;36(5):1066-1073.
15. Hashimoto Y, Yamana H, Michihata N, Shigemi D, Ishimaru M, Matsui H, Yasunaga H, Aihara M. Eye drops for dry eye disease during pregnancy and adverse neonatal outcomes: high-dimensional propensity score analyses. *Ophthalmic Epidemiology*. 2022;29(4):384-393.
16. Hirano T, Nakajima M, Ohbe H, Kaszynski RH, Iwasaki Y, Arakawa Y, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H. Corticosteroid use with extracorporeal cardiopulmonary resuscitation for out-of-hospital cardiac arrest: A nationwide observational study. *Resuscitation Plus*. 2022;12:100308.
17. Hirano Y, Kaneko H, Konishi T, Itoh H, Matsuda S, Kawakubo H, Uda K, Matsui H, Fushimi K, Daiko H, Itano O, Yasunaga H, Kitagawa Y. Short-Term Outcomes of Epidural Analgesia in Minimally Invasive Esophagectomy for Esophageal Cancer: Nationwide Inpatient Data Study in Japan. *Annals of Surgical Oncology*. 2022;29(13):8225-8234.
18. Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H. Association between nurse aide staffing and patient mortality after major cancer surgeries in acute care settings: A retrospective cohort study using a national inpatient database. *Nursing and Health Sciences*. 2022;24(1):283-292.
19. Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H. Dose-response association between nurse staffing and patient outcomes following major cancer surgeries using a nationwide inpatient database in Japan. *Journal of Clinical Nursing*. 2022;31(17-18):2562-2573.
20. Honda A, Michihata N, Iizuka Y, Uda K, Morita K, Mieda T, Takasawa E, Ishiwata S, Tajika T, Matsui H, Fushimi K, Yasunaga H, Chikuda H. Risk factors for severe lower extremity ischemia following venoarterial extracorporeal membrane oxygenation: An analysis using a nationwide inpatient database. *Trauma Surgery & Acute Care Open*. 2022;7:e000776.
21. Hosoi T, Yamana H, Tamiya H, Matsui H, Fushimi K, Akishita M, Yasunaga H, Ogawa S. Association between comprehensive geriatric assessment and polypharmacy at discharge in patients with ischaemic stroke: a nationwide retrospective cohort study. *eClinicalMedicine*. 2022;50:101528.
22. Inooka Y, Yamana H, Shinoda Y, Inokuchi H, Matsui H, Fushimi K, Yasunaga H, Haga N. Predictive factors for oral intake recovery after acute stroke: analysis of a Japanese nationwide inpatient database. *Dysphagia* 2022;37(6):1623-1632.
23. Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H. Prevalence, incidence rate and risk factors of medication-related osteonecrosis of the jaw (MRONJ) in patients with osteoporosis and cancer: a nationwide population-based study in Japan. *Journal of Oral and Maxillofacial Surgery*. 2022;80(4):714-727.
24. Ishimaru T, Fujiogi M, Michihata N, Konishi T, Morita K, Matsui H, Uda K, Fushimi K, Kawashima H, Fujishiro J, Yasunaga H. Perioperative Outcomes of Laparoscopically

- Assisted Anorectoplasty versus Conventional Procedures for Anorectal Malformation: A Retrospective Nationwide Database Study. *Pediatric Surgery International*. 2022;38(12):1785-1791.
25. Itoh H, Kaneko H, Okada A, Suzuki Y, Fujii K, Matsuoka S, Michihata N, Jo T, Nakanishi K, Takeda N, Morita H, Node K, Di Tullio MR, Homma S, Yasunaga H, Komuro I. Age-Specific Relation of Cardiovascular Health Metrics With Incident Cardiovascular Disease. *American Journal of Cardiology*. 2022;177:34-39.
 26. Iwasaki Y, Ohbe H, Shigemi D, Fushimi K, Yasunaga H. Effect of antithrombin III among patients with disseminated intravascular coagulation in obstetrics: a nationwide observational study in Japan. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2022;129(5):805-811.
 27. Kamijo K, Nakajima M, Shigemi D, Kaszynski RH, Ohbe H, Goto T, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H. Resuscitative endovascular balloon occlusion of the aorta for life-threatening postpartum hemorrhage: A nationwide observational study in Japan. *The Journal of Trauma and Acute Care Surgery*. 2022;93(3):418-423.
 28. Kamon T, Kaneko H, Itoh H, Okada A, Matsuoka S, Kiriyama H, Fujii K, Morita K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I. Sex difference in the association between lipid profile and incident cardiovascular disease among young adults. *Journal of Atherosclerosis and Thrombosis*. 2022;29(10):1475-1486.
 29. Kaneko H, Suzuki Y, Ueno K, Okada A, Fujii K, Matsuoka S, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Node K, Yasunaga H, Komuro I. Association of Life's Simple 7 with incident cardiovascular disease in 53,974 patients with cancer. *European Journal of Preventive Cardiology*. 2022;29(18):2324-2332.
 30. Kaneko H, Yano Y, Lee HH, Lee H, Okada A, Itoh H, Morita K, Fukui A, Fujii K, Suzuki Y, Matsuoka S, Nakamura S, Michihata N, Jo T, Takeda N, Morita H, Yokoo T, Nishiyama A, Node K, Viera AJ, Muntner P, Oparil S, Kim HC, Yasunaga H, Komuro I. Medication-Naïve Blood Pressure and Incident Cancers: Analysis of Two Nationwide Population-Based Databases. *American Journal of Hypertension*. 2022;35(8):731-739.
 31. Kaneko H, Yano Y, Suzuki Y, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Node K, Viera AJ, Lima JAC, Oparil S, Lam CSP, Carey RM, Yasunaga H, Komuro I. Reduction in blood pressure for elevated blood pressure/ stage 1 hypertension according to the ACC/AHA guideline and cardiovascular outcomes. *European Journal of Preventive Cardiology*. 2022;29(14):1921-1929.
 32. Karakawa R, Konishi T, Yoshimatsu H, Fuse Y, Hashimoto Y, Matsui H, Fushimi K, Yano T, Yasunaga H. Comparison of short-term outcomes between pedicled- and free-flap autologous breast reconstruction. *Breast Cancer*. 2022;29(6):1067-1075.
 33. Kodama S, Jo T, Yasunaga H, Michihata N, Matsui H, Kumazawa R, Shirota Y, Fushimi K, Toda T, Hamada M. Outcomes of gastrointestinal cancer surgeries in Parkinson's disease patients: a nationwide study. *Parkinsonism and Related Disorders*. 2022;96:45-49.
 34. Konishi T, Fujiogi M, Michihata N, Kumazawa R, Ohbe H, Matsui H, Fushimi K, Ogita M, Tanabe M, Seto Y, Yasunaga H. Interstitial lung disorders following postoperative radiotherapy with concurrent or sequential hormonal therapy for breast cancer: a nationwide database study in Japan. *Breast Cancer*. 2022;29(4):688-697.
 35. Konishi T, Fujiogi M, Michihata N, Matsui H, Tanabe M, Seto Y, Yasunaga H. Association between body mass index and incidence of breast cancer in premenopausal women: A Japanese nationwide database study. *Breast Cancer Research and Treatment*. 2022;194(2):315-325.
 36. Konishi T, Fujiogi M, Michihata N, Morita K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Association of operative day of the week with length of stay and total hospitalization costs in patients with partial mastectomy: a nationwide database study in Japan. *JMA Journal*. 2022;5(3):319-327.
 37. Konishi T, Fujiogi M, Michihata N, Ohbe H, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Cost-effectiveness analysis of trastuzumab monotherapy versus adjuvant chemotherapy plus trastuzumab in elderly patients with HER2-positive early breast cancer. *Japanese Journal of Clinical Oncology*. 2022;52(10):1115-1123.
 38. Konishi T, Fujiogi M, Sato A, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Short-term outcomes following breast cancer surgery with and without neoadjuvant chemotherapy: a nationwide administrative database study in Japan. *Annals of Surgical Oncology*. 2022;29(5):3000-3010.
 39. Konishi T, Fujiogi M, Shigemi D, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Risk factors for postoperative bleeding following breast cancer surgery: a nationwide database study of 477,108 cases in Japan. *World Journal of Surgery*. 2022;46(12):3062-3071.
 40. Konishi T, Fujiogi M, Shigemi D, Nishioka K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Surgical and obstetric outcomes of breast cancer surgery during pregnancy: A nationwide database study in Japan. *Breast Cancer Research and Treatment*. 2022;195(3):289-299.
 41. Konishi T, Goto T, Fujiogi M, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using a Japanese nationwide inpatient database. *Surgery*. 2022;171(4):1036-1042.

42. Konishi T, Sakata A, Inokuchi H, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Treatments and outcomes of adult parapharyngeal and retropharyngeal abscess: 1,882 cases from a Japanese nationwide database. *American Journal of Otolaryngology*. 2022;44(2):103770.
43. Konishi T, Takamoto T, Fujiogi M, Hashimoto Y, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H. Laparoscopic versus open distal pancreatectomy with or without splenectomy: A propensity score analysis in Japan. *International Journal of Surgery*. 2022;104:106765.
44. Konishi T, Yoshimoto T, Fujiogi M, Yamana H, Tanabe M, Seto Y, Yasunaga H. Validity of operative information in Japanese administrative data: a chart review-based analysis of 1221 cases at a single institution. *Surgery Today*. 2022;52(10):1484-1490.
45. Kumazawa R, Jo T, Matsui H, Fushimi K, Yasunaga H. Direct oral anticoagulants versus warfarin for secondary prevention of cerebral infarction and bleeding in older adults with atrial fibrillation. *Journal of the American Geriatrics Society*. 2022;70(7):2029-2039.
46. Kurakawa KI, Okada A, Bessho K, Jo T, Ono S, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yamaguchi S, Yamauchi T, Nangaku M, Kadowaki T, Yasunaga H. Major complications after percutaneous biopsy of native or transplanted liver in pediatric patients: A nationwide inpatient database study in Japan. *BMC Gastroenterology*. 2022;22(1):395.
47. Kutsuna S, Ohbe H, Matsui H, Yasunaga H. Effectiveness of fluoroquinolone antimicrobials in addition to tetracyclines for Japanese Spotted Fever: A retrospective analysis using a national inpatient database. *International Journal of Infectious Diseases*. 2022;123:70-75.
48. Kutsuna S, Ohbe H, Matsui H, Yasunaga H. Delayed Tetracycline Initiation Increases Mortality Risk in Patients With Japanese Spotted Fever: Retrospective Analysis Using a National Inpatient Database. *Open Forum Infectious Diseases*. 2022;9(11):ofac573.
49. Matsuda K, Jo T, Matsui H, Fushimi K, Yasunaga H, Sugimoto K. Institutional factors associated with early mortality of newly diagnosed acute promyelocytic leukemia. *Blood Cancer Journal*. 2022;12(12):167.
50. Matsuda K, Jo T, Shimura A, Honda A, Taoka K, Masamoto Y, Matsui H, Fushimi K, Yasunaga H, Kurokawa M. Risk of febrile neutropenia in very elderly patients aged ≥ 80 years receiving their first cycle of R-CHOP regimen: a nationwide real-world study in Japan. *British Journal of Haematology*. 2022;197(3):e37-e41.
51. Matsui H, Yamana H, Fushimi K, Yasunaga H. Development of deep learning models for predicting in-hospital mortality using an administrative claims database. *JMIR Medical Informatics*. 2022;10(2):e27936.
52. Matsuoka S, Kaneko H, Kamon T, Suzuki Y, Yano Y, Okada A, Itoh H, Morita K, Fukui A, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Yokoo T, Nishiyama A, Node K, Yasunaga H, Komuro I. Incorporation of Retinal Arteriosclerosis into Risk Stratification of Blood Pressure Category According to the 2017 ACC/AHA Blood Pressure Guideline. *Journal of Atherosclerosis and Thrombosis*. 2022;29(10):1487-1498.
53. Matsuoka S, Kaneko H, Okada A, Fukui A, Yano Y, Itoh H, Morita K, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Nakamura S, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Association between Proteinuria and Incident Colorectal Cancer: Analysis of a Nationwide Population-Based Database. *BMJ Open*. 2022;12(4):e056250.
54. Matsuoka S, Kaneko H, Okada A, Itoh H, Suzuki Y, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I. Association of retinal atherosclerosis assessed using Keith-Wagener-Barker system with incident heart failure and other atherosclerotic cardiovascular disease: Analysis of 319,501 individuals from the general population. *Atherosclerosis*. 2022;348:68-74.
55. Matsuoka S, Kaneko H, Okada A, Morita K, Itoh H, Michihata N, Jo T, Takeda N, Morita H, Fujiu K, Nakamura S, Node K, Yasunaga H, Komuro I. Age Modified Relationship between Modifiable Risk Factors and the Risk of Atrial Fibrillation. *Circulation: Arrhythmia and Electrophysiology*. 2022;15(1):e010409.
56. Michihata N, Shigemi D, Yamana H, Matsui H, Jo T, Yasunaga H. Safety of Japanese Herbal Kampo Medicines for the Treatment of Depression During Pregnancy. *International Journal of Gynecology and Obstetrics*. 2022 159(3):865-869.
57. Michihata N, Suzuki T, Honda A, Oka I, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Clinical features of Kawasaki disease with atlantoaxial rotatory fixation. *Pediatric Infectious Disease Journal*. 2022;41(8):626-630.
58. Michihata N, Suzuki T, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Association between intravenous immunoglobulin dose and outcomes in patients with acute Kawasaki disease. *European Journal of Pediatrics*. 2022;181(10):3607-3615.
59. Miyachi H, Konishi T, Kumazawa R, Matsui H, Shimizu S, Fushimi K, Matsui H, Yasunaga H. Treatments and outcomes of generalized pustular psoriasis: a cohort of 1516 patients in a nationwide inpatient database in Japan. *Journal of the American Academy of Dermatology*. 2022;86(6):1266-1274.
60. Miyachi H, Konishi T, Shigemi D, Matsui H, Shimizu S, Fushimi K, Matsue H, Yasunaga H. Association between psoriasis and short-term outcomes of acute myocardial infarction: a matched-pair cohort study using a nationwide inpatient database in Japan. *JAAD International*. 2022;8:21-30.

61. Miyakawa T, Michihata N, Morita K, Matsui H, Honda M, Yasunaga H. Ambient temperature and hospital admissions for acute cholecystitis: A nationwide inpatient database study in Japan. *HPB*. 2022;24(3):398-403.
62. Miyamoto Y, Iwagami M, Aso S, Uda K, Fushimi K, Hamasaki Y, Nangaku M, Yasunaga H, Doi K. Postoperative outcomes of cancer surgery in patients with and without kidney failure with dialysis therapy: A matched-pair cohort study. *Clinical Kidney Journal* 2022;15(6):1137-1143.
63. Miyamoto Y, Matsuyama T, Goto T, Ohbe H, Kitamura T, Yasunaga H, Ohta B. Association between age and neurological outcomes in out-of-hospital cardiac arrest patients resuscitated with extracorporeal cardiopulmonary resuscitation: a nationwide multicentre observational study. *European Heart Journal: Acute Cardiovascular Care*. 2022;11(1):35-42.
64. Mizuno S, Ono S, Takano A, Yasunaga H, Iwase H. Dental characteristics associated with methamphetamine use: Analysis using forensic autopsy data. *BMC Oral Health*. 2022;22(1):141.
65. Nagata S, Ohbe H, Jo T, Matsui H, Fushimi K, Yasunaga H. Glucocorticoids and rates of biphasic reactions in patients with adrenaline-treated anaphylaxis: A propensity score matching analysis. *International Archives of Allergy and Immunology*. 2022;183(9):939-945.
66. Nishimoto Y, Ohbe H, Matsui H, Nakajima M, Sasabuchi Y, Goto T, Morita K, Fushimi K, Sato Y, Yasunaga H. Predictive ability of the sequential organ failure assessment score for in-hospital mortality in patients with cardiac critical illness: a nationwide observational study. *EJH Acute Cardiovascular Care*. 2022;11(4):312-321.
67. Ogasawara H, Obe H, Yasunaga H. Oral food challenge management in Japan: A retrospective analysis of health insurance claims data. *Clinical & Experimental Allergy*. 2022;52(7):898-900.
68. Ogawa Y, Yamana H, Noda T, Kishimoto M, Yoshihara S, Kanaoka K, Matsui H, Fushimi K, Yasunaga H, Kawaguchi M, Imamura T. Severe complications after general anesthesia versus sedation during pediatric diagnostic cardiac catheterization for ventricular septal defect. *Journal of Clinical Medicine*. 2022;11(17):5165.
69. Ohbe H, Goto T, Matsui H, Fushimi K, Yasunaga H. Associations of government-issued ICU admission criteria with clinical practices, outcomes, and ICU bed occupancy. *Annals of the American Thoracic Society*. 2022;19(6):1013-1021.
70. Ohbe H, Goto T, Nakamura K, Matsui H, Yasunaga H. Development and validation of early prediction models for new-onset functional impairment at hospital discharge of ICU admission. *Intensive Care Medicine*. 2022;48(6):679-689.
71. Ohbe H, Matsui H, Kumazawa R, Yasunaga H. Intensive care unit versus high dependency care unit admission after emergency surgery: a nationwide in-patient registry study. *British Journal of Anaesthesia*. 2022;129(4):527-535.
72. Ohbe H, Matsui H, Kumazawa R, Yasunaga H. Postoperative intensive care unit admission following major elective surgery: a nationwide inpatient database study. *European Journal of Anaesthesiology*. 2022;39(5):436-444.
73. Ohbe H, Matsui H, Yasunaga H. ICU versus high-dependency care unit for patients with acute myocardial infarction: a nationwide propensity score-matched cohort study. *Critical Care Medicine*. 2022;50(6):977-985.
74. Ohbe H, Sasabuchi Y, Kumazawa R, Matsui H, Yasunaga H. Intensive care unit occupancy in Japan, 2015–2018: a nationwide inpatient database study. *Journal of Epidemiology*. 2022;32(12):535-542.
75. Ohbe H, Sasabuchi Y, Matsui H, Yasunaga H. Impact of the COVID-19 pandemic on critical care utilization in Japan: a nationwide inpatient database study. *Journal of Intensive Care*. 2022;10(1):51.
76. Ohbe H, Tagami T, Ogura T, Matsui H, Yasunaga H. Low-flow duration and outcomes of extracorporeal cardiopulmonary resuscitation in adults with in-hospital cardiac arrest: a nationwide inpatient database study. *Critical Care Medicine*. 2022;50(12):1768-1777.
77. Ohbe H, Tagami T, Uda K, Matsui H, Yasunaga H. Incidence and outcomes of in-hospital cardiac arrest in Japan 2011–2017: a nationwide inpatient database study. *Journal of Intensive Care*. 2022;10(1):10.
78. Ohisa K, Yamana H, Morita K, Matsui H, Fushimi K, Yasunaga H. Association between subcutaneous steroid injection for extravasation of vesicant anticancer drugs and skin ulcers requiring surgery. *European Journal of Oncology Nursing*. 2022;58:102119.
79. Okada A, Hashimoto Y, Goto T, Yamaguchi S, Ono S, Ikeda Kurakawa K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T. A machine learning-based predictive model to identify patients who failed to attend a follow-up visit for diabetes care after recommendations from a national screening program. *Diabetes Care*. 2022;45(6):1346-1354.
80. Okada A, Yamana H, Yamaguchi S, Kurakawa K, Michihata N, Matsui H, Fushimi K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T. Effect of branched-chain amino acid infusion on in-hospital mortality of patients with hepatic encephalopathy and endstage kidney disease: a retrospective cohort study using a national inpatient database. *Journal of Renal Nutrition*. 2022;32(4):432-440.
81. Okada A, Yasunaga H. Prevalence of non-communicable diseases in Japan using a newly developed administrative claims database covering young, middle-aged, and elderly people. *JMA Journal*. 2022;5(2):190-198.
82. Okubo Y, Nishi A, Uda K, Miyairi I, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yasunaga H. Financial Incentives for Infection Prevention and Antimicrobial Stewardship to Reduce Antibiotic Use: A Japan's Nationwide Observational Study. *Journal of Hospital Infection*. 2022;131:89-98.

83. Ono S, Ishimaru M, Yokota I, Konishi T, Okada A, Ono Y, Matsui H, Itai S, Yonenaga K, Tonosaki K, Watanabe R, Hoshi K, Yasunaga H. Risk of post-extraction bleeding with direct oral anticoagulant compared with warfarin: Retrospective cohort study using large scale claims data in Japan. *Thrombosis Research*. 2022;222:24-30.
84. Sakamoto T, Fujiogi M, Ishimaru M, Matsui H, Fushimi K, Yasunaga H. Comparison of postoperative infection after emergency inguinal hernia surgery with enterectomy between mesh repair and non-mesh repair: a national database analysis. *Hernia*. 2022;26(1):217-223.
85. Sato M, Yamana H, Ono S, Ishimaru M, Matsui H, Yasunaga H. Trends in prophylactic antibiotic use for tooth extraction from 2015 to 2018 in Japan: an analysis using a health insurance claims database. *Journal of Infection and Chemotherapy*. 2022;28(4):504-509.
86. Shibahashi K, Konishi T, Ohbe H, Yasunaga H. Cost-effectiveness analysis of termination-of-resuscitation rules for patients with out-of-hospital cardiac arrest. *Resuscitation*. 2022;180:45-51.
87. Shibahashi K, Ohbe H, Yasunaga H. Association between intensive care unit admission practices and outcomes in patients with isolated traumatic subarachnoid hemorrhage: A nationwide inpatient database analysis in Japan. *Neurocritical Care*. 2022;37(2):497-505.
88. Shibahashi K, Ohbe H, Yasunaga H. Urban-rural inequalities in care and outcomes of severe traumatic brain injury: A nationwide inpatient database analysis in Japan. *World Neurosurgery*. 2022;163:e628-e634.
89. Shibahashi K, Ohbe H, Yasunaga H. Adjuvant oral tranexamic acid and reoperation after burr-hole surgery in patients with chronic subdural hematoma: Propensity score-matched analysis using a nationwide inpatient database. *Journal of Neurosurgery*. 2022;138(2):430-436.
90. Shigemi D, Tabuchi T, Okawa S, Yasunaga H. Association between health literacy and COVID-19 prevention behaviors among pregnant and postpartum women. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2022;35(25):9971-9977.
91. Sugawara C, Yamana H, Sasagawa E, Yonezawa K, Hikita N, Morita K, Matsui H, Yasunaga H, Haruna M. Factors associated with surgical treatment in postpartum women with mastitis or breast abscess: A retrospective cohort study. *Breastfeeding Medicine*. 2022;17(3):233-238.
92. Sugihara T, Kamei J, Yasunaga H, Sasabuchi Y, Fujimura T. Prescription of Choreito, a Japanese Kampo medicine, with antimicrobials for treatment of acute cystitis: a retrospective cohort study. *Antibiotics*. 2022;11:1840.
93. Suzuki J, Sasabuchi Y, Hatakeyama S, Matsui H, Sasahara T, Morisawa Y, Yamada T, Fushimi K, Yasunaga H. Additional effect of azithromycin over β -lactam alone for severe community-acquired pneumonia-associated acute respiratory distress syndrome: a retrospective cohort study. *Pneumonia*. 2022 ;14(1):1.
94. Suzuki T, Michihata N, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H. Low- versus high-concentration intravenous immunoglobulin for children with Kawasaki disease in the acute phase. *International Journal of Rheumatic Diseases*. 2022;25(5):576-583.
95. Suzuki Y, Kaneko H, Okada A, Itoh H, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Impact of Glucose Tolerance on Incident Proteinuria in Prediabetes: Analysis of a Nationwide Population-Based Dataset. *American Journal of Nephrology*. 2022;53(4):307-315.
96. Suzuki Y, Kaneko H, Okada A, Itoh H, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Yano Y, Kamiya K, Matsunaga A, Ako J, Fukui A, Yokoo T, Nishiyama A, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Risk for Proteinuria in Newly Defined Hypertensive People Based on the 2017 American College of Cardiology/American Heart Association Blood Pressure Guideline. *American Journal of Cardiology*. 2022;168:83-89.
97. Suzuki Y, Kaneko H, Okada A, Itoh H, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I. Comparison of cardiovascular outcomes between SGLT2 inhibitors in diabetes mellitus. *Cardiovascular Diabetology*. 2022;21(1):67.
98. Suzuki Y, Kaneko H, Okada A, Itoh H, Morita K, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Fukui A, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I. Change in Cardiovascular Health Metrics and Risk for Proteinuria Development: Analysis of a Nationwide Population-Based Database. *American Journal of Nephrology*. 2022;53(2-3):240-248.
99. Suzuki Y, Kaneko H, Okada A, Matsuoka S, Itoh H, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Node K, Yamauchi T, Yasunaga H, Komuro I. Prediabetes in Young Adults and Its Association with Cardiovascular Health Metrics in the Progression to Diabetes. *Journal of Clinical Endocrinology & Metabolism*. 2022;107(7):1843-1853.
100. Suzuki Y, Kaneko H, Okada A, Nangaku M, Yasunaga H, Komuro I. Kidney outcomes in patients with diabetes mellitus did not differ between individual sodium-glucose cotransporter-2 inhibitors. *Kidney International*. 2022;102(5):1147-1153.
101. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I. Glycemic status and the association of change in blood pressure with incident cardiovascular disease. *American Heart Journal*. 2022;254:48-56.
102. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H,

- Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Matsunaga A, Node K, McEvoy JW, Lam CSP, Oparil S, Yasunaga H, Komuro I. Association of Cardiovascular Health Metrics with Risk of Transition to Hypertension in Non-Hypertensive Young Adults. *American Journal of Hypertension*. 2022;35(10):858-866.
103. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujiu K, Yamaguchi S, Michihata N, Jo T, Takeda N, Morita H, Node K, Kim HC, Viera AJ, Oparil S, Yasunaga H, Komuro I. Age-Dependent Relationship of Hypertension Subtypes with Incident Heart Failure. *JAHA*. 2022;11(9):e025406.
104. Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Ueno K, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yokota I, Node K, Yasunaga H, Komuro I. Dose-dependent relationship of blood pressure and glycemic status with risk of aortic dissection and aneurysm. *European Journal of Preventive Cardiology*. 2022;29(18):2338-2346.
105. Tagami T, Yamakawa K, Endo A, Hayakawa M, Ogura T, Hirayama A, Yasunaga H. Japanese Multicenter Research of COVID-19 by Assembling Real-world Data: A Study Protocol. *Annals of Clinical Epidemiology*. 2022;4(3): 92-100.
106. Taguchi R, Shigemi D, Yasunaga H. Trends in antihypertensive prescription for pregnant women with hypertension and their peripartum outcomes before and after label and guideline revisions in Japan. *Hypertension Research*. 2022;45(11):1823-1831.
107. Tahara S, Hattori Y, Aso S, Uda K, Kumazawa R, Matsui H, Fushimi K, Yasunaga H, Morita A. Trends in Surgical Procedures for Spontaneous Intracerebral Hemorrhage in Japan. *Journal of Stroke and Cerebrovascular Diseases*. 2022;31(9):106664.
108. Ueno K, Kamiya K, Kaneko H, Okada A, Itoh H, Fujiu K, Takeda N, Morita H, Michihata N, Jo T, Yasunaga H, Komuro I. Acute-phase initiation of cardiac rehabilitation for improvement in activities of daily living in patients hospitalized for acute heart failure. *Journal of Cardiovascular Development and Disease*. 2022;9(4):97.
109. Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Node K, Yamauchi T, Yasunaga H, Komuro I. Relationship of normal-weight central obesity with the risk for heart failure and atrial fibrillation: Analysis of a nationwide health check-up and claims database. *European Heart Journal Open*. 2022;2(3):oeac026.
110. Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Yasunaga H, Komuro I. Clinical utility of simple subjective gait speed for the risk stratification of heart failure in a primary prevention setting. *Scientific Reports*. 2022;12(1):11641.
111. Urushiyama H, Jo T, Hasegawa W, Yokoyama A, Ando T, Sakamoto Y, Kumazawa R, Uda K, Michihata N, Awano N, Hiroki M, Fushimi K, Yasunaga H, Nagase T. Effect of nintedanib on acute exacerbations of fibrosing interstitial lung diseases: A national database study in Japan. *ERJ Open Research*. 2022;8(4):00209-2022.
112. Wada Y, Takahashi H, Matsui H, Yasunaga H, Fujiwara H, Sasabuchi Y. Adhesion barriers and intraperitoneal or uterine infections after cesarean section: A retrospective cohort study. *Surgery*. 2022;172(6):1722-1727.
113. Yamana H, Sasabuchi Y, Matsui H, Michihata N, Jo T, Yasunaga H. Maoto plus neuraminidase inhibitor versus neuraminidase inhibitor alone for reducing hospitalisation in older adults with seasonal influenza. *Annals of Clinical Epidemiology*. 2022;4:120-128
114. Yamana H, Tsuchiya A, Horiguchi H, Morita S, Kuroki T, Nakai K, Nishimura H, Jo T, Fushimi K, Yasunaga H. Validity of a model using routinely collected data for identifying infections following gastric, colorectal, and liver cancer surgeries. *Pharmacoepidemiology and Drug Safety*. 2022;31(4):452-460.
115. Yamazaki R, Ohbe H, Matsuda Y, Kito S, Shigeta M, Morita K, Matsui H, Fushimi K, Yasunaga H. Early electroconvulsive therapy in patients with bipolar depression: a propensity score-matched analysis using a nationwide inpatient database. *Journal of Affective Disorders*. 2022;312:245-251.
116. Zhang L, Michihata N, Matsui H, Fushimi K, Yasunaga H, Tsuda Y, Tanaka S, Kobayashi H. Preoperative Arterial Embolization and Wound Complications After Resection of Malignant Bone Tumor in the Pelvis: A Nationwide Database Study. *Japanese Journal of Clinical Oncology*. 2022;52(10):1176-1182.

H. 知的財産権の出願・登録状況

なし

図1. Q1.本講義の難易度

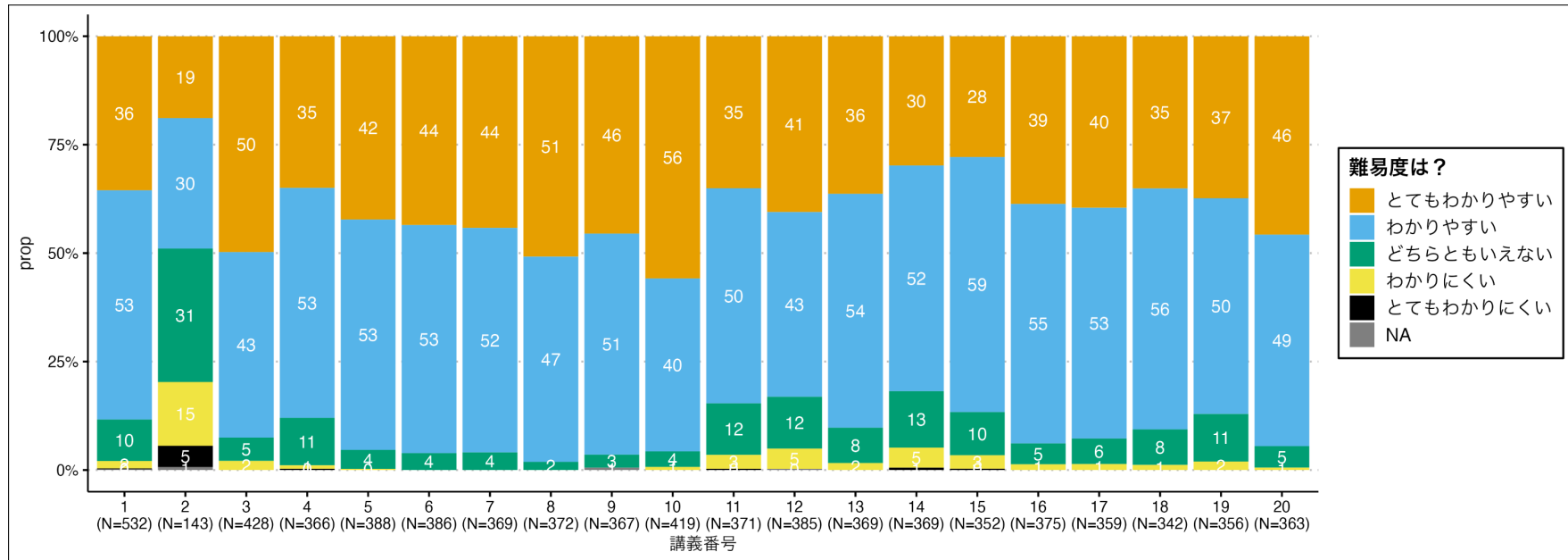
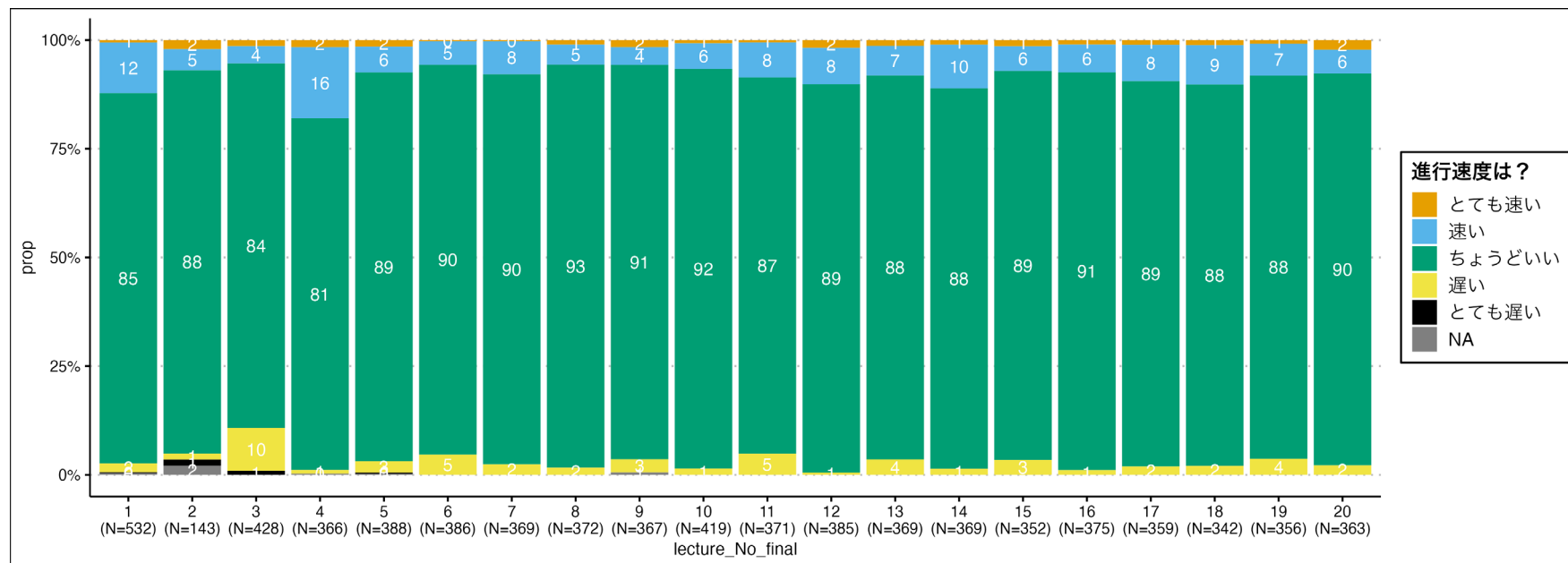


図 2. Q2. 本講義の進行速度



Q3. 本講義の満足度

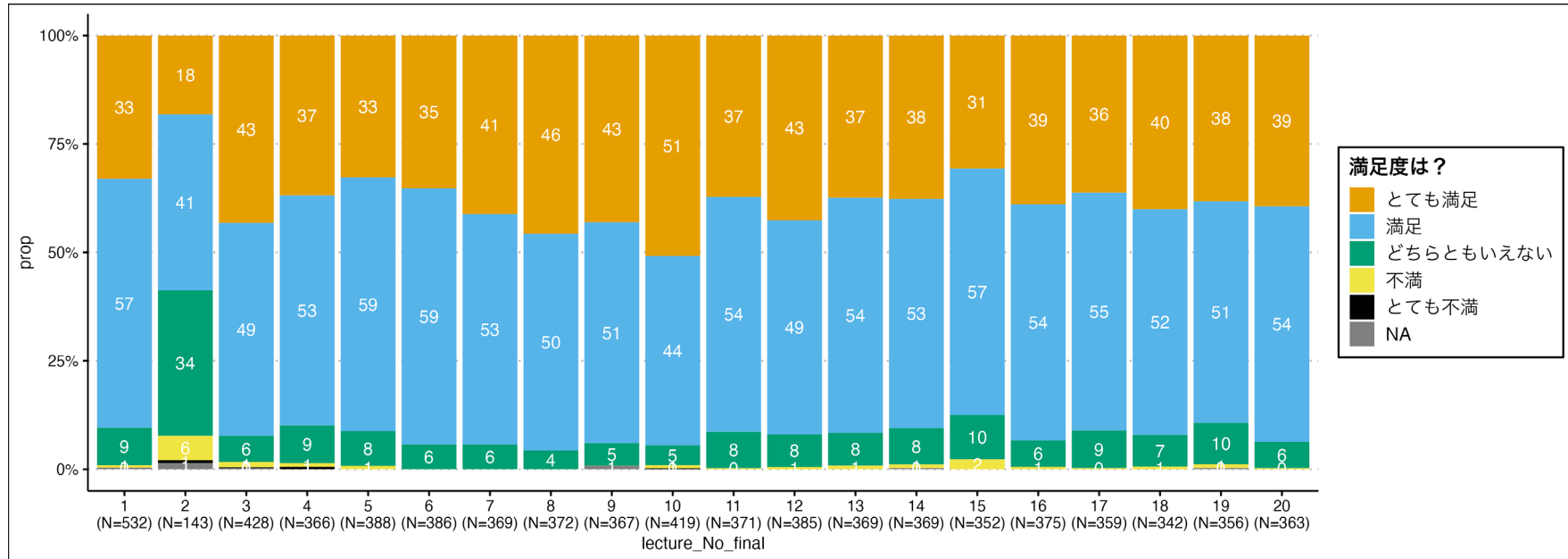


表 1. Q4.本講義で学べて良かったこと（自由記載）（抜粋）

講義番号	Q4
1	(長大な内容である)NDB ガイドラインの内容について、わかり易く説明していただき、理解できた。
1	・ NDB, DPC データ, 介護 DB の特徴が分かりやすく整理されていた ・ 格納されているデータの連結方法が分かりやすかった。
1	・ 各データベースの特徴及び共通の特徴 ・ 解析をする際の物理的、技術的安全管理措置の厳密さ ・ 不適切利用の内容及び責任の重大さ ・ 将来の各 DB の動向
1	DB について整理され、それぞれの特徴を分かりやすくお伝えいただいた。また、各手続きなどにも触れられていてよかった。
1	DB にどのような情報がどのように格納・連結されるのかを学ぶことができたため良かった。
1	DB のそれぞれの特徴が簡潔に述べられており、分かりやすかった。
1	DB の概観を理解できた
1	DB の概要を網羅的に知ることができた。
1	DB の今後の方向性について
1	DB の種類、いつから収集されているのか、どのような流れで情報が格納されているのかなど基本的なことを分かりやすく説明されており、勉強になった。
1	DB の特徴と利用の概要と課題の一部を理解できました。
1	DB の連結についての概要がわかりました。
1	DB 毎の違いがわかりやすかったです
1	DB 利用の流れや今後の展望がわかりやすかったです
1	DB 利用方法について知ることができてよかった。
1	DPC が取得される仕組みは知らなかったので、勉強になりました。
1	DPC データについての概要が理解できた。
1	DPC データの第三者提供について知ることができた。経時的な個人追跡が可能というのがこれまでの DPC データの弱点をうまくカバーしている印象です。また介護保険のデータで社会経済情報が得られるようになったのも注目すべきことだと感じました。
1	DPC データの動向が知れた
1	DPC データベースに関する話は初めてお聞きしました。勉強になりました。
1	DPC データベース以外のデータベースの概要について学ぶことができた
1	DPC データを利用した研究
1	Health Interagency Cloud が開発され Big data が利用できることは素晴らしい。
1	ID5 が追加されたこと。
1	ID5 など DB の改善により、DB は今後さらに使用しやすいものとなるのは必須で、DB の構造を理解するためにもどのように整備されて行って完成したものなのか知れてよかった
1	ID5 について知らなかったので、ためになりました。
1	ID5 の導入は 2022 年 4 月からだが、2020 年 10 月まで遡っていることを知りませんでした。
1	ID のリンク方法
1	ID の種類がよくわかりました。
1	ID の連結について理解できた。

1	ID 等についての具体的な構造
1	LIFE を初めて知ったのでよかった
1	LIFE 情報とは何なのか。
1	NBD,DPC の詳細な説明や規定などが知れてよかったです。
1	NBD、DPC、介護 DB の違い
1	NBD のしくみ、基本的な項目や利用可能な事項について
1	NDB, DPC データベースに関する使用方法
1	NDB, DPC データベースの仕組み
1	NDB, レセプト以外の DB について詳細が学べたのでよかった。
1	NDB,DPC にはどのような情報が掲載されるのか、特徴などを説明されていて、具体的なイメージがわいた。
1	NDB、DPC DB 等の格納データ、格納されるまでの流れについて学べたことです。
1	NDB、DPC、介護 DB などの特徴や使用上の注意点を知らなくて良かった。
1	NDB、DPC がどのような内容を根拠としたデータベースかについて理解できた
1	NDB、DPC だけでなく、様々な公的 DB の内容を網羅的に勉強できたこと。
1	NDB、DPC などの基本が学べてよかった。
1	NDB、DPC などの公的データベースに格納されているデータ内容について、概要を理解できました。
1	NDB、DPC などの公的データベースに関する最近、今後の動向について学ぶことができました。
1	NDB、DPC などの公的データベースに関する最近の動向や今後の DB の発展を知ることができてよかったです。
1	NDB/DPC などのデータベースの実情が理解できた。
1	NDB/DPC 等、どの様にして連結させるべきかなど扱い方のイメージができました。
1	NDB・DPC・介護 DB の共通の特徴や差異について体系的に学ぶことができ理解を整理できました。
1	NDB・DPC・介護 DB 利用の手続きについて
1	NDB・DPC データの構造と動向がよく理解できました。 また、データのセキュリティ対策が徹底されていることも知ることができました。
1	NDB・DPC データベースのデータ、データ登録までのフロー、注意点について理解の整理をすることができた。
1	NDB・DPC データ利用における手続き、違い、今後の展望について理解することができました。
1	NDB・DPC などの公的データベースの基本と最新の動向について知ることができてよかった。
1	NDB および介護 DB のデータ収集の流れは大変分かり易いものでした。また、今後の公的データベースの活用ならびに利用環境のクラウド化について大変参考になりました。ありがとうございました。
1	NDB が第三者提供されることがわかった。
1	NDB だけでなく、他の公的データベースについての情報と、今後の方向性がしれたのでよかったです。
1	NDB データベースの実践的な利用のみでなく基礎的な事象を学べた。

1	NDB データ使用の申請の仕方がわかった
1	NDB と DBC だけでなく難病データベースとの比較もわかりやすかった
1	NDB などの公的データの使用に関して、現状と今後について分かりやすく説明頂き、非常に勉強になりました。
1	NDB において所得階層情報が含まれているのは知らなかったのが勉強になりました。うまく活用して解析できないか検討してみたいと思いました。
1	NDB において被保険者番号の履歴の古いものを匿名化して追跡可能としていることなど
1	NDB についての見識が深まった。 他の公的 DB については存在自体を知らなかったものもあり、関心を持つきっかけとなった。
1	NDB に格納されているデータについての解説がとても分かりやすかった。
1	NDB に含まれるコンテンツ、利活用時の注意点などを幅広く網羅されていて、勉強になりました。
1	NDB の各 ID 体系がどういうものがあるのかや、他の DB とどう結合できるのか、といった周辺知識が得られてよかった
1	NDB の匿名化 ID の種類や連結可能な DB について学習ができました。複数の公的データが連結されることで医療・介護分野におけるデータ活用の将来性を感じました。
1	NDB の特性などを概観することができた 介護、DPC の状況がよく分かった また、データの概要や収集・提供の流れがよく分かった。 今後、連結を予定している DB の種類がよく分かった また、それぞれを連結するための ID がよく分かった 審査の観点がよく分かった
1	NDB の民間事業者への提供が法的に可能になっていることや、DPC や介護 DB との連結の状況を知ることができて良かったです。
1	NDB は全国の医療施設のレセプトを集計したデータベース DPC は対象病院のみ、データ構造が複雑ではない。
1	NDB や DPCDB など各データベースの概要および連結申出ができるものもあること。今後の動向。
1	NDB や DPC データの歴史や特徴の概要をお聞きすることができ、大変勉強になりました。
1	NDB や DPC について初めて学んだため、それぞれがどのようなものかを理解することができ良かった。
1	NDB 他 DPC などの第三者提供のデータベースの概要をまとめていただき、その違いを学ぶことができた。また ID4 や ID5 などのデータベースの連結に用いる KEY 項目についても、用語としては理解していたが、その内訳はきちんと理解できていなかったのが、今回の講義で学ぶことができたのは良かったです
1	こういった研究に関して学ぶことのできる場が限られているため非常に勉強になりました。
1	このようなまとまった情報というのはあまり目にする機会がなかったので、現在利用可能な DB のオーバービューを知ることができて、とても参考になりました。
1	この講義で紹介されたデータを東京や京都以外のものが使用するのはかなり難しそうだ、ということ。
1	そもそも NDB というものの詳細が分かっていたいなかったため、概略がわかった。
1	そもそも公的データベースの種類について理解していなかったのが、そこからわかってよかった
1	それぞれの DB のについての個々と共通の特徴について知ることができた。また、そのデータの格納と連結、今後の動向についても学ぶことができた。
1	それぞれの DB の特徴、相違点、今後の連結予定も含めてまとめられていた点

1	それぞれの DB の特徴がよくわかった
1	データ利用申請の流れや、審査で特にみられる点等をご教授いただき大変ありがたく存じます。
1	どのようなデータベースがあるのか、大まかな概要が理解できた。 データベースに格納される情報や、それぞれの ID の定義などが何となく分かった。
1	どのようなデータベースが整備されていて、そのデータベースがいつごろから開発されてきたのかが明確になってよかった。
1	どのような公的データベースがあるのか、具体的にどのようなデータが提供できるのかの概要がわかった。
1	ブラックボックス化しやすい審査でのポイントを明示して頂けているので参考になりました。
1	まだ敷居が高そうということがわかりました。民間の DB の方しかアクセスいづらいのではと思いました
1	医療介護データの解析基盤の構築準備がなされていることを初めて知ったので、今後 NDB データの利活用の円滑化が進められていることを知って勉強になりました。
1	介護データベースオープンデータが公開されていることや LIFE の存在を初めて知りました。2 年前にも本 Webinar を受講しましたが、こういった情報は日々アップデートする必要があることを痛感しました。
1	介護レセプトと LIFE の関係が分かってよかったです。
1	各 DB の概要が理解できた
1	各 DB の特徴があいまいだったことが、明確になった。
1	各公的データベースの概要と連結、セキュリティに関する要件などについて知り、実際に研究を行う際にどのように利用するのかイメージすることができました。
1	各種 DB に関して概要が整理されており理解しやすかった
1	各種 DB の全体的な概要や複数年にまたがるデータを利用したい時の注意点について知ることができた。今後利用できる DB が増えることや、被保険者番号が変わった場合でもと違うできる仕組みが出来ていることが知れてよかった。
1	現在公的データを解析ベースに乗せるために整理しているので、提供データや追跡・連結に関するお話が聞けて業務に活かせたらと思った。
1	個票により DPC でも追跡可能になること 第 3 社提供されるデータベースの種別がしれた データの利用には NDB 同様にわりと高いハードルが必要であること
1	公的なデータベースの全般的な事項をすることができて良かった。 公的なデータベースが連結できるような動向を知ることができ、今後の進展に期待したい。
1	厚生労働省が第三者提供する公的 DB の概要が把握できたこと、今後の動向について情報が得られたこと
1	刻々と状況が変化していく中で、最新の動向を知ることが出来たこと、特に利用可能な公的データベースの状況を網羅的に把握することが出来たので大変有益でした。有り難うございました。
1	次年度より、介護/LIFE と NDB/DPC を連結した研究を考えているので、十分な理解やイメージには至っていませんが、とても参考になりました。
1	実際の NDB・DPC の取り扱いやデータの内容の概要、今後の変更点などがわかりやすかつかめた。
1	手数料、公益性の担保など、具体的なハードルについて知れたことが良かったです。また、LIFE のデータ利用についても学ぶことが出来た点がよかったです。
1	情報の流れや ID の紐付けなどがわかりやすく解説されていてとても参考になった。クラウド対応したらこういった DB をぜひ利用してみたいと思った。
1	日々チェックしているレセプトがどのように活用されているかわかった。NDB データベースが現在も改善、他のデータベースとの更なる連結を目指していることがわかった。

1	日本の公的データベースを使用するにあたっての厳しい制限や利用規定について、（文書で読み取れる内容以上に）詳しく知ることができた。
1	費用面がわかったことがよかった。
1	未来軸での展望も含め、各公的データベースのまとめを学ぶことができ有用な情報でした。
1	民間活用も含めてカバーしていただいて大変ありがたかったです。最近の動向がよくわかりました。
1	予防接種や感染症 DB などとの連結により、ワクチン接種による経済的効果など、今後の研究の可能性が広がることが予想され、今後の展望に期待を持つことができた。
1	様々な DB があるなかでそれぞれの対象や取り出し方を具体的に開けて今後の研究をすすめるうえで参考になった。
1	様々な HP、会議体資料に散在している基礎知識が集約されていて、知識の確認に大変役に立ちました。
1	利用可能なデータベースが丁寧にまとめられており、研究課題の検討に資する動画であった。今後の動向までカバーしている点が特によかったです。
1	利用可能な公的データベースについて知ることができた。探索的利用環境を使ってみたいです。
1	連結用 ID について 具体的な審査ポイントについて 今後の連結予定について
2	NDB ガイドラインを読むきっかけを作ったこと
2	1 回目の講義で行ったことがよく復習できた
2	NDB ガイドラインをこの機会に読み込むことができました。
2	NDB ガイドラインの概要を学べた
2	NDB ガイドラインの詳細、申請方法
2	NDB ガイドラインの内容を網羅していたこと 申請する際に知っておくべき内容が確認できたこと
2	NDB ガイドラインを読むのは初めてでしたが、テスト問題に回答することによって自分の理解の甘い部分を明確でき、再度学習を行うことで理解を深めることもできて大変勉強になりました。
2	NDB データを今後利用する場合に必要な知識をテストを通じて学べたこと
2	NDB データ利用申請を実際に行う際にはかなり細かい事柄まできちんと理解していないといけないことがよくわかりました。現時点では申請予定はありませんが、将来的に申請する場合にはその時点の最新のガイドラインを熟読してから手続きしなければならないと思いました。
2	NDB のガイドラインやマニュアルなどをしっかり読み込む良い機会になった。
2	NDB の申請等、把握すべき事項を確認することができた。
2	NDB の本来の目的を含め、根拠となる法や、具体的に利用する際に必要なルールを知ることができてよかった。
2	NDB を利用する上での実際がわかりました。
2	ガイドライン読むのが大変です
2	ガイドライン内容
2	ルールに関する詳細を整理できた。
2	一問一答形式で理解しやすかった。
2	回答のフィードバックがあったこと

2	具体的に申請する流れをよりよく理解することができました。
2	厚労省作成の資料が多数あるためいずれの資料がより参考資料として適当にご教示いただけますと幸いです。
2	講義を拝聴しているよりこのような形式の方が、理解が深まると感じました。
2	私は理解できていないと言うことが認識できた
2	手続きなどについて学ぶことができて良かったです。
2	詳細事項の確認となった
2	申請方法の具体的な流れの理解が高まりました。
2	成果物の公表の認識が不十分でした。(原則として、公表される研究の成果物において患者等の数が10未満になる集計単位が含まれていないことや、年齢区分が、5歳毎にグルーピングして集計されていることというガイドラインを把握できておりませんでした)
2	適度に難しかったので。理解の薬だった
2	難しかったので、調べながら勉強になった点は良かったです。
2	非常に難しかったです。知らないことが多かったのでフィードバックで勉強しました。
2	聞きなれない言葉が多く、難しかったです。回答を表示頂けた事で、勉強になりました。
2	問題を通して調べてみる機会となったことは良いと思った。
2	問題形式により、自身の理解度を図ることができた
3	<ul style="list-style-type: none"> ・ SQL の意義、重要性について ・ 基本的な SQL の使用方法について(基本構文、条件や並べ替えの書き方、集計値記述、変数作成、日付関数、データ結合)
3	DB と SQL に限らず、複数のテーブルに出力されるデータの事前処理に応用できる内容だった。
3	distinct の理解が深まりました。
3	huge データの解析に当たっての流れ、考え方が理解できた。
3	NDB データを扱う上での、特に結合時の SQL の使い方が非常にわかりやすく、参考になりました。
3	Posgresql は過去に使ったことがあり、自院の電カルデータの抽出のために、SQL を書くことがあります。データのハンドリングと確認には、Filemaker を使って概要をつかむ方が楽なので、SQL を各頻度は稀なので、忘れかけていた構文の想起ができました。
3	R しか使ったことがなく不安でしたが、丁寧な説明でとてもよくわかりました。
3	R などの統計ソフトの場合・・・といった形での比較が説明に補足されていたのでとても SQL についてわかりやすかったです。
3	R を使っておりましたが、SQL を使うことにより処理が早くすすむのではないかと思います。食わず嫌いでしたが、利用を検討したいと思います
3	SAS の sql procedure しか使ったことがないのですが概ね同じであることが理解できました。
3	SELECT の活用方法が理解できた。
3	SELECT 文を用いることで、要約統計量を算出することができる。
3	SQL の操作方法については知識があったが、実際の流れで SQL コードを学べるのは良かった。次の講義も楽しみ。
3	SQL を学んだのが10年以上前だったので基礎を思い出せた。
3	SQL、非常に難しかったです。講師の先生のお話が随所にわかりやすく工夫して下さったと思います。統計ソフト、使えないことは初めて知りました。

3	SQLserver の文法は知らなかったので勉強になりました。
3	SQL でできること具体例をしめしてくださり、わかりやすかったです。
3	SQL でどんなことができるのか、基本的なことがわかりました。
3	SQL でのデータ結合の概略について理解できた
3	SQL での具体例を知ることができよかったです。
3	SQL で用いる構文例
3	SQL というものがどういうものであるのかが、全くわからない状態でした。講義後は何となくどういうものであるのかが、ある程度はイメージできるようになりました。機会があれば、SQL に触れてみたいと考えています。
3	SQL とはなにか？ とういときに役立つか？ リレーショナルデータベースとは何か？ とわかりやすく説明いただいた。
3	SQL について、基礎的な内容から学ぶことができた
3	SQL について、やってみたいと思えた。
3	SQL について、初心者にもわかりやすく、どういうものか学べた。具体的な方法について学べて良かった。
3	SQL については全くの初めてですが全体的なイメージがもてました。資料が手元にないとわからないので準備した上で再度視聴いたします。
3	SQL についてほぼ知識がない状態であったが、触りの部分だけでも理解できてよかった。
3	SQL についてまったく何も知らなかったので、基礎を大変わかりやすく説明頂けてよかった
3	SQL について全く理解していなかったので、大変勉強になりました。クエリの書き方まで丁寧に教えてくださり分かりやすかったです、実際にやってみようと思います。
3	SQL について知識が全くなかったのですが、非常にわかりやすかったです。
3	SQL について非常に分かりやすく、実践的な物を講義頂き、勉強になりました。
3	SQL について復習できた
3	SQL のメリットや既存の R などで処理するデータとの違いなどがわかっていなかったが、理解がすすんだ
3	SQL の基礎について具体的事例に即して解説されており、素晴らしいと感じました。
3	SQL の基本がよくわかりました。加えて、統計ソフトではどのように言語化されるか、研究に使う時にはどのような点に困るかなどを教えていただいたので、すぐに使えそうと感じました。
3	SQL の基本について、ビッグデータで研究する前提での知識が身につけられてよかった。
3	SQL の基本に加えて、DeSC などの構造についてもおおまかに知ることができて良かった。
3	SQL の基本を分かりやすく丁寧にご教示いただき、理解できていなかった点も含めて復習になりました。
3	SQL の基本的な使い方を具体例を活用して教えていただいたので大変わかりやすかったです。
3	SQL の基本的な部分を知ることができました。SQL を利用することの利点もわかりました。
3	SQL の重要性について理解できました。非エンジニアのため具体的なクエリの説明は難しく感じました。
3	SQL の必要性について理解できた。結合されるキーを意識してクエリを作るの部分が知りたいことだったのでよかった。
3	SQL の必要性や基本構文を学べて良かった。メリットが分かったのでモチベーションにもなる。

3	SQL の文法はある程度理解していたが、認識していなかった方法も出てきてありがたかった
3	SQL は使った事がなく、非常に怖いイメージがありましたが、少し払拭され、自分でもいじってみようと言う前向きな気持ちになりました。
3	SQL は使用したことがなかったのですが、実例を交えてご説明いただけたので理解しやすかったです。
3	SQL は触りしかやったことがなかったのですが、丁寧な説明で分かりやすかった
3	SQL は便利そうだが統計ソフトで間に合うのでは？という印象でこれまで後回しにしていました。本講義で SQL の使いどころがよくわかったので、今後 SQL の勉強も始めてみようと思います。
3	SQL を作ったことはないですが、今後作る上でこれをみながらやればできそうと思いました。とっつきやすくしていただきありがとうございます。
3	SQL を使いたいと思っていたので基本的な部分を整理できました。とても丁寧に説明されていて初学者にもよく理解できました。
3	SQL を使う意義が分かった。ぜひ一度使用してみようと思う。
3	SQL を使ったことがないのですが、比較的簡単な部分をわかりやすく学べました。R の構文も併記してあり助かります。
3	SQL を使っていたけれど、なぜ SQL の処理が速いかなどはわかっていなかったのが、勉強になりました。
3	SQL 概論、SQL 入門（基本構文、条件や並び替えの記載実例、集計値の実例、変数作成の実例、テーブルの結合）を幅広く紹介いただき非常に参考になりました。
3	SQL 言語の実践的な使い方を学ぶことができて良かった。
3	SQL 初級者であったが、概要を理解することができ有意義であった。
3	この種のご講義を拝聴する機会はありませんためとても有益でした。
3	これから SQL を扱う予定でしたので、丁寧かつわかりやすく非常に参考になりました。統計ソフトとの変数も合わせて掲載していただいたのでより理解しやすかったです。再度スライドを参照しながら自分の手でやりたいと思います。
3	スライドでクエリと結果の両方が示されており、どのような命令文でどのような処理がされるか一目でわかるのが良かった。
3	テーブル結合の 丁寧な例がわかりやすかったです。Distinct の使い方を駆使して導きたいものを淘汰したいとも思います・
3	テーブル結合の注意点と解決方法について学べて良かったです。
3	テキストを読むだけでは理解できなかった SQL 文の作り方について、よく理解できました。試してみようと思えました。
3	なかなか学ぶ機会がなかった SQL 入門としてとてもわかりやすかった。
3	ビッグデータを扱う上での障壁と、それを攻略するための方法を概観できた。
3	まず SQL が何かという点と、その作成・使用内容について理解できた
3	まずビッグデータ活用のおいて、なぜ SQL なのか、その理由がよくわかりました。また基本の関数をビッグデータを例にしながら学べたのはとても良い機会であった。
3	ミスが発生しやすいテーブル結合の部分で、概念図とクエリ情報が1枚になっており理解しやすかった
3	リレーショナルデータベースの仕組みがよくわかりました。
3	リレーションズを用いて RDB の解説がされていてよかった。
3	レセプトデータのデータ構造 SQL を用いたデータ結合の実際

3	医療者に向けた SQL の資料などはあまりないと思うのでとても勉強になりました
3	基本的な SQL 構文と細かい注意点が分かり、さらに STATA や R の場合のコマンドもあった点が良かったです。また、要約量計算ができることも分かりました。
3	記号のうちかたなど基礎的なことも丁寧に教えてくれたこと。テーブルの結合方法。
3	具体的なデータを表示していただいたので、クエリと実際のデータの動きを視覚的に理解できた
3	具体的にデータを示して説明いただいたため、使い方を明確に理解することができた。また、R のコードも合わせて説明いただいたため、R ユーザーとしてはとても有り難かった。
3	具体例ありイメージが付きやすかったです。 基本的にテーブル結合ばかり使用していたので、他の SQL も積極的に使用していきたいと思いました。
3	具体例が多くてよかったです。 JOIN のテーブルの図が分かりやすかったです。
3	敬遠していた SQL の概略を学ぶことができました。ハードルが少し下がりました。ありがとうございました。
3	結合で左外部結合以外を忘れがちなので、整理してご説明いただき、記憶しやすく大変助かりました。
3	見よう見まねでやっていた SQL 操作の意味がいろいろわかりました。
3	後半の結合の細かい内容が助かりました。
3	実際に SQL を使う上で、問題になることがポイントを押さえて説明があり、大変ためになりました。
3	設定の仕方を丁寧に説明していただき、今度 SQL 使用するときの参考にしようと思いました。
3	早速トライしてみます。
3	大規模データでの SQL の扱いが勉強になりました。
3	定期的なまとめや、Stata のコマンドなどでの同様の例も含めていただき大変わかりやすかったです。丁寧なご説明感謝です。
3	定期的なまとめスライドを入れてくださり、とても分かりやすかったです。
3	統計ソフトだけでは不十分な場合が多いこと。(解析する前にデータセット作成に使える必要性が高い所)
3	統計ソフトをもちいてデータの抽出や結合を行っていたが、SQL を使った方が効率的かつ早いことが分かった。 基本的な構文やそれによる結果が端的に分かりやすくまとまっていた。
3	入門編としての説明が大変分かりやすかったです。データの切り出しに苦勞しており、医療データポイントを絞って説明いただいたことで、一度試してみたいと思います。
3	非情報系の方に向けてわかりやすくまとめられていたとおもいます
3	普段から R を使用しているのですが、SQL も文法に気をつければ対応できそうなことがわかりました。
4	DPC データの構造が非常に複雑なため、演習を通じて体系的に集計方法を学ぶことでできて良かった。
4	NDB からの目的データの抽出、特に ID の追跡について具体的な方法を学べた。
4	NDB が複雑だと聞いたことがあったが、どう複雑なのかよく理解できた。
4	NDB データに触れる機会がないので、構造理解が進みました。
4	NDB データに対する SQL の活用の仕方を学べて良かった。
4	NDB データのハンドリングについて必要な知識を学ぶことができた

4	NDB データのハンドリングに必要な SQL の知識が少し理解できた
4	NDB データの構造について理解が深まった。データ利用についてもイメージができた。
4	NDB データの構造の概要を学ぶことが出来ました
4	NDB データの構造やハンドリング方法はとても複雑であることをよくわかりました。
4	NDB データの構造やレコード別の分類の方法等を具体的に教えていただき大変勉強になりました。SQL のクエリについても提示していただき、データ抽出のイメージがより明確になりました。演習のページも構築していただき、実際に手を動かしながら学ぶことができました。
4	NDB データの構造をイメージすることができた。また SQL の実践を理解することができた。
4	NDB データ構造と、実際のハンズオンでやってみることができたのがとてもよかったです。難易度が非常に高く理解できないことが多かったのですが、講義はとても分かりやすく、資料も充実していてとてもよかったです。
4	NDB について理解を深めることができました。実際にデータを見せながらご説明くださったので、理解しやすかったです。
4	NDB のハンドリングに必要な SQL が演習付きで学べたこと
4	NDB の構造について、参照すべき資料を交えて進めていただいた点
4	NDB の成り立ち、構造、利活用事例や ID 構造と縦断データ等、あまり NDB を使用したことがないものにとっては非常に参考になりました。演習問題が多くありましたので、実際に試すことができ良かったです。
4	NDB を使った経験がなかったため、データの構造を始め、実際のデータハンドリングの手順の概要を説明いただきとても有益でした。
4	NDB を用いた研究で実際に使用するデータシートの数がどの程度あるのかを把握できた。
4	NDB 分析のための DB 構築を行う参考になりました
4	postgresql について知ることができてよかった
4	SQL だけでなく正規化された商用データベースの前のレセプトの構造やマスタについてもご紹介いただき大変勉強になりました。
4	SQL については全く触ったことが無かったため、データハンドリングの流れが分かった。実習については自分の理解がついていかなかった
4	SQL に触れることができる環境を実際にオープンソースで提示して頂けるのはとてもありがたかったです。
4	SQL を使用した NDB のハンドリング方法を充実した演習で手を動かすことで学ぶことができました。今後困った際にも参考にさせていただきます。
4	ウェビナー用にページを設けていただいて自分の手元で操作できた点。
4	ウェブサイト が洗練されてとても楽しく講義を受けることができました。演習問題も今後の研究の基礎となりとてもありがたいです。
4	エクセルで表示した時に隠れていたセルが見えたこと。 SQL の演習とデータセットの例示の説明が具体的だったことがよかった。 厚労省のデータ構造の開示がわかりやすいという説明も良かった。
4	オンサイトリサーチセンターでは PostgeSQL が入っていること。
4	クエリの発行方法およびレセプトデータの取り扱いで多用するテクニック・頻発する状況を知ることができ、勉強になりました。
4	これまで自分なりの勉強でデータ処理していましたので、ハンドリングの流れを体系的に学べとても良かったです。演習もありがたいです。
4	データの見方について知りたいことを調べるのにどこをしらべればいいのかわからなかったが、それが明確になって良かったです。

4	データベース形式について詳しく解説していただき、助かりました。ありがとうございました。
4	ハンズオン形式の講義、とても勉強になりました。初めてレセプトデータに触ったので非常に難しく、混乱も多かったのですが、利用への1歩目を踏み出せたと思います。もっと触ってみます。
4	レセプトデータのCSVファイルについて詳しく解説されていたことが良かったです。
4	レセプトデータの実際の様子を今まで見たことがなかったため、勉強になりました。一方で研究に利用できるような状態に変換することが非常に重要だと改めて実感しました。
4	演習形式で実施環境まで準備してくださっていて、手を動かしながら実施できたのでSQLの理解が深まりました。
4	演習問題と演習環境が用意されていたため、実際に自分でコードを書きながら理解を深めることができました。また、NDBデータのハンドリングイメージを掴むことができました。
4	過去に自院のレセプトデータについて、厚労省公開の資料を参考に解析していた経験があったので、一般的な手法の理解がすすんだこと、手順が誤っていないことの確認ができたことが良かった。
4	各種レセプト構造を掘り下げ、実際のデータハンドリングまでわかりやすかったです。
4	自身のSQLに関する理解が不足している点を再確認できました。期間内に動画を見直したいと思います。
4	実際にサンプルデータを用いて実践できたのがよかったです。
4	実際にレセプトがどう表示されるか見せていただき、手順も手取り足取りでみせていただきわかりやすかったです。演習環境を準備していただきありがとうございました。
4	実際のNDBデータ（サンプルデータ）を用いて構造を説明していただけたので、理解しやすかった。
4	実際の演習があったのが良かったです。また、NDBデータの構想を理解することができたのも大変参考になりました。
4	匿名化ID1と2の違いと追跡方法を学ぶことができてよかった。実際のレセプトデータのみながらの授業でわかりやすかったです。
4	難易度は高いですが、実際に使うときにまた復習できそうです
4	別添8 申出依頼テンプレート（抽出）にNDBの列名が隠されているのは知らなかったもので、分かってよかった。
5	<ul style="list-style-type: none"> ・先行NDB研究レビューの要約を見ることでこれらのデータベースを活用した研究の動向を把握することができた。 ・実際に行われた研究を見ることで、研究の具体的なイメージを持つことができた。 ・
5	<ol style="list-style-type: none"> 1. 検索エンジンごとの検索用語を示していただいたので、論文レビューの参考になりました。 2. NDB研究のトレンドを把握できました。
5	DBの概要、研究の流れについて知ることができた。
5	DBを用いた研究の具体例が分かってよかった。分析疫学研究が多いと思っていたので、記述疫学研究が多くて驚いた。
5	DB研究のレビューは見たことがなかったので非常に良かった
5	DPCデータを用いた研究の具体的な紹介があり、どのような研究がなされているかのアイデアが得られた
5	DPCを用いた研究が増加しており、国民全体をカバーできることが参考になりました。
5	NBD,DBCデータを利用した論文投稿が盛んになっており、先行研究の内容を例にデータベースの活用によって明らかになったことを知れて良かったです。

5	NDB、DPC データなどを用いた研究に関して詳しく学ぶことができました。感謝申し上げます。
5	NDB、DPC データなどを用いた研究の近年の動向を知ることができた。
5	NDB、NDB オープンデータ研究論文数の推移をお示しいただいたこと
5	NDB データで Target trial emulation がされたというのは知りませんでした。読んでみたいと思います。
5	NDB データに関する研究で診療科ごとの件数を学べたこと。広い知識も必要だが、NDB 研究をするうえで疾患を絞り込むときに今回教えてもらった診療科数を用いて考えたい（今年一年のうちでは）
5	NDB データの限界がよくわかりました。
5	NDB データの弱点について触れていること
5	NDB データの利活用事例が学べ、視点が広がりました。
5	NDB データベースの強みや様々な研究の実例を知ることができ、学びが多かったです。
5	NDB データを用いた研究について、近年の動向を知ることができました。実際の先行研究について、ご教示いただくことでより具体的に理解できました。
5	NDB データを用いた先行研究をご紹介いただき、NDB がどういった研究に向いているか理解できました。
5	NDB データを用いた先行研究をしれてよかった
5	NDB データを用いた臨床研究の具体例をご教示いただき、使用に関してイメージが湧いた。NDB に歯科のデータが多数含まれている点が印象に残った。
5	NDB データを利用した研究ができるようになることを目標にしているのので、講義内容すべてが大変参考になりました。
5	NDB データを利用した研究をしたいと考えており、その準備に必要な基礎的知識を得ることができたこと。
5	NDB の強みとして歯科の検診データが含まれるという点がよく理解できた。
5	NDB の傾向がよくわかった。また診療科毎により研究されていない診療科があったというのは大変興味深かった
5	NDB をはじめとした代表的な RWD を用いた DB 研究の近年の状況や具体的な研究例等を知ることができた
5	NDB を使うにはハードルを高く感じていたので、改めてデータの概要を学ぶことができた
5	NDB を使用した研究の現状が分かりやすくまとめて頂いて、先行研究の理解につながりました
5	NDB を用いた研究の動向や、研究の例についてもご紹介していただき、NDB でどのような研究が行えそうかというイメージを持つことができました。
5	NDB を利用した研究に興味があったので、具体例をしめしていただけてとても勉強になりました。
5	NDB 関連の研究を網羅的に調査されており、今後の研究立案の際に非常に役立つような内容で、勉強になりました。
5	NDB 研究で実際にどのようなことが行われているのか、リアルワールドデータを用いた研究がどの程度行われているのかわかってよかった。
5	NDB 研究のレビューは実はそれほど多く行われていない 調整結果における研究分野について、内科の DB が一番使われている DB によって、記述研究や生活習慣病研究など得意とする分野が異なる
5	PubMed 等での先行研究の調査により、毎年 NDB 等のデータベース研究が増加していることは理解していたが、該当する論文数が 1000 報を超えていることには驚きました。ありがとうございました。

5	データベースを活用した研究は増加傾向であること、 limitation もあることがよく理解できた。
5	どのような NDB 研究があるのか学ぶことができ良かった。データベースにより診療科ごとの研究割合が異なる点も初めて知ることができ勉強になった。
5	近年リアルワールドデータの研究が増えてきているのがはっきりと分かり有益な情報でした。
5	研究内容や見たいアウトカムによって、DB 選択が重要であること。 実際の先行研究について limitation も含めて解説いただき非常にわかりやすかった
5	後ろ向き研究においてエビデンスをどのように創出するかが分かった
5	国内での RWD 研究の動向が聞けた
5	参考書やネットで自己学習している中で解決できなかったことが分かりました。
5	資料がシンプルでわかりやすかったです。
5	歯科と医科のデータを連携した研究が興味深かった
5	歯科について触れてくれる機会があまりないので良かったです。
5	先行研究によりどのような PECO に落とし込めばいいのかイメージができました
5	先行研究の具体例が今後の研究テーマを考えるにあたってよかったです。
5	他データベースと比較した NDB の利用状況や研究事例を把握できて勉強になりました。
5	多くの国民をカバーしているため、どんな疾患でもある程度妥当性を持たせられる強みがあるということを知ることができた。
5	日本国内の医療受給者の動向を 98%把握できる database は 価値があると思います。今後の活用が期待されます。
5	文献の検索式も提示してくださったのが、実際に自分で検索する際にも参考になるのでありがたかったです。
6	1つの研究について限界も含めて丁寧に説明いただいたので、今後研究を実施する際に参考になりました。
6	DB 活用方法
6	DPC データではどんなことがわかるのか、また、DPC データ研究の強みを理解することができた。
6	DPC データで出来る事、限界などがわかりやすかった。
6	DPC データについて、その活用法や活用事例について理解できました。
6	DPC データについては知らないことが多く、非常に理解の助けとなりました。
6	DPC データについて知識がほとんど無かったので、基礎的知識を得ることができてよかった。 恥ずかしながら、NDB から提供される「DPC」との区別について分かっていなかったもので、理解できてよかった。
6	DPC データの概要(どんなデータが格納されているか、その特徴) DB 研究では様々なバックグラウンドを持った方々と協業できること 研究事例(具体的解析手法、薬剤性肺障害の高い療法の特定) 研究の限界(交互作用、交絡の影響の可能性)
6	DPC データの基礎から説明していただき、実際の研究結果も示していただけたので、DPC データ研究についての理解を深めることができました。
6	DPC データの基礎的な内容や得られるデータ内容が分かり、研究計画のイメージが行いやすい内容でした。また、前向き研究の難しいケースなどを知ることができてよかったです。
6	DPC データの研究について、その研究限界含めて、示して頂いたので、大変分かりやすい内容でした。
6	DPC データの研究の方法論が分かりやすくまとまっていた。

6	DPC データの研究計画書作成について学ぶ機会はなかったのでここで学ぶことができて良かったです。
6	DPC データの構造、含まれる情報、方法論、抽出方法、利点と限界に関して詳しく学びました。感謝申し上げます。
6	DPC データの構造の詳細を知ることができました。 事例紹介も、どのような研究をどのように行ったかが分かりやすかったです。
6	DPC データベースに含まれている項目や、それを用いた研究の具体例についてもご教授いただき大変勉強になりました。
6	DPC データベースの説明、DPC データベースに適している研究、DPC データベース研究の限界等、非常に幅広く講義が行われた非常に現状になりました。 特に、多職種が研究に参加可能ということで、研究への意欲も上がってきました。 実際の実例の説明もあり、研究の実例、限界も非常に分かりやすかった
6	DPC データベースの長所を生かした研究の例がわかりやすかった
6	DPC データベースの利点欠点と研究事例を教えていただいてわかりやすかったです
6	DPC データベースを活用した研究で、詳細な解析も可能であることがわかりました。ありがとうございました。
6	DPC データを使うためにどのような項目があり、研究への活用を知ることができた点。
6	DPC データを使う研究の発想の部分、どのデータがある、ないなど。
6	DPC データを用いた研究がかなり具体的に示していただき、これまで以上にこれらの研究がどのようなものであるかのイメージができました。
6	DPC データを利用した研究の方法論について学べたのはよかった
6	DPC データを利用した研究計画書、データ抽出依頼書の作成方法
6	DPC データを利用する際は未測定の交絡や患者が病院を変えると追跡ができないといった限界があることがわかったのでよかったです。
6	DPC データ研究と NDB の違いはがわかりやすかったです
6	DPC データ研究について実例を交えて紹介いただきありがとうございました。適したテーマなどわかりやすかったです。
6	DPC の構造、具体的に含まれているデータや、DPC データを用いて行える研究について学べた。
6	DPC を活用した研究を実施する際に注意すべきことが良くまとまっており、分かりやすかったです。
6	DPC を利用した研究に興味があったので、具体例が知れて大変勉強になりました。
6	DPC 研究で case control study できること
6	DPC 研究について
6	DPC 研究の魅力的な点とそれを具体的にどのようなテーマで実施されたのかが明瞭に理解できた。
6	DPC 導入の影響評価に係る調査を改めて見直そうというきっかけになりました。
6	ILD の定義について
6	Limitation が明示されていて良かったです
6	一般的な DPC データベース研究対象の向き不向き、限界に関することと具体的研究事例
6	基本的な DPC データの情報から、研究の例を詳細にご説明いただきとてもありがたかったです。どのようなテーマが適しているか、制限があるかなどの例もとても有用な情報だと思いました。また必ず見返すスライドの一つになると思います。

6	講義内容すべてがとても参考になりました。計画書や依頼書の作成方法まで含まれ、感激いたしました。
6	実際の解析事例を紹介していただき、DPC をデータベースでどのような研究が可能であるか具体的なイメージを持ってました。
6	実際の臨床に即した研究で興味深かった。たしかに実体験的にも EGFR 受容体拮抗薬は薬剤性肺障害おきやすいな、という印象でしたので、腑に落ちた感じがします。
6	初学者ですが NDB や民間 DB とはデータがかなり異なる印象を受けました。より治療に近いデータが利用できるのかと思いました。またデータの t 取り扱いや解析手法なども参考になりました
6	商用だけでなく、厚生労働省から個票データが提供されている点を知り、勉強になりました。
6	前向き研究がむずかしい研究で、DPC データ利用ができる事がわかって良かった。また、前向き研究の代わりにできる点が学べて良かった。
6	全て。DPC のことがいろいろわかってよかった。
6	対象とする薬剤の選択方法を具体的に説明されていてよかったと思いました。
6	丁寧な処理によって扱いの難しい薬剤性肺障害も研究が可能であることを学ぶことができよかったです。
6	発想が豊かになるほど今後 DPC データは活用される可能性を秘めていることが分かりました。
6	病院を変えると追跡できないこと。
6	有害事象の発生リスクに関する研究にも DPC データベースが利用できることや、実際にどのように利用して研究されるかを知ることができてよかったです。
6	利用実績の多い DPC データベースの概要、実際の研究内容とデータ特性上の限界について学べてよかった
7	DPC が急性期疾患の解析に向いていること
7	DPC データがどのような研究へ発展しているのかがわかりました
7	DPC データから明らかにできる RQ が幅広いことが理解できました。
7	DPC データが救急や外科領域以外で、医療の質評価や政策評価に活用できる点を理解できました。
7	DPC データが向いている研究デザインなどについて学ぶことができよかったです。
7	DPC データに適した研究領域（救急・外科急性期）と実例を見れたこと。
7	DPC データの強みや具体的な事例でどのような調整を行っているのかなどの実例を見ることができ良かった
7	DPC データの利点（大規模、母集団代表制）や欠点（病院を変えると追跡できない、検査結果はなく）、適した研究等に関して説明があり、実際の研究の紹介があり、イメージをすることができた。
7	DPC データはどのような研究に適しているか、実際にどういう研究が可能かという点について大変勉強となりました。
7	DPC データベースのハンドリングについて学ぶことができた。
7	DPC データベースの活用
7	DPC データベースを用いた研究の長所や、限界への対処について具体的な研究をもとにご教授いただき大変勉強になりました。
7	DPC データベースを用いた研究例とその中での注意点などを具体的に学ぶことができました。
7	DPC データを用いた記述疫学研究の実例は大変勉強になりました。論文も拝読させて頂きたいと思います。
7	DPC データを用いた研究が臨床疫学の研究分類においてどの研究に向いているのか、理解できた

7	DPC データを用いた研究について、筆者の立場から手法や limitation について解説があったこと
7	DPC データを用いた研究の具体例を知れて、大変参考になりました。医療政策に影響を与えうるようなデータもあり、大変興味深かったです。
7	DPC データを用いた研究の利点欠点を学べたこと。DPC データを用いた研究について先行研究の解説を通じて方法論を学べたこと
7	DPC データを用いた研究は、全体像を把握するのに有用性が高いことを理解しました。救急、急性期研究例が興味深かったです。
7	DPC データを用いた研究例を複数例知れたことと、各研究でどのような Limitation が発生するかが大変参考になりました。
7	DPC データを用いた実際の研究、その結果について日本語で学ぶことができた
7	DPC データを用いた実際の先行研究例を通じて、DPC データの特徴や可能性について理解が深まった。
7	DPC データを用いた例を通してさまざまな解析手法や制限に対する対策を学べた
7	DPC データを用いて具体的にどのような研究が行われているのか、また限界についてわかりやすかった。
7	DPC は急性疾患の研究に向いている ICU の稼働率という概念を Primary Endpoint に設定し、ICU 患者を特定可能、米国とその大小の比較 分割時系列デザインにより、傾きの評価を行う(上昇傾向かどうか) 真の統合失調症の道程のため、抗精神疾患薬を投与している患者に絞る 多変量解析において p 値を用いずに、 ASD を使用すること(n が多いと p 値では統計的有意差が出やすくなるため)
7	DPC を用いた救急・外科領域の研究事例の概要を学ぶことができた。
7	limitation に対して、様々な手法を適用して、多角的に検討されている点が勉強になりました。
7	Limitation への対処の説明が実際のわかりやすかったです。
7	p 値と ASD の使い分け DPC データの研究の具体例
7	Restricted cubic splines が学べたのがよかった。
7	さまざまな研究について解析法やその意味もふくめて解説されていてよかった。
7	さまざまな手法と組み合わせてデータを処理する必要があることがわかってよかったです。
7	データベースの利点や欠点だけでなく、どのような解析や工夫をしているかがわかってよかったです。
7	医療政策への利用についても利用できるとわかり、とてもよかった
7	課題に対して具体的にどのような手法を用いて対応したか、また結果の解釈を丁寧にご説明いただき大変実用的な講義でした。ご紹介いただいた手法は知らないものが多く (RCS 分析、マルチレベル分析など)、これから学んでいこうと思いました。
7	感度分析のガイダンスを知ることができた
7	希少疾患に DPC 解析を用いるのが有用だとわかりました
7	救急・集中治療医学分野の例を学べた点。後半では、マルチレベル分析を併用して結果を比較する点が参考になりました
7	救急や外科領域での DPC データ研究について知ることができた点。
7	救急領域で DB 研究は強いなと感じました
7	救急領域と外科領域の具体的な研究を簡潔に分かりやすく説明していただきました。

7	救急領域や、外科領域といった短期的アウトカムをみるのに、DPC データの使用が適していることが分かった。
7	具体的な研究事例に加えて、各研究の Limitation や信頼性を改善するための工夫についてもご紹介いただいております、研究を進めるうえで大変参考になりました。
7	具体的な研究事例を基にどのような解析を行い、どのようなグラフを作成していたのかが良くわかった。参考にしたい。
7	具体的な事例を通して DPC 研究の強みや解析法を理解することができました。自分で RQ を考える上でとても役に立つと思います。
7	後半では吹き出しにどうしたら精度が上がるのか等が補足されており、該当箇所の理解が進みました
7	交絡に対しさまざまな方法で対処することを学んだ
7	交絡因子を取り除くための手法が大変工夫されているなと思いました。
7	治療効果のみではなく、政策評価や費用対効果の視点からの事例
7	実際の事例、特にどのような感度分析が実施されているのかについて、興味深く拝聴しました。
7	実際の臨床現場のクリニカルクエスチョンから DPC データを用いた研究を行っていて、モチベーションと参考になった。 入院データを活かした論点なのがよかった。
7	実際の論文から、DPC データから示される数値を論文の表やテーブルにどのように示すのかという点と、それらの理解について学ぶことができた
7	重み付けやマッチングの様々な方法を知ることができたため。
7	色々な解析手法や、代替変数を用いることがわかって良かった。
7	第一線で活躍している研究者の DPC 研究の実例をまなぶことができた点
7	入手不能な因子について、代替変数での評価や疫学的手法での対処方法が参考になりました。
7	配慮すべき注意点とそれへの対処法、頑健性を高めるための疫学的手法の具体例を知ることができた
7	病名と処方との組み合わせで対象者の像を同定する点 紹介研究内の感度分析手法
7	複数の研究を提示していただいたので、実際の研究に DPC を利用するイメージができた。特に政策評価に利用しやすいという点はとても納得できた。
7	分割時系列デザインについて
7	豊富な具体例で、DPC データの使い方について理解を深めることができました。
7	明日からできそうな研究例で参考になった。
7	臨床に則した研究内容であり、とても勉強になりました。
8	DeSC のデータを使ったことがなかったので、概要やどのような研究で適用可能かを学ぶことができた点
8	<ul style="list-style-type: none"> ・ DeSC データの構造 ・ NDC と比較した、DeSC データの利便性 ・ 高齢者を含む研究での優位性
8	<ul style="list-style-type: none"> ・ DeSC データベースの概要 ・ NDB との比較や保険者別での実際のデータ
8	DaSC データは馴染みが薄かったので学べてよかった
8	DeCS-DB は初見だったので概要を知れてよかった
8	DeSC DB の利活用に関して、認識を深めることができました。

8	DeSC-DB の概要と保険者別の代表性をしれました
8	DeSC が他データベースや日本人全体の母集団と比較してどのような特徴を持っているのか、研究で用いる場合に配慮すべき点について学べた。
8	DeSC が母集団をある程度反映しているということがわかった。
8	DeSC データがどのようなもので、どういった研究ができるかを学べてよかった。
8	DeSC データが全年齢をカバーしているのはとても良さそう。
8	DeSC データと NDB データにおける有病割合の比較
8	Desc データというものをはじめて知りました。1000 万人を超えるデータが取得可能で非常に驚きました。
8	DeSC データについてはほとんど知らなかったため、概要を知ることができ非常に参考になりました。
8	DeSC データについては知らないことが多かったので、どのようなデータベースなのか、知ることができて良かったです。
8	DeSC データについてほとんど知らなかったので興味深い講義であった。
8	DeSC データについて初めて学ぶことができ、今後の活用可能性について考えることができた。
8	DeSC データについて知らなかったなので、その概要がわかりました。ありがとうございます。
8	DeSC データの構造等、基礎的な部分から御説明いただくことが出来、研究の具体的な手法や利点について理解が深まりました。
8	DeSC データの詳細を初めて学びました。DeSC データの代表性なども論文として公表していただき大変勉強になります。
8	DeSC データの特性や構造、研究の具体的について知ることができた。保険毎の比較ができること、全年齢対象の母集団性が確保できることを知り、行政の立場から活用したいと感じた。
8	DeSC データの特徴がよく理解できた 他のデータベースとの比較が参考になった
8	DeSC データの特徴と母集団代表性について大変勉強になりました。
8	DeSC データの特徴について調査をもとに説明いただき理解できました。
8	DeSC データの特徴や研究の実例を知ることができて良かった。
8	DeSC データの内容を知りたかったので非常に役に立った
8	DeSc データは、今回の webinar で初めて知りました。大変勉強になりました。
8	DeSC データベースが NDB とどう違うのかがわかった。
8	DeSC データベースについて、本データベースを用いる利点を学べてよかった。
8	DeSC データベースについてこれまで全く知らなかったなので、概要を学ぶことができた
8	DeSC データベースについての研究などが学べて良かった
8	DeSC データベースについてもともとあまりよく知らなかったなので、 大変貴重な機会となりありがたかったです。
8	DeSC データベースについて学べて良かった。特に、初学者としては、データベースごとの違い等がよくわからなかったなので、比較がわかりやすかった。 NDB データ活用のハードルが高いため、まず DeSC データベースを用いた研究を行うのもよい、という事を学べて良かった。
8	DeSC データベースについて初めて知りました。勉強になりました。
8	DeSC データベースについて説明や事例を紹介してもらった

8	DeSC データベースについて全く知らなかったため、NDB 以外のデータベースについて知ることができて良かった
8	DeSC データベースについて聞くのは初めてだったので、その内容や利用法について知ることができた。
8	DeSC データベースに含まれる患者年齢が、日本人の年齢層をある程度反映していることが分かりました。市販データベースとして選択肢が増えたと思います。ありがとうございました。
8	DeSC データベースの詳細がわかり大変興味深かったです。講義の中でも説明がありましたが、NDB のハードルの高さがあることを考慮すると、選定する DB の一つとして考慮していきたいと考えております。
8	DeSC データベースの特徴について、既報論文も引用して紹介していただき大変参考になりました。今後 DeSC データベースを用いた解析を検討しているので、紹介された論文についてもさらに勉強してみようと思います。
8	DeSC データベースの母集団代表性について知りたいと思っていましたので、大変興味深く拝聴しました。
8	DeSC データベースの利点・限界を学ぶことが出来たこと。データベースの母集団代表性研究も非常に勉強になりました。
8	DeSC データベースの利用経験がなく、データベースの特徴を理解することができました
8	DeSC データベースはそもそも知らなかったので勉強になった。
8	DeSC データベースは小児から後期高齢者までの全年齢層をカバーしている事。RWD の利点、欠点。
8	DeSC データベースは全く知らなかったもので、勉強になりました。NDB と大きな差はないこと、保険者ごとの傾向の違いは大変興味深いです。
8	DeSC データベースを使い始めたところだったので、記述統計がとても参考になりました
8	DeSC データベースを用いた研究について、実例を用いてわかりやすく理解できました。
8	Desc データベースを用いた研究の代表例を紹介頂きました。NDB と違い、高齢者のデータが含まれているのは大きなメリットであると感じました
8	DeSC データベース自体を知らなかったもので、利用できるデータベースのひとつとして、参考になりました。
8	DeSC データをあまり知らなかったもので、概要からお話いただけて大変参考になりました。
8	DeSC データを用いた研究についての基本的なところから学ぶことができて良かったです。
8	DeSC では全年齢層がカバー出来ているとのことで、非常に魅力的だと感じました。
8	Desc と NDB との網羅性、アクセスの違いが明解に解説され、役立つ内容でした。
8	DeSC のデータベースの特徴を分かりやすくご説明頂き、大変勉強になりました。特に、各保険者の集団の特徴については、今後の研究立案の際に役立てたいと思います。
8	desc のデータベースを用いて研究を始める所でした。とても参考になりました。査読のディフェンスにも役立つ情報がたくさんあり、よかったです。
8	DeSC の名称は知っていましたが、含まれているデータの種別や年齢分布等が理解出来ました。国勢調査との比較は参考になりました。
8	DeSC の有用性がわかりました、高齢者を含む研究に有用と思いました
8	DeSC は新規のデータベースと思いますが、JMDC などの伝統的なデータベースとの違いがわかりやすくまとまっていると感じました。
8	DeSC を用いた研究事例はあまり詳しくみたことがなかったので分かりやすく解説いただけて有難かったです。

8	DeSC 内で、健保と国保を比較した内容は面白いと思いました。 また、ハードルが高い NDB を扱う前に、まずは DeSC から手を付けるという選択肢があるという事も知れて良かったと思います。
8	JMA Journal の結果の一部は臨床疫学会の DeSC さんブースでもお伺いできましたが、詳しい結果を康永先生からご説明いただけて学びになりました。
8	NDB しか知らなかったのですが、DeSC も知ることができました。
8	NDB データとの違いなども取り上げていただけたのがよかった
8	NDB データベースを用いたことのないものにとって、DeSC データベースが選択肢となりうることを知ることができて良かったです。
8	NDB には含まれない被保険者台帳データの存在を学ぶことができ、勉強になりました。
8	NDB 以外のデータベースの活用例が知れて良かったです。
8	RWD が隆盛するであろう理由や DeSC について学べた点
8	RWD データが今後も多く使用される理由(RCT の限界) RWD の利点・欠点 DeSC データの特徴(NDB との違い、母集団代表性) 健保と国保で各疾患の有病率が異なること
8	RWD を活用した研究が今後増えること。 妥当性、信頼性の高い研究を行うためには DB の利点、欠点を把握することが重要であるということ。
8	このデータベースそのものを知らなかったのですが、概要やできることが理解できてよかった
8	これまで DeSC データベースの詳細を知らなかったのですが、その特徴を理解することができました。
8	これまであまり DeSC データベースに関して情報収集する機会がなかったのですが概要や研究事例を知ることができて良かった
8	データベース間での有病割合などの比較が、興味深かったです。
8	なぜ DeSC データベースが重要なのか、NDB と比較して何が違うのかといった前提条件が非常に丁寧に説明頂いているので以降の理解が進みました。
8	リアルワールドデータの現状と DeSC データの概要について、実際の研究事例を含めて、とても勉強になりました。
8	記述研究の方法やその重要性が理解できました。 DeSC 自体を知らなかったのですが、有用な DB と認知することができました
8	高齢者の研究に興味を持っているので、DeSC データベースが使えることがわかったのがよかった。
8	国保、健保、後期高齢者での各疾患の有病率の相違を知ることができました。
8	国保・後期高齢者のデータを用いて健保での解析と比べてどのような違いが実際出てくるのか（特に高血圧の例）は興味深かったです。
8	今まであまり知らなかった DeSC のデータベースについて、構造や事例を知ることができて良かったです。
8	今回のウェビナー全体を視聴する前は、DeSC データというものさえ知らなかったが、その基本的な内容を理解することができた。
8	使ったことがないデータベースだったので、説明が丁寧でとてもわかりやすかったのがよかったです。
8	初めて DeSC データについて学びました。気になっていた母集団代表性をはじめ、網羅的に扱っていただき理解が深まりました。
8	商用データベースでも NDB を用いた研究と近いものができることがわかりました。NDB よりも利用しやすい点はとても魅力ですし、まずはこれでスキルを身に着けるということも納得できました。

8	先行研究事例がわかりやすかったです。 また、DeSC データの特徴が整理されていて理解が進みました。
8	全年齢層、後期高齢者まで追うことができる DeSC の革新性がよく理解できた。
8	保険者種類別の有病率に差があり、データベース選択時には注意が必要であることを学べてよかったです。
8	母集団の比較、健保と国保の比較が興味深かったです。直結するには至りませんが、自身の研究アイデアの参考になりました。
8	本データベースの特徴や、本データベースの利用が適した事例をご紹介いただきとても有益でした。
8	民間企業でも利用しやすい、NDB の代替となる DB の紹介はありがたい
8	様々なデータ解析で国民全体の政策、結果の有用性を見ることが大切だと改めて思いました。
8	論文で DeSC データの母集団代表性を示して下さったことで、今後 DeSC データを使った研究がしやすくなりました。感謝申し上げます。
9	<ul style="list-style-type: none"> ・JMDC データの構造や他のデータベースと比較した優位性と限界 ・身体疾患を対象とした JMDC データを用いた先行研究事例
9	<ul style="list-style-type: none"> ・JMDC の存在と DB の特徴(利点、限界) ・処方・処置・手術の保険点数データの活用例について ・若年成人、小児を対象とする研究が強みである一方で、高齢者を対象とする研究には不向きである。 ・どの DB にも言えることであるが、データハンドリングは重要である点
9	2 つ目の研究紹介は、JMDC データの強み(若めの母集団 n 数が多い)をイメージできる内容だったので良かった
9	DB の特徴として数は多いが、全年代をカバーしているのではないということがわかり、DB の特徴がわかりました
9	JMDC に含まれるデータやその特色、強みと弱みなどが、実際の研究事例をとおして勉強できました。
9	JMDA データを活用した研究において、解析上注意すべき点、利点欠点について具体的に説明がありよく理解できました。
9	JMDC データの pit fall
9	JMDC (健保データ) の特徴と弱点に分かってよかった。
9	JMDC データについて利点と限界が学べてよかった。 背景や JMDC データを使う理由も説明してもらえてよかった。
9	JMDC データの概要 (含まれる情報、特徴、限界) 先行研究、研究の利点と限界等に関して、研究事例の紹介があり、とてもイメージがしやすかったです。
9	JMDC データの研究利用について JMDC 側ではなく利用者側の視点で知ることができて良かった。
9	JMDC データの構造、含まれる情報、具体的な研究内容を提示して頂いたお陰で、研究の利点と限界を理解することができました。
9	JMDC データの構造、含まれる情報について、また、それを用いた先行研究について学ぶことができました。利点のみならず、限界について理解することが出来て良かったです。
9	JMDC データの構造、含まれる情報を理解できた
9	JMDC データの有用性。検診データとの連結により、腎機能低下や高血圧の程度によって発症する疾患リスクがわかること。大変参考になりました。ありがとうございました。
9	JMDC データの利点と限界について整理することができました。リスク調整のスライドについて、各項目とそれの対応策をご教示いただき大変勉強になりました。
9	JMDC データベースの得意なところ、他のデータベースとの違いについて簡潔にまとめていただきわかりやすかったです。また、それらの特徴を活かした研究を紹介していただくことで、より具体的なイメージを捉えることができました。

9	JMDC データベースの特徴とその事例がわかりましたので、活用する際の参考になると思いました。
9	JMDC データベース研究で high impact journal を目指せる研究ができることがわかりました。
9	JMDC データをあまり知らなかったので、その構造や含まれる情報を理解し、先行研究を学ぶことで、どういった研究に向いているのか理解できたのが良かったです。
9	JMDC データを扱っている職員がいるので、どのようなデータを扱っているのか理解が進みました。
9	JMDC データを研究で用いる際の利点や欠点が具体例を交えて紹介して下さったので理解が深まった。
9	JMDC の活用例をお示しいただいたことで、JMDC に向いている研究とその限界がより深く学べました。また純粋に、2つの研究の内容がとても興味深かったです。特に young adult の研究はまさに自分の世代なので、自分のリスク因子を減らせるよう日常生活を見直したいと感じました。ご講義、ありがとうございました。
9	JMDC の基本的情報や利点・欠点、解析の実際などがよくわかりました。また解析例をお示し頂き、具体的なイメージがわかりました。
9	JMDC を使って Top journal へ掲載されている点について、大変素晴らしいと思いました。改めて論文を読んでみたいと思います。
9	NDB よりもアクセスしやすい JMDC の DB 使用の実例について示していただきながら解説いただいたので、非常にわかりやすかった。 Young Adult 世代に強みという視点はなかったので非常に参考になった。
9	なかなか聞きたくても身近で研究を行っている人が居ないので、内容を知ることができて今後の参考になりました。
9	ビックデータ解析が実際の臨床エビデンスを証明する例に興味深くおもしろかったです。
9	リスク調整（対応策）、交絡因子の調整、ドロップアウトの対応がとても参考になりました
9	リスク調整の対応策が示されており良かった JMDC データの特徴と限界が理解できた
9	リスク調整の対応策の具体例があり大変参考になりました。
9	レセプトデータの有効活用、薬剤疫学研究の具体例が提示され、深く理解することができました。
9	レセ病名やピンポイントで目的の項目がなかった際の対策に関してとても参考になりました。
9	一番、興味を持ち使用してみたいデータベースであった
9	家族 ID であったり、JMDC データの有するデータの内容 JMDC データの利点と限界
9	過去に企業で JMDC データを用いた研究をしたが、こうした内容を事前知っておくべきであった。基礎的内容が理解できました。
9	生活習慣病の予防など伝えたい message がはっきりしている場合、たくさんの RQ が浮かぶことを実感しました
9	特定検診に関する研究を行いたいならば、JMDC データを使用する選択肢があること。 JMDC データの利点と欠点を学べたこと。 先行研究を通して研究手法を学べたこと。
10	・傾向スコア、傾向スコアマッチングの概念、利点、推定・確認の方法、 ・他の分析手法同様、事前の綿密な計画、測定している変数の確認等が重要であるということ。(交絡因子)
10	ASD とか C スコアなど馴染みがなかったですがマッチングの流れや方法がわかり勉強になりました。
10	ATE,ATT どちらも確認できることやしようする際の注意点について理解できた。未測定交絡を高次元傾向スコアでどのように調整できるか他の動画で確認しておきたい。

10	ATT, ATE まで触れられており、わかりやすく解説いただき、理解を深めることができました。
10	ATT や ATE など馴染みがない分野でしたが、初歩的な理解はできたと思う
10	C 統計量など傾向スコアの算出時のチェック項目に関する理解が深まった。
10	PS について研究で使用しているので、マッチングや重み付けの研究における立ち位置が分かってよかった。
10	PS マッチング、逆確率による重み付け、使用する上での注意点についてわかりやすかった。
10	PS マッチングしても患者背景にばらつきがあったりする経験もあり、改めて整理することが出来ました。
10	PS マッチングの method の記載の仕方が丁寧に説明して頂けてわかりやすかったです。
10	オーバーラップ重みづけについて、よく理解していなかったので勉強になりました。
10	オーバーラップ重みづけは初めて知ることができました
10	かんたんな例を示して、傾向スコアの意味をわかりやすく説明していただき、理解しやすかった。注意点もよかった。
10	これまでに傾向スコアを用いた統計解析を行った経験がないためとても勉強になりました。ありがとうございました。
10	すぐに活用できるようなわかりやすい説明でしたありがとうございます。
10	データベース研究における傾向スコアについて活用について具体的に理解を深めることができた。
10	どういう場合に使用する手法なのか 傾向スコアの確認など論文を読む上でも有益な情報
10	バランスが取れなかった場合や傾向スコアの解析が不適切だった場合の対処法などが、実際に研究を行う視点でご説明いただけた点が良かったです。
10	マッチングと逆確率重みづけの使い分けが理解できた。 それぞれの手法は独立していて、どちらを事前を選ぶのかと思っていたが、マッチングの状況に応じて適切な方を選択する、というイメージを持つようになった。
10	ロジスティック回帰との違い。c 統計量による傾向スコアの確認
10	解析を取り入れたかったけれども、これまでなかなか方法を知っている人が周りにいなかったこともあり、手が出せなかったのが、とても参考になった。
10	解析手順の理由が、丁寧に述べてあり、とても分かりやすい講義でした。手順の意味が具体的であり、回答を導きだすための考察過程がいかに重要であるかを理解しました。
10	各値の重要性について丁寧な説明があり、なぜ論文に記載する必要があるのかよく分かったこと。
10	基本的なことに加えて、「何を見たいのか？」という観点で研究テーマや手法を選ぶべきであることを改めて復習することができ、良い気付きになりました。
10	近年多くの論文で見られる傾向スコアについて、改めて傾向スコアの使用方法を理解することができてよかったです。
10	傾向スコアについて非常に理解しやすく、また、注意点等も理解できましたので大変勉強になりました。
10	傾向スコアについて勉強した中で一番わかりやすい説明でした。すごかったです!!!
10	傾向スコアに関して、細かい丁寧な講義で、さらに具体的な研究内容を提示して頂いたお陰で、よく理解することができました。
10	傾向スコアに関してわかりやすく説明いただき理解しやすかった。注意点についても理解が進んだ。
10	傾向スコアに関して学ぶことができました。これまで傾向スコアや逆数重みづけなど論文を読んでいて目にする が、正確には理解できていないものについて、学ぶことができて貴重な経験になりました。

10	傾向スコアの目的から注意点・実施方法までを知ることができた。実際の論文でどのように傾向スコアを実施しているかの例示など具体例があったので、すごくわかりやすかった。
10	傾向スコアの利点だけでなく、注意点まで合わせて教えていただけたのが良かったです。特に解析対象の評価可否については疑問に思っていた部分なので対処法までご紹介いただき助かりました。
10	傾向スコアは、わかりそうでわからない分野でした。数式を使わず具体例を示した説明でよかったですと思いました。
10	傾向スコアは何となくの理解だったのが、理解が深まりました。治療を受ける確率という点の理解が深まった。
10	傾向スコアは前々から聞いたことがありましたが、具体的な内容について学ぶ機会がなかったので、今回具体的な例も示して教えていただけてよかったです。
10	傾向スコアマッチング、重み付けについて、例示のおかげで理解が進んだ。論文中の記載例を示して頂けたのが、参考になった。
10	傾向スコアマッチングを行う上でのバランスのとり方、変数選択について学ぶことができました。
10	傾向スコアや傾向スコアを用いた分析の基礎について、とても勉強になりました。また、傾向スコアマッチング・逆確率による重みづけについて、先行研究もご紹介いただき、注意点や用途など、具体的でとても勉強になりました。
10	傾向スコアを用いた解析は一度学んだことがありましたが、あらためてPS マッチングとIPWの推定している効果の違い等、理解を深めることができました。
10	傾向スコアを用いた論文をきちんと理解することが出来ずモヤモヤしていたが、わかりやすくまとめてあり非常に勉強になった。
10	傾向スコアを利用するときの注意点についてよく理解できました。論文での記載例もわかりやすかったです。
10	傾向スコアを利用する時の解析の流れや注意点が非常にわかりやすく、使用できない場合の代替手法の候補も明示され勉強になった。
10	傾向スコア分析について、基礎から重みづけなどを学べたこと。特に逆確率による重みづけはよくわかっていなかったもので、大変勉強になりました。
10	傾向スコア分析の基礎から丁寧にご教授いただき、傾向スコア分析の勉強する上でのハードルが下がりました。傾向スコア分析を用いた先行研究のご紹介のスライドでは、方法、結果などの各パートを詳細にご説明いただき大変ありがたかったです。どこに着目して論文を読めばよいかという点も学ぶことができましたし、研究を進める上でのポイントが大変わかりやすかったです。誠にありがとうございます。
10	傾向スコア分析の基礎に関して、私自身のもやもや感をすっきりさせていただき有難うございました。全体的に、幅広く説明いただきましたので、とても参考になりました。
10	今までの傾向スコア分析の中で一番分かり易かったです。具体的な手法から限界、感度分析まで触れてもらったのが特によかったです。
10	今まで見た傾向スコアの解説の中でもとてもわかりやすかったです。
10	最近傍マッチングのキャリパーの決め方は知りませんでした。
10	最後の「傾向スコア解析の対象」のスライドが、日頃から疑問に思っていたところなので特に勉強になりました。
10	使用方法や概念だけでなく、制限事項等も知ることができ、非常に実用的な内容を学ぶことができたと感じた。
10	事後の事象を傾向スコアに含めてはいけないことを理解できました。
10	治療選択より後に起きるものは投入しないことを、あまり意識できていませんでした。
10	初級者のため、傾向スコアの使い方について順序だてて説明させていただき大変分かりやすい動画でした。実際の手順や、傾向スコアの適応が適さない場合、IPTW やオーバーラップとの考え方の違い等、これまで漠然と疑問に感じていた点を整理いただいております、何度も見返したいと思います。

11	hdPS についてはきちんと学んだことがなかったため、全体像を掴むことができ大変有用だった。
11	HdPS の適用場面やチェックリストについて学ぶことができたのは、非常に有意義であった。
11	hdPS を使う場面が具体的に想像できたこと。
11	High dimensional の PS マッチングについて恥ずかしながら初めて聞きました、理解が深まりました。
11	PS が普及している昨今でも、なかなか高次元 PS を学ぶ機会はないため、大変良い機会であった。
11	PS は知っていたものの、hd-PS は初めて聞いた言葉だったので知れてよかった。
11	このような解析があるのは初めて知った。リハビリテーションの分野ではあまり馴染みがなかった。
11	このような方法があるということを知らなかった
11	コンセプトについて、冒頭のフレイルの例はとても直感的に理解できました。また利点、限界点のサマリが大変分かりやすかったです。
11	スコア作成について、表の活用や説明部分を赤字にするなど対応されているので、何を説明したいのかが分かりやすかったです。
11	チェックリストやガイドツールなども紹介されており、勉強になりました。
11	はじめて高次元傾向スコアを知った
11	はじめて知る概念だったが理解しやすかった
11	やり方がいまいちわからなかった高次元傾向スコアのやり方をわかりやすく学べたこと
11	一般的な PSM しか考慮していなかったため高次元 PSM が学べてよかった
11	一般的な傾向スコアと高次元傾向スコアのちがいをあまり理解していなかったため、チェックリスト等のガイダンスがあったのがよかった
11	気になっていた hdPS について実際に利用して研究されている先生に講義をいただけて勉強になりました。
11	傾向スコアの応用方法について理解することができ またその利点についても理解できた。
11	傾向スコアの拡張の話は聞いたことがなかったので良かったです
11	健診データを高次元傾向スコアの算出に用いていない点、納得感があると共に事例として勉強になりました。
11	高次元傾向スコアに関して、詳細かつ丁寧な講義で、さらに具体的な研究内容を提示して頂いたお陰で、よく理解することができました。感謝申し上げます。
11	高次元傾向スコアに関して、体系的に学べる機会がなかったため、非常にわかりやすかったです。
11	高次元傾向スコアのメリット(特に、変数選択を自動選択してくれることと、データを余すこと使えること)
11	高次元傾向スコアの基本や高次元傾向スコアの利点及び高次元傾向スコア分析を用いた先行研究についての紹介があり参考になりました。以前に実施した研究に関して理解できるようになりました。
11	高次元傾向スコアの限界：中間因子の選択など疑問に感じましたが、動画中に解説があり理解できました。
11	高次元傾向スコアの未測定交絡を間接的に評価できるものと初めて知ることができ、大変参考になりました。
11	高次元傾向スコアの有効性、限界、社会経済学要因の情報が未測定であることなど。
11	高次元傾向スコアの理解ができていなかったのですが、機械的に行う方法だと言うことと、その限界について。高次元傾向スコアのチェックリストのご紹介もありがとうございます。

11	高次元傾向スコアをプロトコールにどのように落とし込むのか解説がありとても良かったです。
11	高次元傾向スコアを始めて知り、今後データベース研究で使用できるよう理解を深めていきたい。
11	高次元傾向スコアを推定する上での細かい計算方法が丁寧に説明されており、詳細を知ることが出来た点が非常に良かった。
11	高次元傾向スコア分析を用いるためのチェックリストをご紹介いただき大変勉強になりました。また、石丸先生が実施された研究の中で、特定検診の受診状況など欠測値が多い変数が、高次元傾向スコアで調整したあと群間で揃ったということがわかり、高次元傾向スコアを使用する意義をより明確にイメージできました。
11	高次元傾向初めて知りました。 ところどころ初めて聞く言葉がありました。 後で調べてみたいと思います。
11	中級者向けで難しい内容だろうが、初心者でもなんとかついていける程度にわかりやすく工夫してくださっていて助かった。
11	難しい分野であるが、実際の大規模データに当てはめた実例は大変ありがたいです。大規模データでRCT類似手法として今後身に着けていきたいです。
11	普段気軽に学べない内容を効率的に学べた。
11	本にも詳しく書いていないことが非常にわかりやすく解説されていてよかった
11	未測定交絡に対する対処方法を知ることができた。
11	未測定交絡の間接的調整 高次元傾向スコアの作成や評価
11	未測定交絡の調整ができる点や変数選択の恣意性の減少などが魅力的に感じた
12	・変数操作法が論文においてどのように使われるのか等、実務に寄せた内容も学ぶことができた点が非常に良かった。
12	3つの前提条件についてのご説明が分かりやすかったです。
12	CACEの計算式についても、分解して丁寧に説明してくださったので、よく理解できました
12	CACEの説明が数式込みでなされたので理解が深まりました。
12	DB研究における操作変数法について理解することができた。
12	Monotonicityに関する詳しい解説
12	Unmeasured confoundersと操作変数法について学ぶことができた。
12	Unmeasured confoundersの対処が可能である操作変数法の基礎理論
12	これまでよく理解できていなかった操作変数の仮定における単調性の必要性について特によくわかりました。また、適切な操作変数を用いない場合にどのような影響が出るのかについてもよくわかりました。
12	やや理解が追いつかない部分もあったが、操作変数法の概要を学ぶことが出来て勉強になった。
12	仮定、なぜバイアスが大きくなるのか、なにの効果を見ているのか、解析手法として取り上げる場合の注意点を網羅的に提示してくださった点
12	検証できる仮定と臨床的直感が大事になる仮定、それぞれが及ぼすバイアスなどがわかりやすく説明してもらえて良かったです。
12	事例の紹介を通して、操作変数法を適用する際の注意点や、仮定からの逸脱を評価した結果の示し方について学べたのが良かったです。
12	実際に過去の研究を紹介いただいた点：自分で遡及して読みに行くことができる。
12	実際の研究を組み立てるにあたって、どのような仮説と結果を想定すればよいかについての重要な内容を学ぶことができた

12	実際操作変数法がどのような時に用いられるか、また”操作変数”自体がどのような物を用いるべきかイメージする事が出来ました。
12	初学者でもわかりやすい資料がいままであまりなかったので非常に勉強になりました。
12	書籍で学習した際にわかりづらかった操作変数法についてわかりやすく学習できた
12	詳細を理解することは難しかったですが、論文を読む際に視界が深まったように感じます、あり
12	図が直感的でわかりやすかった
12	図と説明が非常にわかりやすかった
12	数式も詳しく説明して頂き、非常に理解しやすかったです。
12	数式を用いてさらに説明を口頭で追加していらしたので、注意しながら学ぶことができました
12	選択する変数が仮定を満たすかどうかをどのように判別すべきか、とても難しいと考えていましたが、感度解析として傾向スコア分析と併用されることが多い点など、先行研究の事例に沿って利用していくのが現実的であると理解することができました。
12	前提条件について理解できた。また、限界については十分理解できておらず、具体的に説明していただき、学べてよかったです。
12	操作変数が RCT のくじ引きの役割を果たす という例えが分かりやすかった。 証明できない仮定を置いている点についても分かりやすかった。
12	操作変数という概念が図や実例を交えてわかりやすかった。
12	操作変数方が対処できる具体例として 実際の医療現場、住んでいる環境を加味した研究がとても興味深かったです。
12	操作変数法の仮定などをていねいに説明いただいて、理解が深まりました。また、推定方法の種類などもまとめていただき、勉強になりました。
12	操作変数法の手法と未測定の変数因子の調整ができること、操作変数を見つけるむずかしさや限界が多いことを学べてよかったです。
12	操作変数法の前提条件やそれを解説した論文についてもご紹介いただき大変勉強になります。万能な解析手法があるわけではなく、それぞれの解析手法の利点と欠点を整理することが重要であると学びました。
12	操作変数法の内容解説に留まらず、操作変数法の限界に触れており、理解が深まりました。
12	操作変数法は理解しにくいと思っていたので、わかりやすく解説してもらい助かりました
12	非常にスライドが見やすく、説明も分かりやすかったです。4つの仮定も、詳しく説明してくださりありがとうございました。
13	<ul style="list-style-type: none"> ・不連続回帰デザインが、治療の判断に使用される連続する変数に、閾値がある場合に（検査値により治療が変わる場合や政策等で時期により治療方針が変化した場合の検証に）適用されること ・差の差分分析は、介入群と対象群における経時変化の差を取ることで、介入の影響を検証する手法であること
13	GPSP 省令下で行われるリスク最小化策の効果測定で使えそうなデザインだと思いました。
13	regression discontinuity という考え方を初めて知り、血圧 1mmHg の差が RCT のコイントスと同じような使われ方をするという考え方に興味を持ちました。 限界の外的妥当性について重要だと思いました。
13	Sharp RD と Fuzzy RD の違い
13	これまで使用したことのない手法ではありますが、手法への興味がわきました。今後の選択肢の一つにします。
13	データベース研究において、医療政策の介入効果を調べる方法の一つとして、適した分析方法であることがよくわかりました

13	どちらのデザインも初めて聞くデザインでした。 前後比較は、問題が多いことが多いのでいいデザインを紹介していただきよかったですと思いました。
13	回帰不連続デザインと差の差分分析の表面的なことは知っていましたが、各仮定と先行研究を学ぶことができました。
13	回帰不連続デザインによって、ランダム割り付けに近い状況を作り出す事ができ、アウトカムを比較できる事がわかった。
13	回帰分析について理解を深められ、また新たなことを学ぶことができたのが良かったです。
13	経時変化を評価するセッションでは、段階的に変化する概念図を使用されておりイメージが非常に理解しやすかったです
13	最近の日本からの研究事例をお示しただけなので、各デザインにおける仮定が満たされているかどうかについて、感覚を持ちながらさらに勉強していきたいと思います。非常にわかりやすい誤解左折をありがとうございます
13	少しトリッキーな両手法ですが、具体例を示していただくことでどのような研究テーマに使うことができるのかイメージを掴むことができました。
13	年度等で前後比較をする場合にどのような解析が適切か悩んでいましたので、具体的な研究例とともにご教授いただき大変勉強になりました。
13	不連続回帰デザイン、差の差分分析がどのような臨床疑問の検証に適しているのかを活用例を示してご説明いただき、わかりやすかったです。
13	不連続回帰デザイン・差の差分分析について、手法、留意点はもちろんのこと、どのような臨床研究に適しているかについて理解ができました。違うすぎるのもの
13	不連続回帰デザイン・差の差分分析の基礎的知識のみならず、リサーチクエスト・クリニカルクエストに即してどのような研究デザインを選択するかの概要が理解できた。
13	不連続回帰デザイン分析および差の差分分析の手法・研究デザイン・先行研究について学ぶことができた。これまでの経験の中であまり聞く機会のなかった研究手法であったため、今回の研修で学ぶことが多かったです。今後は自習して更なる理解を深めることをしていきたいと考えております
13	理論的な面だけでなく、実際の論文での例を多く紹介してもらえたため、理解が難しい内容ではあったが、分析の意味する内容を多少なりとも学ぶことができた気がする
13	両デザイン共にあまりなじみがなかったのでその概要、仮定、先行研究について学べてよかった。
13	閾値の近傍でアウトカムの変化を比較する。 導入の前後比較には経時変化の効果が含まれている。経時変化の差を取る。
14	DAG を使った説明が大変わかりやすかったです。
14	DB 研究における時間依存性交絡と周辺構造モデルの活用について知ることができた。
14	MSM-IPTW の考え方についてわかった。DAG がわかりやすかった。
14	いかにバイアスを調整することが大切か学びました。神のみが知るという実際の治療研究を回避する研究方法を学びました。
14	とても難解そうなことを噛み砕いて時間依存性交絡などについて解説して頂けたこと。
14	なかなか教科書にないところでしたのでよく学びました
14	以前、別の論文を読んでいる時に、時間依存性交絡のことが気になったものがありました。当時はそれを時間依存性交絡ということを知りませんでした。今回学べてすっかり理解できました。ありがとうございます
14	疫学的な暴露のほとんどは時間依存性の交絡の可能性が高いことがわかり、その交絡を調整して解析する周辺構造モデルの端緒を理解できた点
14	過調整バイアスについて
14	時間依然性交絡の話や時間性交絡の問題点の一連の講義をあまりきいたことがありませんので、非常に参考になりました。

14	時間依存性の共変量の取り扱いをどうするか、悩んでいたところだったため、わかりやすく教えていただき、大変勉強になりました。
14	時間依存性の治療や交絡の概念から基礎部分に加え、先行研究についても学ぶことができました。
14	時間依存性の治療や交絡の概念について初めて学ぶことができました。モデルの基礎を見直して研究に役立てられるように見直しを繰り返します。
14	時間依存性交絡と周辺構造モデルに関して、詳細かつ丁寧な講義で、さらに具体的な研究内容を提示して頂いたお陰で、よく理解することができました。 感謝申し上げます。
14	時間依存性交絡について、実際に研究で使うまでにはもう少し理解を深めないといけないと感じておりますが、まずは基本的な理論を理解することができたので良かったです。
14	時間依存性交絡の DAG での表現がとても分かりやすかったです。confounder にも mediator にも collider にもなり得ることが分かりました。時間依存性治療のパターンを明確に定めること=consistency という仮定ということも分かりました。
14	治療薬の変遷で前時点での検査値や治療が時間依存性交絡として次の治療に影響与えるとして対応が必要になったケースがあり、理解が深まりました
14	実際の論文を例に時間依存性の項目をどう処理しているのかについて学ぶことができた
14	周辺構造モデルについて簡単に勉強することができた。時間依存性交絡の概要についてわかった。
15	CIF と部分分布ハザードが混合してしまうのでわかりやすかったです。
15	KM を引くことはよくあり、競合リスクを考慮することの重要性が改めて整理できました。
15	カプランマイヤーの原理と、競合リスクモデルの説明がわかりやすかった。
15	カプランマイヤーやログランク検定など臨床研究で用いていたものを基礎から学び直せて良かった。
15	シェーンフィールド残差の解釈
15	それぞれのモデルについて、先行研究を用いて具体的に説明していただいて大変わかりやすかった
15	過大評価、過小評価の議論が参考になった。
15	競合リスクの概念について理解できた。また、生存時間分析については知っていたが、競合リスクを考慮した生存時間分析の手法については知らなかったため、とても勉強になった。
15	競合リスクを考慮した解析の必要性が想定される研究を計画していたため、わかりやすく参考になった。
15	具体的に対応プログラムも紹介されていてわかりやすかったです。
15	原因別ハザード比と部分分布ハザード比の違いがよく分かりました。
15	使い分けや論文の提示があったため、理解しやすかったです。
15	生存時間分析において競合リスクが存在する場合、どのような分析手法があるか学ぶことができてよかった
15	生存時間分析についてわかったつもりでいたが、まだまだ理解が甘かったことがわかった。競合リスクモデルについてもよく理解できた。
15	専門書を読んでも苦戦していたので、モデルの概要、仮定、解釈が学べて良かったです。
15	打ち切りの種類についてわかりやすく整理されており、理解が深まりました。ほかの解析方法との意味合いの違いやどのようなバイアスが発生するかについて、例をあげていただきよくわかりました。
15	特に、どの統計ソフトで行える解析かがわかりやすく分類しており、ありがたかった。
15	比例ハザード性、競合バイアス、打ち切りのバイアスについて、理解ができていなかったことに気づきました。

15	例題を示した説明で分かりやすかったです。 参考書籍や文献は、最後にまとめられていてよかったですと思いました。
16	LOCF 法、多重代入法のアルゴリズム
16	MAR、多重代入法についてよくわかりました
16	MCAR,MAR,MNAR の分類の説明や単一代入法の具体的な内容が簡潔でわかりやすかった
16	MCAR,MAR,対処法(hot-deck 法・cold-deck 法、回帰代入法、LOCF 法、FCS 法)
16	MNAR, MAR など混合してしまうので、それぞれに対する対処もセットで学ぶことができました。
16	Rubin's rule について具体例を交えた解説がわかりやすかった。
16	なぜ多重代入法が推奨されるのか、多重代入法のしくみについてよく理解できました。
16	メーカー側で臨床試験の実施に携わっていたが、こうした学術的な理解により更に欠測や多重代入法について基礎的な理解を深めることができた。
16	各種代入法のメリット・デメリットが学べてよかった。また、実例を踏まえた説明もわかりやすかった。
16	欠測が起こるメカニズムから、その対処方法まで丁寧にご教授いただき大変勉強になりました。また、論文の方法で多重代入法に関する記述例を紹介していただき大変ありがたいです。
16	欠測が起こるメカニズムから MI まで体系的に学べる機会は多くないため、大変理解しやすかった。
16	欠測値があった場合にどのように対応すべきか、具体的にわかりやすく説明されていてとても勉強になりました。
16	欠損データの扱いは難しいと感じていたので、この講義で実例を見ながら学べて非常に役に立った。アウトカム変数の欠損にも使えるというのは、この講義で初めて知った。
16	欠損値はこれまで分析対象から除外する方法を多くのケースで採用していたが、値を推定し、代入することで対応することができることを今回の講義で学ぶことができました。
16	欠損値は欠損値が存在する限り行われる議論なのではないかと考えています。自分自身では lightgbm や misforest による補完などを使用し対応したことがありました。このような大規模な臨床研究においてどのように対応するかそのことについてしれたのはとても勉強になりました
16	欠損値補完は Database 研究に必要な操作であるため、色々な手法を知ることができた
16	推奨されていない補完方法についても理由付きで解説されていた点が良かったです。
16	多重代入法において、各データセットの分散平均より各データセット間の分散が小さい場合には、代入データセットを増やしても精度が上昇しないことが理解できた
16	代入値算出の理論の詳細まで理解に至っていませんが、欠測メカニズムを特定できない状況では、多重代入法の適応が一番無難であることを認識することができました。
16	代入法が非常によくわかりました。いろいろと複雑に思っていたものが整理されました。ありがとうございました。
16	同種の講義では欠損値のメカニズムの分類を教えて対処法の時間があまり取られないというものが結構多いとおもっています。この講義ではそれらを分かりやすくまとめた後に、実際の対処法について時間を割いて解説いただいているのがとても素晴らしく感じました。ありがとうございます。
16	様々な欠損データの処理方法について、利点も含めて説明されており、しっかりとこの内容について知識を得ることができた
17	”階層構造を考慮しないとサンプルをコピーしているかのような解析になってしまう”という点、このテーマの考え方の重要性をよく理解できました。
17	DB 研究におけるマルチレベル分析の活用を知ることができた。
17	データの階層構造をなぜ考慮する必要があるか、具体例をまじえてご教授いただき大変わかりやすかったです。

17	データの階層構造を意識しながらデータ分析を行うことが必要であること、分析するにはマルチレベル分析を用いることを学ぶことができた。
17	マルチレベルモデルにおける切片項（あるいは傾き項）の意味合いについて理解が深まりました。
17	マルチレベルモデルの種類について、整理していただき勉強になりました。
17	マルチレベル分析が必要になるデザインは多いと思いますが、体系的に学んだことがなかったので、初心者向けかつ網羅的に構成して頂き勉強になりました。
17	マルチレベル分析と混合効果モデルなどの言葉の整理ができた
17	マルチレベル分析の基礎の概要から統計モデルの基本的な内容、実際の研究紹介をしていただき、わかりやすかった。
17	マルチレベル分析をどのような目的で使用するのか、という点を明確にすることができ勉強になりました。
17	階層を考慮した/しない分析を示していただいたことで、分析結果のピットフォールやアルファエラーの増大などがよりイメージしやすかった。また、マルチレベル分析の各種方法が分かりやすく説明されておりよかった。
17	階層を考慮しないと結果が全く変わってしまうことがあることを学べたのが良かった。
17	階層構造について基礎から説明していただき、マルチレベル分析を用いる例を具体的に説明していただいたのでとても理解が深まりました
17	本で読んだり、ネットで調べたりしても、いまいちしくりきいていなかったもので、講義形式を聞いてよかった
17	用語に混乱していたので整理できました。階層構造の解釈が確認できました。
18	SCCS が何なのかという基礎から、その研究の生まれた過程、論文での実例について順序立てて説明されており、学びやすかった
18	SCCS が既存の研究方法と比べて優れている点とそうでない点について、具体例を通して解説していただいたのがわかりやすかったです。また、SCCS を使うための必要条件とその判断方法についても具体的に示していただき、参考になりました。
18	SCCS が必要となるシナリオの説明と導入の背景からしれてよかった。教科書的に聞いたことがあってもよく理解できていなかった為、理解が深まった。
18	SCCS が有用な場合の具体例を学ぶことができよかったです。
18	SCCS で risk period が indefinite とすると検出力が下がる、という点、気づきがありました。またよくある質問について触れていただいたのは非常に有用でした。
18	SCCS では時間非依存性交絡因子がキャンセルされること、これまでの研究デザインと併用することで研究の質を高められること。
18	SCCS について体系だった理解が深まり、よかったです。SCCS を使いたいと考えるモチベーションの部分から説明いただき、納得しました。
18	SCCS の応用について利点と注意点 コホートと SCCS の点推定値の乖離についての意味 実際に解析を行う際の参考文献
18	SCCS はなじみがなかったので、基本について学べて良かったです。
18	SCCS は疫学をやる上で必要知識ですので、ここでしっかりと学べて良かったです。
18	SCCS を用いた具体的な研究をもとに解説していただき大変勉強になりました。また、ある解析手法を用いたいというモチベーションで研究を行うのではなく、自身のリサーチクエスチョンに適した手法を用いるという重要な点を改めて認識することができました。
18	SCCS 自体初めてでしたので概要に触れることができて勉強になりました。
18	どのようなリサーチクエスチョンが向いているか、データベース研究での強みを生かせることがよくわかり、また R でのコマンド例なども提示頂き、よかったです

18	プログラムサンプルがあり大変参考になりました
18	まだあまりなじみがないデザインなので、基本的な理解を得ることができてよかったです。
18	ワクチンに関する研究でこの手法のことが気になっていたため、大変参考になりました。
18	最後によくある質問について触れていただき、自分も同じような疑問を持っていたのでとても良かったです
18	最後の SCCS についての質問とその回答が、今後実際に研究を行う際に役立つと思う。
18	参考になるホームページや教科書の紹介がありよかったですと思いました。
18	自己対照ケースシリーズの基礎を知ることができてよかった。統計ソフトのコマンドのおすすめもあり、その点についてもよかった
18	自己対照ケースシリーズについて、論文で読んでもいまいち理解できていなかったが、とても分かりやすくてよかった。
18	自己対照ケースシリーズについてあまり知らなかったため解析手法の新たな選択肢を学べて良かった
18	自己対照ケースシリーズの利点と欠点が整理されていたため実際に自分が活用する場面をイメージしながら聞くことができた。
19	・機械学習、ディープラーニング、AI 等の概念が自分の中で混乱していましたが、本講義を聴くことにより少し整理することができた。
19	DB 研究における機械学習の活用について知ることができ、今後業務での活用に取り組んでいきたい。
19	Lasso 回帰、XGBoost について言葉は知っていたが、仕組みがわかっていなかったので勉強になった。
19	lasso 回帰などの解説を丁寧にしていただいたので、ハードルが下がった点が良かったです。論文を見てみたいと思いました。
19	あまり慣れ親しんだことがなく苦手意識があった機械学習について、用いられる方法・用語を分かりやすく解説いただき理解が深まりました。
19	かみ砕いて Lasso 回帰や決定木、ランダムフォレスト、XGboost について説明いただいたので、よく理解できました。
19	クロスバリデーションや K-fold 法が、モデル作成用データ内での話だということが学べてよかった。今まで勘違いしていた。
19	とりくみづらい内容だったので、押えておくべきポイントを概観できて、たいへん有用でした。先行研究を踏まえると、今後予後予測スコアに関する研究がすすむのではないかと感じました。
19	ハイパーパラメータの論理的な選び方
19	バギングと勾配ブースティングの違いが分かりやすく、大変勉強になりました。
19	ブラックボックスだった機械学習のメソッドの書かれていない部分が理解できてよかったです
19	まさに対象者にぴったり合致する立場だったので、どの内容も大変参考になりました。これで導入本での学習にも進めそうです。特に予測と推論の目的によって手段である機械学習の使い方が変わってくところや各種機械学習手法の導入的な説明がとてもわかりやすかったです。
19	医療分野での機械学習の利用例を知ることができて勉強になりました。
19	一番興味のある内容で、講義もとても面白かったです。
19	機械学習が医療データに対してどのように使用されているかの例を論文を通じて学ぶことができてよかった
19	機械学習について初学者でありついていけるか心配でしたが、基礎的な部分もご教授いただき、機械学習を学ぶハードルが少し下がったように思います。
19	機械学習のハードルが高く感じられ敬遠していたのですが、本講義のおかげで基礎が理解できました。

19	機械学習の概念についての理解が不十分でしたので、理解を深めるとも貴重な機会となりました。
19	機械学習の基本的概念はもともと理解があったが、データベース研究に適用するためにどのようなことをしているのか、基礎的理解を得ることができた。
19	機械学習の有用性を学びました。正確な法則性を見つけることが可能性をもっと広げることを知りました。
19	機械学習や深層学習など、いくつか資料にあたり、概要は理解しておりましたが、商業的なものや、診断などに用いるもの、といった AI はイメージできるものの、研究への活用のイメージがなく、活用にむけた理解がえられていませんでしたが、そのあたりの理解が進んだように思います
19	機械学習を用いた研究を開始していますが、研究活動における具体例を学習することができ、非常に勉強になりました。また、ISE で λ を決定する方法は知らなかったもので、自分の研究活動にも活用していきたいと思います。
19	機械学習自体の内容を身近なテーマ（臨床疫学）を通じて学ぶことができた点が非常に良かった。
19	数式と文字のバランスがちょうどよく、導入する上で XGBOOST や決定木に関する概要を学ぶことができた。
19	普段 AI/機械学習/deep learning 等や、Lasso/リッジ回帰等の様々な用語を耳にする事がありますが、非常にわかりやすく違い等も含め説明して下さい、理解できました。
20	DB のバリデーション研究の重要性や方法を学ぶことができた。
20	DPC データを用いる場合に ER 受診を対象にすると、既往歴など情報が欠損しやすい
20	RWD におけるバリデーション研究の必要性、重要性
20	これならば、大変だけど自施設でも可能かと感じた。
20	データの妥当性がわからないと、真の結果とは異なる結果が導かれる可能性があることがよく分かった。
20	データベース研究とは聞いたことある程度でしたが、基礎から実際の研究まで丁寧に解説いただきありがとうございます。
20	データベース研究を行っていく上でのバリデーションの必要性についてよくわかりました。特に、妥当性の低さを示すための引用にも用いられていることは初めて知りました。
20	バリデーション研究に非常に興味がありましたので、バリデーションとは何か、バリデーションの研究の種類（目的、方法）なぜ必要なのか、研究の活用例、バリデーション研究の現状等を幅広くごせつめいただき非常に参考になりました。
20	バリデーション研究のノウハウについてよくわかった
20	バリデーション研究の概要と課題について、学ぶことができた。個人情報の観点から現資料に戻って確認を行うことは色々壁があることを改めて実感することができました。
20	バリデーション研究の重要性ならびに大変さがわかりました。感謝申し上げます。バリデーション研究を行われた先生を尊敬しております。
20	バリデーション研究の方法とその重要性が理解できた点。トラブルなど実際に研究を行う中での課題が丁寧に説明されていました。
20	バリデーション研究の方法についてまなべた。
20	バリデーション研究の方法について学べて、とても良かったです。
20	バリデーション研究の方法や課題に加え、やってみる中での具体的な大変だったことの話がとても参考になりました。
20	バリデーション研究をどうしていいかわからなかったもので、その方法を学べたことが良かったです。
20	山名先生のバリデーション研究に救われている研究者がたくさんいます。先生のバリデーション研究に対するありのままの感想を聞くことができ、興味深かったです。

表 2. Q5. 本講義で疑問に感じたこと、不明点（自由記載）（抜粋）

講義番号	Q5
1	「今後追加される」と説明のあったデータの追加がいつ完了するのか知りたい
1	「難病・小児慢性特定疾患 DB」の民間事業者への提供は今後検討されないのか気になりました。
1	・格納されているデータの形式、連結方法や追跡方法が実際に自分がどのように行うかのイメージをするのは難しかったです。
1	・今後の流れとして、複数の DB が独立して作成され(て研究者自身が結合す)るのか、あるいは 1 つの大きな DB を作成するのか、いずれでしょうか。 ・时期的に新しいデータが多い印象です。全体的に 2010 年以前のデータは今後も入手が難しいのでしょうか。
1	1 個人についてどのようなアウトカムがどの程度の期間で追跡可能か例を挙げていただければより良い。
1	CRO などの企業は、公的データベースをどれくらい現状利用可能であるのか、今後変わっていくのかなどが気になりました。
1	Data linkage 機能がより洗練されていくことを希望します。急患、救急車の活用 Data.
1	DB 利用に掛かる費用について、もう少し具体的な例を聞きたかった。
1	DPC の個票データを申請してみたいと感じました。NDB などとの連結申し出を行わない場合の Pitfall について、知りたいと思いました。
1	HIC は具体的にどれくらいの期間で施行的利用され、いつ頃から本格利用されるのを知りたかったです。
1	ID1-5 に関しての具体例など詳細がもう少し知りたかった
1	ID4, ID5 などを使うと、その精確性はどのくらいでしょうか？
1	ID については技術的に難しいと感じました。
1	ID の連結が大変そうな印象を持ちました。また、データ構造が複雑とのことでしたが、データ構造を変更したり、データクリーニングを自ら行うことになるかと思いますが、その際の留意事項等について、知りたいと思いました。
1	LIFE についてももう少し詳しくお聞きしたいと思いました。
1	NDB データベース利用が年 50 件ほどであるということについて、今後の利用拡大に向けた国の取り組みはいかがでしょうか。NDB 申請から利用までに時間を要することは以前より問題となっていたかと思われます。
1	NDB における「慎重な審査」に関する具体的な事例をお伺いしたかったです。
1	NDB のオンサイトセンターで使用できる統計ソフトの中に SAS が含まれていないのはなぜでしょうか
1	NDB の使用までに時間を要する理由
1	NDB の担当はかなり少ないと思います。今後検討いただける事項が多いために感じますが、HIC の本格運用時期の目安が分かると嬉しいです。
1	NDB は利用のハードルが高いと聞いたことがあるが、今後どのように改善されるのかが気になりました
1	NDB を個人で利用する割合が増えているのかお聞きしたい
1	NDB を使用した経験はないのですが、今後、人口動態調査票に基づく死亡情報が収録されるというのは非常に有益だと思います。これらの情報が十分に集積されて妥当な研究を実施できるようになるのは、まだ数年先という理解で宜しいでしょうか。
1	NDB 使用にかかる審査について、手続きの具体的な流れについてご教示いただきましたら幸いです。
1	NDB 利用における課題や現在の取り組みについて、スライドにて図示があると、より理解が進むと感じました。

1	NDV データベースを利用するにあたり各施設で具体的にどのような苦勞があるのか、可能であれば知れると、より現実味が湧いたと思います。
1	オンサイトセンターで SAS は使用できないのでしょうか。SAS のインストール媒体をオンサイトセンターに持参する等。
1	オンサイトリサーチセンターの利用について興味があるので、もう少し詳細に聞きたかった。
1	クラウド環境での利用におけるセキュリティ
1	クリニカルクエスチョンに応じてどのように使用するデータベースを決定するのがよいか。
1	それぞれの DB を購入し使用する団体の種別や目的が気になった
1	データの具体的なかたちがイメージできないため、ダミーデータでもよいのでお示しいただきましたら幸いです。
1	データベースごとの、データ提供までの期間はどの程度かが気になった。
1	データベースにアクセスするためのハードルがあり、なかなか臨床等をしながらでは難しく、臨床疫学者との連携が必要と感じた。
1	データベースに入力する項目が今後さらに増えていくことが予想されるが、それに伴い、現場の医師の負担が増えることが懸念される。この懸念点に関しても並行してなにか制度が進んでいるのか、疑問に感じた。
1	データを取る際の具体的な金額
1	データ抽出や連結の複雑なところや限界がわかったが、逆にどのようなデータなら抽出可能で、どのような事ができるのかが、まだイメージしにくかった。
1	データ分析用のデータ取得には、まだハードルが高いと思った
1	データ利用申請を行う上での落とし穴や対処法などを可能な範囲で知りたいと思います。
1	なぜ企業での利用が増えづらいのか
1	マイナンバーカードとの関連性について知りたかった。
1	もっと手軽にデータベースが利用できるように法整備が進めばいいのと思います。
1	違反事項に関し、実際の過去の事例（どんな不備があり、どんな罰則になったか）を教えていただければ幸いです。
1	医療介護データ等の解析基盤について、いつ頃運用開始になるのかももう少し具体的に知りたかったです
1	一つの研究で複数の DB を利用する（利用することを検討している）場合、それぞれの DB の特徴や注意点があれば、比較しやすいと思いました。
1	介護 DB の ID は医療系 DB の ID（2020 年 10 月時点の被保険者番号）と同一なのでしょうか マイナンバーの利活用について構想や計画はあるのでしょうか
1	各 DB から抽出できる項目（変数）の粒度。例えば、介護 DB であれば、サービスの種類、利用回数や日数、利用開始日など、どこまで可能なか等。
1	各 DB について、具体的なデータテーブルについてもう少し説明して下さると理解しやすかった気がする。
1	各 DB の将来的な発展をもう少し知りたい。
1	各データベースについて、マスタはどのような種類のものが提供されているか。
1	感染症 DB の内容は感染症法の発生届に準じたのでしょうか。
1	含まれている内容、含まれない内容の具体例をもう少し教えていただきたかったです。
1	企業が公的データベースを利用する際の留意点が知りたかった。
1	現実的に、申請、審査から利用までに大変な時間がかかる上、連結する DB 毎に申請、審査が必要となると、制限付きの研究費を用いた研究では利用できないように感じています。統合運用に向けた法整備が望まれます。

1	公的データベースを扱うためには、想像以上に制約があると感じました。利用者はどの様にしてこのハードルを乗り越えているのかも知りたいところです。
1	公的データ使用のハードルがどの程度高いものか、具体的な例があるとよいと思った。
1	講義のスコープとは異なると思いますが、相当の公益性、という条件について、利用が進むためには具体的な要件が示されていくことが望ましいと考えた。
1	今後、他のデータベースとのリンケージ時期
1	今後の他の DB との連結予定について、大体いつ頃を目指しているのか
1	今後の動向がメインであり、定義書も公開くださっていますが、各データベースの限界もぜひ教えていただけますと幸いです。特に DPC データは年度ごとに入力割合や入カルールも変更になる場合があると認識しています。
1	今後解析基盤が整備される予定ということであるが、例えば、利用申請などの手続きについては、どのくらいのハードルがあるものか。
1	使ってみたいと思えるほど敷居が低くはなかったです。
1	私共製薬メーカーがいつか使えるようにはなるのでしょうか。
1	次世代 DB との連結で期待している事、ビジョンについて知りたかったです。
1	将来的な個人番号との連結について
1	申し出がされて提供までに時間がかかる理由をもっと具体的に知りたかった。 (申請時の注意点やよくある指摘事項など)
1	申請から入手までの期間や利用費等の具体的な情報
1	第三者提供の機関/企業の内訳やテーマの実際の動向はどんな傾向にあるのか
1	転記情報について、人口動態調査票との連携について、2023 年以前は連結できるのか、よくわからなかった
1	難病・小児慢性疾患も民間活用への制限が解除されるような動向はありますでしょうか。
1	民間における利用において「公益性がある」と判断された研究の例などが知りたい
1	利用に際し、専門委員会で承諾されなかった場合、拒否された理由は通知されるのか、再申請は可能かについて疑問に感じました。
2	NDB ガイドラインの内容は、Open NDB でも、特に、結果の公表や集計方法に関して、反映させる必要があるのか
2	データ提供まで 300 日かかるとのことですが、その理由や、短縮するにはどうすればいいかを知りたかったです。
2	最新のガイドラインのリンクを提示していただくとありがたかったです。
2	参考にすべき NDB ガイドラインがどこにあるかがわかりにくかった。 URL があると良かった。
2	自力で解答にたどり着けない項目も多かったので、できれば動画での講義→e-learning でのテストという流れだと吸収しやすい気がした。
2	提供申出の承諾後から NDB データが提供されるまでかなりの期間かかるが、その理由は何であるのか疑問に感じた。
2	利用申請に時間がかかるのは、リソースの問題になるのでしょうか。あるいは、確認事項が多いことが原因になるのでしょうか
3	「SELECT 文③：条件」というタイトルのスライドの右下図は年齢の列も表示されているが、ID と性別だけ表示されるのか、それとも WHERE を入れると年齢も表示されてしまうのか、わからなかった。
3	ID に基づくリレーショナルデータベース構造の具体図が、全ての項目に ID が記載されており少し難しかったです。

3	INNERJOIN と OUTERJOIN (LEFT/RIGHT/FULL)の説明をされていた、空集合の図ですが、LEFT と RIGHT の図が私の認識と異なっていました。
3	LEFT JOIN, RIGHT JOIN の説明のベン図が間違っているように思います。 共通項の ID の部分も含まれるはずです。
3	MySQL や PostgreSQL などの中で、医療データベース研究に用いるにはどれがいいのか。
3	Null を 0 にした際に、0 歳の患者が対象となる場合不都合があるのか少し気になりました。
3	PC の性能によるとは思いますが、どれくらいの n 数の研究から SQL を使い始めたほうがいいでしょうか？
3	SELECT DISTINCT [ID]を実行すると、抽出されるのが[ID]の列だけになってしまいます。ID の重複を除いた上で、他の必要な列も抽出する方法が知りたいです。
3	SQLserver のインストール、設定、NDB データを取り込むテーブル作成の部分
3	SQL の利点について、データ処理量の話はあったと思います。最終的には python との連携も視野に入ってくると思いますので、そこに向けた注意点などあればより解説して頂けたらと感じました。
3	SQL は、条件に合ったデータを集計するのに使用するのわかりました。 実際の研究に使用する場合どのように使用すればよいのか？ 例を示して、いただければもっとよかったですと思いました。
3	SQL をさらに学ぶ際、おすすめの書籍などがあれば教えて頂きたい。
3	SQL をどのようにダウンロードして立ち上げて動かすかが知りたいです(MAC)。
3	SQL 上だと何かエラーが生じている時に気付く術はあるのかどうか少し気になりました。また、論文中に SAS や R を用いた点はよく記載されていますが、SQL を使っている点は触れられていないのがほとんどのように思います。基本は前処理としてのみ使用する SQL については記載不要ということで良いか気になりました。
3	結合時にデータ量が膨大になるのを防ぐために、複数の変数を一致させて結合を行うことを一つの解決策として、示されていました。複数の変数を一致させて結合する際に起きうる問題点について、さらに詳しく知ることができればよかったですと感じました。
4	NDB の構造の excel で、再表示すると隠れた table 名が出てくることを初めて知りました。このことは、NDB の HP のどこかに書いてあるのでしょうか。
4	SQL でデータ整理を行い、その後他のツールで解析を進めていくことになるかと思います。SQL で整理したデータはどのような形でエクスポートされるのでしょうか。また、R 等でもデータの結合など様々データ処理を行うことができるとは思いますが、どこまで SQL でデータ処理を行うものなのでしょうか。
4	SQL のおすすめテキストや Web ページ等もご紹介いただけるとよかったです。SQL は情報があふれすぎて逆に迷子になりがちなので、入門者向け、レファレンス用などに分けてご紹介いただけるとたいへん助かると思います。
4	スタイルガイドなどの説明があった方が良いように思いました。 クエリを書く環境構築についても使いやすいツールなどの説明があると助かるかと思いました (VScode etc.)。 (いずれも説明を見逃していたらご容赦ください)
4	演習問題が盛りだくさんで 焦りました。講義演習の予測完了時間を示して頂けると 自分がどの位置にいるのか目安になるとは思います、もしくは、予測時間に合わせて問題を解く時間の予定が組めそうです。
4	有料でもいいのでぜひレセプトハンドリングについて e-learning 講座を開催してほしいです。
5	・先行 NDB 研究レビューにて、臨床医学、薬理学の分野について件数の記載がありましたが、データとして活用できるにもかかわらず活用が不足している分野はどこなのか疑問に感じました。社会医学、理学療法や看護といった臨床医学以外の分野の活用はなかったのでしょうか。 ・抽出対象とした NDB、NDB Open Data、JMDC、DPC の違いについてより具体的に知りたいと思いました。 ・研究紹介にありました口腔ケアの論文について、周術期口腔機能管理料のデータ格納の状況について関心を抱きました。
5	NDB での死亡情報は 100%捕捉できているわけではないと聞いたことがありますが、NDB 先行研究①の死亡の定義方法が気になりました。

5	NDBにおける病名の精度について（例 肺炎など）。
5	NDBの研究分野において精神科が2位だったので、その研究詳細を知りたかった。
5	NDBを用いた研究を行うにあたって、かかった費用と時間を具体的に知りたいと思いました。
5	どのような研究デザインは不向きなのかという limitation について少し知りたかった
5	なぜ NDB が他 DB より論文数が少ないのかを検討すべきだと思いました。
5	リアルワールドデータでの研究は増加傾向にあるが、一方で研究が進んでいない分野もあるとの事ですが、その具体的な要因は当該分野での研究者が少ない事が要因でしょうか？
5	リアルワールドデータ研究のレビューにおいて、各データベースについて、研究分野別の件数のほか、疾患領域別の件数も教えていただきたいです（考察で少し述べられていましたが）。また、研究紹介については、解析手法のリミテーションなども気になったので、論文を拝読したいと思います。
5	一人の研究者だけで NDB 研究に取り組むのは難しそうで、こういったチーム（人的体制）で臨むのが効率的かなど
5	研究が盛んな領域と盛んでない領域の違いは何でしょうか。データベースのリミテーションに依存するのでしょうか。
5	研究分野で活用が盛んでない部分はデータ内容による要因なのか、その分野ではデータベースの活用よりも基礎研究のような実験が主なのかが疑問に感じました。
5	今後のセキュリティ等の制約の変更などにより NDB の研究にどれぐらいの影響が出てくるか、増加に影響するかどうかも知りたいと思います。
5	論文数が増えていることはわかりましたが、その質（例えば IF）については言及されなかったので、疑問に思いました。
6	1:4 の症例対照マッチングについて、対照を何度も繰り返し使用可能な場合についてもう少し解説してほしいです。
6	DPC データベースから薬剤の処方情報を取得することはできませんが、服薬コンプライアンスについては把握できないと思われます。薬剤の効果について解析する際、コンプライアンスが考慮できない点はどのようにケアするのでしょうか。（limitation として考えるのみなのでしょうか）
6	研究の限界以外の困難な点や他にどんな研究に応用ができるか結果からこういった意思決定をされているのか
6	研究を実施するための幅広い分野の人材確保にあたっての課題などはないのか（必要な人材が思うように集められる/参加・協力してもらえるのか）
6	公的な DPC データと商用の DPC データでできることや結果の解釈に違いはあるか気になりました。
6	講義内でも限界点として説明がございましたが、別病院で受診された場合は別症例として区別されることは同一症例が複数カウントされることになり、こういった同一症例の重複がどの程度発生しているかは評価可能なのでしょうか。
7	DPC データには外来データも含まれていると思いますが、外来データを利用した研究はおこなわれない、または研究には不向きということでしょうか？
7	DPC で入院費用がわかるとのことですが、それは包括ではなくて全てに使用した医療費、ということになりますでしょうか。
7	Exposure や Outcome の定義、背景因子の選択理由、欠損値の扱いに関してより詳細な解説が欲しかった。結論に影響を及ぼしうる Limitation があっても研究成果の論文文化は可能だということがわかりましたが、これらの研究成果は臨床的意義はどのように評価されているのでしょうか？臨床現場や政策の意思決定の根拠とされるようになるのでしょうか？
7	PS マッチングにおいて、1:4 や 1:2 などの使い分けは、研究デザインとして決めたのか、実際にマッチングする例が少なかったからなのか、知りたい

7	具体的な研究内容の紹介に出てくる研究の限界で、不足している情報(統合失調症の乳がん手術の研究の場合、統合失調症の重症度や、術前化学療法)がどの程度その研究の分析に影響してくるのかを事前にどう考慮すべきなのかが難しいと感じました。
7	信頼性を高めたり、限界に対して考察を得るために様々な方法を適用されたかと思いますが、査読対応ではどのような対応が生じたか興味を持ちました。このような様々なアプローチを適用できる引き出しを持つことの重要性を認識しました。
7	統計的な有意と臨床的な意味についてより具体的に伺いたいです
8	DeSC データにかかる費用がどれくらいかに興味があります
8	DeSC データの入手手続き、コスト等
8	DeSC データベースと国民健康栄養調査との有病率を用いた比較において、どの程度の差であれば、許容範囲であるのか、もう少し詳しく知ることができれば嬉しかったです。
8	DeSC データベースは今後どの程度の規模の症例数になる見込みでしょうか。
8	DPC データと DeSC データはどのような研究で使い分けるのか？
8	JMDC と同じように台帳情報はあるのでしょうか。
8	NDB と DeSC を比較した場合のメリット、デメリットをもう少し知りたかったです。
8	NDB の利用ハードルが非常に高い現状で DeSC は次善策として非常に魅力的だと思いましたが、n 数以外にはほかにどのような欠点、弱点があるのかをもう少し知りたかったです。
8	NDB 利用のハードルがたかく、DeSC で傾向をつかむような利用のされ方が今後増えていくのでしょうか？
8	後期高齢者へ移行する年齢帯においては、フォローアップができなくなるように思われます (ID のリンケージはできないため、後期高齢者医療保険で見ているレセプトは初発診断ではない可能性が高いように思います)。そのあたり、高齢者を対象とした研究をしたい際にどのように考え、扱うべきか疑問でした。
9	DeSC データを用いた研究セッションを見た後にこのセッションを見たため、前半パートや、研究紹介 1、3 など、DeSC データでも同様の内容や結果が言える(のではないか)と思う点が多々合った。JMDC データを選択したからこそできた(分かった)研究、のような内容が知りたかった。
9	JMDC データの実際取り扱い時の注意点、データクリーニング時の pit fall など practical な部分も教えて頂ければよりよかったです。
9	JMDC データ利用に不向きなものについての内容があったが、実際にやろうとしたが、やってみたら出来なかった (適さなかった) 研究などがあれば知りたいと思いました。
9	JMDC は比較的大企業の健保を組み入れているという事で、SES の高い集団という selection bias と傷病名の妥当性についての欠点を研究時にどのように考慮すべきか、妥当な結果が得られないかもしれないという疑問が浮かんだ。
9	データを申請する際にある程度リサーチクエスチョンが成立しているかとは思っていますが、実際にデータを受け取ってみると立案時には予期しないことなども生じうると思います。どのように対応されていますか？
9	家族 ID が入手できる点について、どのような研究に生かせるのか知りたい
9	各データベースの特徴をお聞きすると、各データベースの特徴を意識しながら、研究に最も適したデータベースをいかに選定するのか、何を重視してデータベース選定をするのか等のデータベース選定の重要性を感じました。
9	保険から逸脱しない限り追跡可能なのに、観察機関が短いという限界について疑問に感じました。
10	「PS を用いる際に、過剰適合や多重共線性を考慮しなくてよい」「傾向スコアの算出に用いなかった別の交絡因子でさらに調整する方法は適切でない」とご説明がありましたが、その理由についての解説、もしくは文献の紹介があれば、より深い学習につながったと思います。
10	・逆確率重み付けとオーバーラップの使い分けはどのようにするのでしょうか。 ・ATE と ATT の用語の違いはわかりましたが、結果の解釈でどのように異なるのかが知りたいです。

10	ATE と ATT の違いをもう少し説明していただきたかったです
10	overlap による重み付け、感度分析についてももう少し詳細な説明が欲しかった。
10	PSM 後に SD0.1 以上になり、その後に工夫して揃えたケーススタディが知りたい
10	スライド 26 の重みづけの話とセットでロバスト分散の話も入れて欲しいと思いました。
10	データベースで重要な背景因子が収集できない場合、どのような対処が可能でしょうか。
10	どうして逆確率やオーバーラップ重み付けを行うのか、"逆確率"や"オーバーラップ"で重み付ける意義のようなところがわからなかったです。
10	なぜ、標準偏差×0.2 が閾値として用いられることが多いのでしょうか？
10	傾向スコアとロジスティック回帰のどちらを選択するか迷った際の判断基準などあるのでしょうか？
10	傾向スコアによる調整前の分布の評価のために C 統計量を測定することを推奨していますが、一方で本法の使用を推奨されない、不要であるという意見も聞いたことがあります。査読などでこれが求められることは多いのでしょうか？
10	抽出方法で、復元と非復元方法を選択する際には、どのような点で違いがあるのでしょうか？（講義では 1 対多数の場合などは復元を利用している例を挙げていただいていた）
10	二群のバランスをとるのに難渋する経験があるのですが、どうしても SD が 0.1 を上回る（例えば 0.2 など）際には、それでよしとするのか、なんとしてでも 0.1 を下回るように工夫するのでしょうか。
10	論文や発表で示す際の注意点に関する質問です。マッチング後に群間のバランスを整える際、二乗項やカテゴリ変数に変えても良いとありましたが、それらを選んだ理由やバランスが取れるまでの過程は示すべきでしょうか？
11	「高次元傾向スコアのコンセプト」未測定交絡を間接的に調整することと、高次元傾向スコアの分析方法が結びつかなかった。今回ご紹介いただいたような統計的変数選択をすることで、本当に未測定交絡の間接的に調整につながるのか、未測定交絡因子と測定できる因子の関連に関する情報は反映しなくていいのか、がわからなかった
11	hdPS の算出アルゴリズムもさまざまな提唱があるとのことで、なかなかお手軽に手を出せなさそうな手法である印象を持っていました。自身の臨床疑問を解決するための研究としてどこまでマッチし得るものか、計画がとても重要だと感じます。引き続き勉強させていただきます。ありがとうございました。
11	hd-PS を行うに統計ソフトは何がおすすめでしょうか
11	あまり先行研究が多くはない理由は何なのでしょう
11	感度分析として高次元傾向スコアを用いた解析する、というケースが多いのか、はじめから高次元傾向スコアのみを用いた解析をするのか、教えていただきたかったです。
11	欠点に計算リソースが必要という点もあるように感じました。
11	高次元傾向スコアの場合も、マッチングや重みづけ後に SMD など用いてバランスを確認するのか？
11	高次元傾向スコアの場合も傾向スコアと同様に C 統計量が 0.6-0.9 の間の値でないと適切ではない理解で間違いないか疑問に感じております。
11	先行研究②で推定に用いていない変数についてもバランスが取れていたが、この現象はいつでも起こるものなのか、この研究ではたまたまなったのか、これは毎回確認した方がよいことなのかということが疑問に感じました。
11	予め投入すべき変数がいくつか挙げられていましたが、それらが選択された理由（年齢などは言わずもがなだと思います）を知りたかったです。
11	臨床的に意味のない変数が含まれたときに、入れる場合と入れない場合などの、感度分析を行ったりするものなのでしょうか？

12	2 SLS で C (背景因子) が①式と③式の両方で用いられる理由が理解できませんでした 2 SRI で予測値そのものではなく、予測値の残差を用いる理由が理解できませんでした。
12	CACE の解釈 (何の効果をみているのか) が難しいと思いました。
12	NDB や商用データベースで操作変数となるようなデータがあるのでしょうか。またバイアスの問題もあり利用は難しいように思いました。またどの程度論文で使用されているのでしょうか
12	propensity score analysis と結果が乖離した場合、どちらがよりバイアスが少ないのかを判断することは難しいように思いますが、そうすると、propensity score analysis と結果が同様だった時のみ報告されるというバイアスにつながるように思いました。
12	この手法では、操作変数として何をを用いるかが非常に重要と思います。適切な操作変数を見るけるために、どのような対処をされているのでしょうか。
12	傾向スコア分析と操作変数法を用いた分析を併用する場合、どちらを主解析にしたらよいですか？
12	最後のいくつかの実例において、各操作変数が本当に「アウトカムと直接関連していないか」「アウトカムと共通原因をもたないか」の仮定を満たしているか怪しいと思いました。どの程度まで受け入れられるのか、定量的な評価は難しいと思いますが、なにか指標やガイドラインのようなものがあればご紹介いただければさらによかったです。
12	商用化されたデータベースを用いた研究の場合、操作変数を見つけることが困難で、操作変数法を適用するのは難しいのでしょうか。
12	操作変数をどのように見つけるのか。難しいなかでどのように行っているのか興味を持ちました。
13	Fuzzy RD と Sharp RD をどこで区別してよいのか、十分に理解できませんでした。 もう少し、それぞれに合致する具体例をご説明いただければ嬉しかったです。
13	ガイドラインの改訂によるアウトカムへの効果を見る場合、不連続回帰と差の差分析どちらがよいのでしょうか
13	この分析方法を利用した CQ を探すことも難しいと感じた。
13	どの程度の時系列でトレンドを確認した場合、並行トレンドが仮定できたといえるのでしょうか。
13	仮定を満たしているか確認が難しいとのことですが、不適切な解析を行った具体例などがあればわかりやすいと思いました。
13	前提となる仮定を満たしているかの検証を具体的にどうしていくべきか、また論文にする際にどこまでそれについて詳細を記述すべきか、についてもさらにお聞きできたらうれしく思いました。
13	前提条件、仮定が当てはまりづらい場合にどこまでこの分析が使えるのか。
13	適応するケースがまだ腹落ちできていない
13	臨床的な研究の具体例も挙げていただけると今後の研究に活かせるため助かります
14	exchangability などは勉強が必要だと感じた
14	HIV 研究の紹介において調整後ハザード比が後の RCT で正しいことが証明されたとのことがありました。 その文献番号も紹介していただけるともっとよかったですと思いました。
14	weight 作成時の P0, P1,....を実際どうやって求めるのかわかりませんでした。
14	どこまで時間依存性を考慮するか
14	観察研究の中では同じタイミングで来院や検査や行われるわけではないので、時間依存性交絡の欠損も多いかと思えます。そういった場合はどのように取り扱うのが普通なのでしょうか。一般的に行われる代入法で適切な結果が得られるものなのでしょうか。
14	逆確率重みづけ法によって作成された重みの分布は、どのような分布であれば良いのか、Truncate が必要になる場合はどんな分布であるのかが不明だったため、具体的なグラフなどがあるとありがたいです。
14	欠測がある場合に MICE で補完しつつ、MSM-IPW を適用することは可能ですか？

14	合流点層別バイアスが生じる、という点が理解できなかった。CD4 数 1 は CD4 数 0 と Zidovudine 0 の Collider であるが、死亡には直接関係していないのに、なぜバイアスになるのかが理解できなかった。
14	周辺構造モデルのステップは理解できたが、どのような計算をしているかがなかなかイメージできなかった。
15	composite outcome にして通常のハザード解析をするのと、結果の解釈がどう変わるのか
15	競合リスクに対する対処法で、統計的に対処するのと、複合エンドポイントにするのでは、どのような使い分け、選択をすればいいのか。
15	競合リスクモデルの際にハザード比の推定の方法の選択について復習を繰り返したいと感じました。
15	原因別ハザードと部分分布ハザードの違いがわかりにくかった。
15	死亡が追跡中に発生するがん領域の安全性評価で製薬後 DB 調査を行う場合、死亡は打ち切りではなく、競合リスクとして扱うべきでしょうか。死亡による打ち切りが特段他の打ち切りとリスクが変わらない場合はバイアスなしと想定して、シンプルに打ち切りとしてハンドリングすればいいのでしょうか。ご紹介いただいた成書で詳細を確認しておきたいと思います。ご説明ありがとうございます。
16	・ 血圧のみが欠損の場合はイメージがわかりましたが、血圧、腹囲、ヘモグロビン A1c のように複数の欠損がある場合、それぞれに対して 20 セット以上のデータセットを作成するのでしょうか？ (A さんは血圧のみ、B さんは血圧と腹囲、C さんは 3 つとも、の場合) ・ 欠損値の取り扱い、解析方法は、研究デザイン立案時に決めておくべきでしょうか？
16	・ 多重代入法で使用するデータセット数はどのように決定するのでしょうか。 ・ データを解析する際、様々な変数で欠測が発生すると思いますが、そのすべてをこの多重代入法で保管するという理解で合っていますでしょうか。
16	hot-deck 法が可能な状況であれば cold-deck 法を用いる必要はないということで良いのか疑問に感じております。
16	Imputation は何%以上の欠損があるとやったほうが良いということはあるのでしょうか。また解釈として差があったものが Imputation で差がなくなるのは欠損の影響はありそうだとありますが、差がなかったものが差がでてしまうと解釈が難しいように思えますがいかがでしょうか。
16	LOCF は不適切な代入法なのか、議論がもう少しあると良かった
16	PS を求めるために使用する変数において欠損値があった場合はどうすればよいでしょうか。
16	Rubin's rule のところがもう一つ理解ができなかったです。
16	多重代入法でデータセットをいくつ作成するのがよいのかよくわからなかった
16	代入のメカニズムとして 3 つあるとの事で、今回は FCS 法を主に紹介して頂きましたが、どのような時に FCS 法を用いるのが良いか、もう少しゆっくり説明頂けるとありがたかったです。
16	単一代入法が推奨されないとのことですが、眼科の抗 VEGF 薬の RCT で LOCF 法が用いられる場合が多いのは何故か疑問に感じました。
16	日本の DPC data を使うとき、特に様式 1 は欠測データがよく起こりますが、今までの事例ではよく多重代入法を使いましたでしょうか？
16	補助変数を見つけるのが大変そうですが、臨床的感覚で判断できるものなのでしょうか。
17	どういったときに、マルチレベル分析を考慮したほうがいいのか、考える指針があるとよい（講義の中で言われていて、聞き逃していたらすみません）。病院や最初の障害の程度が違うなどはとても想像がつくのですが。
17	どのような研究が適しているのか、はたまた例が Multilevel を考慮しないとどうなったのかが知りたかったです
17	マルチレベルを用いた研究の例ではランダム切片モデルのみの紹介だったため、ランダム切片・傾きモデルの例やその使い分けについて疑問に思った。
17	ランダム切片モデル、ランダム傾きモデル、ランダム切片・傾きモデルはそれぞれどのように選択すれば良いのでしょうか。
17	一般化推定方程式との使い分けについてもう少し理解を深めたかったです。

17	社会疫学の利用例も紹介していただけるとよかった
18	・ Risk period の設定の基準は？（研究者の判断？） ・ 残差交絡
18	・ 適切なリスクピリオドの設定方法。一旦、「因果関係なし」という結果が出ても、リスクピリオドを変更したら「因果関係あり」となったりすることはないのでしょうか？ ・ コホート研究と SCCS の結果が食い違った先行研究事例でコホート研究に問題があった（残差交絡が起きていた）可能性があると考えられていましたが、逆に SCCS に問題があったとされたケースは無いのでしょうか？
18	4つの必要条件それぞれを満たさない場合に起こりうる問題点
18	Exposureにより Risk period の設定方法が難しいと思いました。また、絶対リスクが分からないことだけがデメリットなのか、気になりました。4つの assumptions を満たさないといけないことが厳しいので、コホート研究の方がよく用いられるのでしょうか。
18	コホート研究と SCCS を用いた研究を両方行った場合、評価指標が異なる可能性があると思いました。例えば、コホート研究はハザード比、SCCS は発現率比、このような場合、結果が逆転していないかを確認するといった理解でよろしいでしょうか（点推定値や信頼区間の値自体に比較可能性がないと考えてよろしいでしょうか）。
19	Lasso と XGBoost で両方とも高い予測精度と判断された場合の結論づけが難しいと思いました。少ない変数で予測できる Lasso が実装には適していると結論づけるか、今回のデータセットでは違いが無かったが、より多様なデータの場合 XGBoost の方が予測性能が良い可能性があるとするのか、実装で使用するべきモデルの選択についてはまだよく分かっておりません。
19	Lasso 回帰のλの設定方法についてもう1歩ご説明いただけると助かります。
19	Lasso 回帰のキーワードは理解できたが、Lasso 回帰自体が何なのかは分からなかった。
19	xgboost hyperparameter の設定を知りたかったです。
19	ブートストラップ法と K-fold 法の使い分けについて
19	モデル作成用データと検証用データの比率はどのようにして決めるのでしょうか。
19	ランダムフォレストは何組作ればよいのか。
19	ロジスティック回帰の式は、 $\text{logit}(y)=\ln(y/(1-y))=X\beta$ でしょうか
19	過学習との境目はどのように判断するのでしょうか
19	作成した予測モデルでその後事業につながった例など具体的な活用方法が気になった
19	正則化やハイパーパラメータが意味するところがイメージしにくかった。Lasso 回帰、Ridge 回帰、Elastic net がどのように使い分けられるのか疑問に思った。
19	良好な予測能を示したモデルについて、臨床現場等で実際に使用するまでのハードルはどのようなものがあるのでしょうか？
20	DB は匿名化されていることや、独立した2人でカルテレ뷰を行うとすれば、制約の大きい環境の中でどのように実現にもっていくかが問題のように思いました
20	DB 研究を行う際は必ずバリデーション研究を合わせ行わないといけないのか、バリデーション研究にどの程度汎用性があり、既存のバリデーション研究を活用することがどの程度現実的なのか、疑問に思いました。
20	DPC レセプトではない情報のバリデーションがどのように行われるか。どこまで行えるか。
20	PPV もしくは NPV どちらを優先すべきか、指針などがあればお教えいただきたいです。
20	validation 研究における妥当性の指標は感度特異度一択という感じなのでしょうか。
20	validation 研究を行った時に苦労した点で、再度病院を訪れられない点を挙げていたが、具体的にどのような理由で再度病院を訪れたいと思ったのか。カルテレ뷰が、予定の数よりも少なかったためか。

20	カルテレレビューまでできない時に、先行研究との比較で許容される条件
20	カルテレレビューをする人によって、かなり効率も変わるし、クオリティがバラつく事があるのではないかと疑問に感じた。
20	カルテ上に病名が出現して、それが間違いなこともあったりするのでしょうか？
20	バリデーション研究、特に、傷病名等に関しては、難病や難病指定に関するものを実施してしまうと、患者様への迷惑や難病指定を取るために、医師側の過大申告等も懸念され、対象とする研究を良く考慮しないと難しいと思っております。
20	バリデーション研究がその対象となったデータベースからの外挿可能性に乏しい、という議論があるのかなと思いました。
20	バリデーション研究が実施されていないDBの利用方法が分からなくなった
20	バリデーション研究について正しい共通認識をより多くの人が持てばより研究が進むこともあると感じましたが何か取り組みなどあるのでしょうか
20	バリデーション研究にはお金がかかりそうですが、研究費はどこから獲得することが多いのでしょうか。
20	バリデーション研究の実施は大変なことを再認識できました。一方で、先人の取り組んだ貴重な研究結果を共有いただくような取り組み（バリデーションが取れたアウトカムのレポジトリ）等があれば、ご教示いただけたら幸いです。
20	バリデーション研究の重要性について、まだまだ周囲の関係者の理解が得られにくいので、引き続きこのようなセミナーをはじめ啓発・教育が必要のように思っております。
20	バリデーション研究は評価されにくい分野であることは理解できました。それゆえに、どうしたら評価されやすくなるのかについて、より深ぼりして頂けたらよかったですと感じます
20	バリデーション研究を行わなくてもDB研究はできてしまうため、（Limitationに明記されていたとしても）信頼性の低い論文が量産されてしまうのが現状のように感じました。 製薬メーカーのように、データの所有者ではない組織ではバリデーション研究を実施するのはさらに困難です。どうすればよいのでしょうか？
20	医師間の診断のばらつきの影響はあるのでしょうか。また、影響がある場合、影響度が大きい対象疾患はどの領域でしょうか。
20	一施設だと入力における考え方の違いからバイアスが大きいと思いますが何施設くらいあれば十分でしょうか
20	一般的に感度・特異度がどの程度なら査読者から許容されるか、目安があれば教えて欲しい。
20	感度、特異度、陽性的中率、陰性的中率。例えば陽性的中率だけしか示していないバリデーション研究もあり、こういった場合にどの指標で妥当性を示すべきか
20	感度、特異度、陽性的中率、陰性的中率がどの程度であれば適切か、こういった影響があるか。問題となる事例はほかにどのようなものがあるか。
20	感度・特異度が分からないデータベースは活用しない方がいいのか。
20	現在は、山名先生のように、研究者が苦勞してバリデーション研究を行うケースが多いと思いますが、提供するデータベースの会社が実施すべきではないかと考えています。 また、研究者が3Kで苦勞しながら、バリデーション研究を行う事を考えると、評価指標のような著作権制度を導入してもいいのではないかと考えています。
20	講義中紹介されていたように、バリデーションの重要性は高いものの、実施はハードルが高い状況かと思えます。一方で、今後NDB等を活用していくためにはバリデーション研究の推進が欠かせないと思うのですが、公的な研究実施計画等あるのでしょうか
20	国を挙げてバリエーション研究を行い、色々な人が活用できる仕組みがあれば詳しく知りたいと思いました。
20	次世代医療基盤法、医療DXの進展により、バリデーション研究はより容易になり、多くの研究が行われるようになるのでしょうか。

20	製販後 DB 研究実施のために、企業が積極的に実施しようと考えるとなかなか難しい面もあるのかと感じました。
20	先行研究の事例の通り、必ずしもバリデーション研究で得られた結果がちっとも valid ではなく、結局は結果の解釈は実臨床での揺らぎと判断されることが多いと思う。
20	全世界ではどれくらいの研究があるのか（日本がどれだけ遅れているのか、その原因はなにか）というところを知りたいです。
20	多施設でバリデーション研究を行う際に、NHO のデータベースのような大規模な匿名化前のデータにもアクセスできるような立場でなくては実施は難しいように思えました。ご紹介いただいた形以外で多施設のバリデーション研究を実施した実例があればご教示いただけますと幸いです。
20	倫理委員会や電子カルテ委員会の説得が最大の壁のようにも感じており、申請書類にどう書くかでちょうど悩んでおります。

別紙1 Webinar 概要

NDB・DPC データベース研究人材育成 Webinar のご案内

～保健医療介護のデータベース研究に必要な知識と技能を学びませんか？～

◆開催概要

名称	NDB・DPC データベース研究人材育成 Webinar
配信期間	2023年1月10日(火)～1月20日(金)
形式	Webによる動画配信(講義とeLearning) ※事前に登録していただいたアドレスに動画のURLをお送りします ※配信期間中は全ての講義をいつでも視聴することが可能です
主催・共催	主催：厚生労働省科学研究・保健医療介護ビッグデータ人材育成研究班 (研究代表者：東京大学 康永秀生) 共催：日本臨床疫学会
申し込み期間	2022年12月12日(月)正午～2023年1月9日(月)
料金	無料(事前登録制)

◆受講申し込み

[\[申し込みボタン\]](#)

(※2022年12月12日(月)正午からお申込みいただけます)

◆セミナー概要

近年、NDB・DPCなどの保健医療ビッグデータや介護ビッグデータを用いた大規模データベース研究が拡大しています。これらの研究をさらに発展させるためには、臨床医学・疫学・統計学の知識に加えて、データベースに関連する知識や技能に精通した人材の育成が必須です。本セミナーでは、保健

医療介護ビッグデータ研究に興味のある方々を対象とし、各種大規模データベースの概要や研究計画の立案、統計解析についてオンラインでの学習機会を提供します。

◆本セミナーの特徴

- ・保健医療介護ビッグデータ研究で実績のある講師陣による講義・演習を受けけることができる
- ・短期集中で大規模データベース研究の計画立案から統計解析まで学習できる

◆受講対象

保健医療介護ビッグデータ研究に興味のある方。

- ・医療・介護従事者
- ・大学関係者
- ・各種研究機関に勤務する研究者
- ・企業に所属する者
- ・医療政策の担当者

など

◆オンラインコンテンツの概要

	講義内容	講師
1	NDB・DPC等の公的データベースの動向	牧戸香詠子
2	NDBガイドラインの理解度テスト	康永秀生、橋本洋平
3	SQL入門	岡田啓
4	SQLを用いたレセプトデータのハンドリング	松居宏樹
5	NDBを用いた研究	佐藤壮
6	DPCデータを用いた研究1	城大祐
7	DPCデータを用いた研究2	大邊寛幸・小西孝明
8	DeSCデータを用いた研究	康永秀生
9	JMDCデータを用いた研究	金子英弘
10	傾向スコア	小西孝明
11	高次元傾向スコア	石丸美穂
12	操作変数法	大野幸子
13	不連続回帰デザイン・差の差分析	山名隼人
14	時間依存性交絡と周辺構造モデル	大邊寛幸
15	生存時間分析における競合リスクモデル	道端伸明
16	欠測データの取り扱いと多重代入法	麻生将太郎
17	マルチレベル分析	笹渕裕介
18	自己対照ケースシリーズ	橋本洋平
19	機械学習	岡田啓
20	データベースにおけるバリデーション研究	山名隼人

※それぞれ 40-50 分

※各プログラムの詳細は[こちら](#)に掲載。

◆アンケートのお願い

今回はオンライン講義の形式であるため、参加登録された場合には全ての講

義を配信期間中何度でも受講することができます。

各講義の後に無記名アンケートを実施いたします。アンケートは事務局でデータ処理し、セミナーの向上に利用されます。アンケートにお答えいただくと講義資料の pdf がダウンロード可能になります。

◆お問い合わせ

NDB・DPC データベース研究人材育成 Webinar 事務局

東京大学大学院医学系研究科公共健康医学専攻 臨床疫学・経済学分野

113-0033 東京都文京区本郷 7-3-1 東京大学医学部 2 号館本館 3 階

clinepi_adm★www.heer.m.u-tokyo.ac.jp (※ 送信時は★を半角アットマークに変更して下さい)

【ご注意ください】

本件について、日本臨床疫学会にはお問合せしないでください。お問い合わせはすべて、上記の NDB・DPC データベース研究人材育成 Webinar 事務局までお願いいたします。

別紙2 Webinar 詳細

NDB・DPC データベース研究人材育成 Webinar

期間： 2023年1月10日（火）～ 1月20日（金）

形式： Webによる動画配信

共催： 日本臨床疫学会

<お願い>

各講義の後に無記名アンケートを実施いたします。

アンケートにお答えいただくと講義資料の pdf がダウンロード可能になります。

01	NDB・DPC等の公的データベースの動向
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 牧戸香詠子
	<目標> NDB、DPCなどの公的データベースに関する最近の動向を理解する。 公的データベースの連結に関する今後の動向を理解する。
	<想定する対象者> NDB、DPCなどの公的データベースの利用を検討している、あるいは着手している方々

02	NDBガイドラインの理解度テスト(eLearning)
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学

	<p>教授 康永秀生 特任研究員 橋本洋平</p>
	<p><目標> NDB ガイドラインの詳細について、eLearning による設問に解答しながら学ぶ。 NDB データの申請方法について理解する。</p>
	<p><想定する対象者> NDB の利用を検討している方々</p>

	<p>SQL 入門</p>
	<p><講師> 東京大学大学院医学系研究科 糖尿病・生活習慣予防講座 特任講師 岡田啓</p>
0 3	<p><目標> データベースの制御言語である SQL の基礎を理解する。</p>
	<p><想定する対象者> レセプトデータなどのビックデータのハンドリングを検討している方々</p>

	<p>SQL を用いたレセプトデータのハンドリング</p>
	<p><講師> 東京大学大学院医学系研究科 臨床疫学・経済学 助教 松居宏樹</p>
0 4	<p><目標> NDB データに含まれる情報を理解する。 NDB データの構造を理解する。 NDB データのハンドリングに必要な SQL の知識を身に着ける。</p>
	<p><想定する対象者> NDB データの利用を検討している、あるいは着手している方々</p>

05	NDB を用いた研究
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 佐藤壮
	<目標> NDB、DPC データなどを用いた研究の近年の動向を理解する。 NDB データを用いた先行研究について学ぶ。 NDB データを用いた研究の利点と限界を理解する。
	<想定する対象者> NDB データの利用を検討している、あるいはその研究に着手している方々

06	DPC データを用いた研究1
	<講師> 東京大学大学院医学系研究科 ヘルスサービスリサーチ講座 特任准教授 城大祐
	<目標> DPC データの構造、含まれる情報を理解する。 DPC データを利用した研究の方法論を理解する。 DPC データを利用した研究計画書、データ抽出依頼書の作成方法を理解する。 DPC データを用いた研究の利点と限界を理解する。
	<想定する対象者> DPC データの利用を検討している、あるいはその研究に着手している方々

07	DPC データを用いた研究2
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 大邊寛幸、小西孝明
	<目標> DPC データを用いた先行研究について学ぶ。 DPC データを用いた研究の利点と限界を理解する。
	<想定する対象者> DPC データの利用を検討している、あるいはその研究に着手している方々

08	DeSC データを用いた研究
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 教授 康永秀生
	<目標> DeSC データの構造、含まれる情報を理解する。 DeSC データを用いた先行研究について学ぶ。 DeSC データを用いた研究の利点と限界を理解する。
	<想定する対象者> DeSC データの利用を検討している、あるいはその研究に着手している方々

09	JMDC データを用いた研究
	<講師> 東京大学大学院医学系研究科 先進循環器病学講座 特任講師 金子英弘
	<目標> JMDC データの構造、含まれる情報を理解する。 JMDC データを用いた先行研究について学ぶ。 JMDC データを用いた研究の利点と限界を理解する。
	<想定する対象者> JMDC データの利用を検討している、あるいはその研究に着手している方々

10	傾向スコア
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 小西孝明
	<目標> 傾向スコア分析の基礎を理解する。 傾向スコアマッチング・逆確率による重みづけについて学ぶ。 傾向スコア分析を用いた先行研究について学ぶ。
	<想定する対象者>

	臨床研究・疫学研究の初級者以上
--	-----------------

11	高次元傾向スコア(High dimensional propensity score: hdPS)
	<講師>東京医科歯科大学健康推進歯学分野 プロジェクト助教 石丸美穂
	<目標> 傾向スコアを大規模データの性質に合わせて拡張した高次元傾向スコアの基礎を理解する。 高次元傾向スコアの利点を理解する。 高次元傾向スコア分析を用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の中級者以上

12	操作変数法
	<講師>東京大学大学院医学系研究科 イートロス医学講座 特任講師 大野幸子
	<目標> Unmeasured confounders の対処が可能である操作変数法の基礎理論を理解する。 操作変数法の3つの前提条件を説明できるようになる。 操作変数法を用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の初級者以上

13	不連続回帰デザイン・差の差分析
	<講師>東京大学大学院 医学系研究科 ヘルスサービスリサーチ講座 特任助教 山名隼人
	<目標> 不連続回帰デザイン・差の差分析の基礎を理解する。

	それぞれの手法がどんな臨床疑問に適した研究デザインなのかを説明できる。 不連続回帰デザイン・差の差分析を用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の中級者以上

	時間依存性交絡と周辺構造モデル
	<講師> 東京大学大学院医学系研究科 臨床疫学・経済学 大邊寛幸
1 4	<目標> 時間依存性の治療や交絡の概念を理解する。 時間依存性交絡と周辺構造モデルの基礎を理解する。 時間依存性交絡と周辺構造モデルを用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の中級者以上

	生存時間分析における競合リスクモデル
	<講師> 東京大学大学院医学系研究科 ヘルスサービスリサーチ講座 特任助教 道端伸明
1 5	<目標> 競合リスクとは何かを理解する。 生存時間分析における競合リスクモデルの基礎を理解する。 生存時間分析における競合リスクモデルを用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の初級者以上

	欠測データの取り扱いと多重代入法
1 6	<講師> 東京大学大学院医学系研究科 生物統計情報学 特任助教 麻生将太郎
	<目標>

	<p>欠測が起こるメカニズムを理解する。</p> <p>欠測データの取り扱いと多重代入法の基礎を理解する。</p> <p>多重代入法を用いた先行研究について学ぶ。</p>
	<p><想定する対象者></p> <p>臨床研究・疫学研究の初級者以上</p>

17	マルチレベル分析
	<p><講師>自治医科大学 データサイエンスセンター</p> <p>講師 笹渕裕介</p>
	<p><目標></p> <p>データの階層構造を考慮したマルチレベル分析の基礎を理解する。</p> <p>マルチレベル分析を用いた先行研究について学ぶ。</p>
	<p><想定する対象者></p> <p>臨床研究・疫学研究の初級者以上</p>

18	自己対照ケースシリーズ(Self-controlled case series: SCCS)
	<p><講師>東京大学大学院医学系研究科 臨床疫学・経済学</p> <p>特任研究員 橋本洋平</p>
	<p><目標></p> <p>自己対照ケースシリーズの基礎を理解する。</p> <p>自己対照ケースシリーズの利点と欠点を説明できる。</p> <p>自己対照ケースシリーズを用いた先行研究について学ぶ。</p>
	<p><想定する対象者></p> <p>臨床研究・疫学研究の中級者以上</p>

19	機械学習
	<講師> 東京大学大学院医学系研究科 糖尿病・生活習慣病予防講座 特任講師 岡田啓
	<目標> 機械学習を用いたデータベース研究の基礎を理解する。 機械学習を用いた先行研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の中級者以上

20	データベースにおけるバリデーション研究
	<講師> 東京大学大学院医学系研究科 ヘルスサービスリサーチ講座 特任助教 山名隼人
	<目標> データベースの妥当性の検証を行うバリデーション研究の重要性について理解する。 バリデーション研究の方法を学ぶ。 既存のバリデーション研究について学ぶ。
	<想定する対象者> 臨床研究・疫学研究の初級者以上

NDB・DPC等の公的データベースの動向

東京大学大学院医学系研究科臨床疫学・経済学
 牧戸 香詠子

NDB

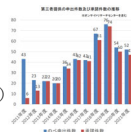
- 第三者提供は2011年から開始
- 2020年10月より高齢者の医療の確保に関する法律に第三者提供等を規定
- 提供するデータは特別抽出、サンプリングデータセット、集計表情報



https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/hoken/reseputo/index.html

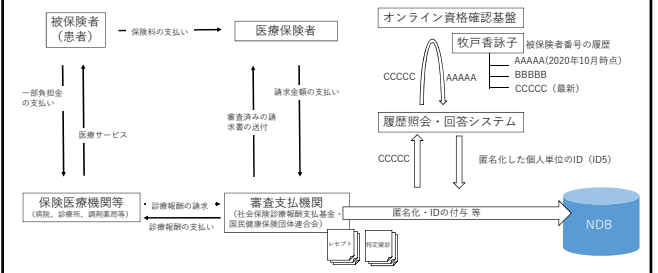
特徴

- 全国の病院・診療所のレセプトデータおよび特定健診情報
- レセプトデータは月1回（2009年4月診療分～最新：3～6ヶ月前）、特定健診情報は年1回（2008年度～最新：2年前）DBへ格納
- 第三者提供の実績は多い
- オンサイトリサーチセンターの設置がある
- オープンデータを公表（年1回）
 （2013年度～特定健診情報・2014年4月診療分～レセプト情報）
- データ構造が複雑



<https://www.mhlw.go.jp/content/12400000/000947953.pdf>

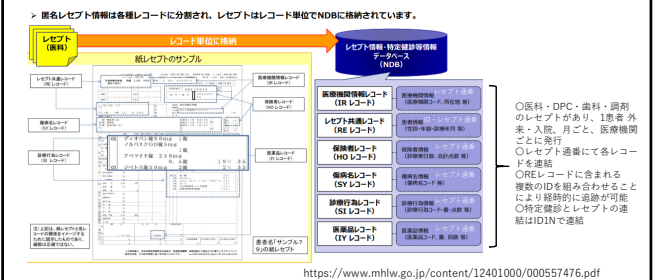
レセプト等がNDBに格納されるまでの流れ



格納されているデータ

- 被用者保険（協会けんぽ、健康保険組合、共済組合等）
- 国民健康保険
- 後期高齢者医療制度
- 医療扶助
- 特定健診情報（40歳～75歳）
 - × 紙レセプト
 - × 自費診療
 - × 労災保険、自賠責保険
 - × 妊娠・出産（保険診療外）
 - × 事業主健診等
- その他（2022年4月～所得階層情報・住所情報が格納）

格納されているデータ



格納されているデータ（抽出テンプレート）

【別添8：抽出テンプレート】

項目	データ形式	注	出力	項目名	出力項目名
1. 患者ID	文字	1	1	患者ID	患者ID
2. 診療科目	文字	1	1	診療科目	診療科目
3. 医師	文字	1	1	医師	医師
4. 患者氏名	文字	1	1	患者氏名	患者氏名
5. 生年月日	文字	1	1	生年月日	生年月日
6. 性別	文字	1	1	性別	性別
7. 年齢	文字	1	1	年齢	年齢
8. 入院日	文字	1	1	入院日	入院日
9. 退院日	文字	1	1	退院日	退院日
10. 手術開始日時	文字	1	1	手術開始日時	手術開始日時
11. 手術終了日時	文字	1	1	手術終了日時	手術終了日時
12. 手術時間	文字	1	1	手術時間	手術時間
13. 手術部位	文字	1	1	手術部位	手術部位
14. 手術器具	文字	1	1	手術器具	手術器具
15. 手術内容	文字	1	1	手術内容	手術内容
16. 手術経過	文字	1	1	手術経過	手術経過
17. 手術結果	文字	1	1	手術結果	手術結果
18. 手術費用	文字	1	1	手術費用	手術費用
19. 手術リスク	文字	1	1	手術リスク	手術リスク
20. 手術後経過	文字	1	1	手術後経過	手術後経過
21. 手術後経過	文字	1	1	手術後経過	手術後経過
22. 手術後経過	文字	1	1	手術後経過	手術後経過
23. 手術後経過	文字	1	1	手術後経過	手術後経過
24. 手術後経過	文字	1	1	手術後経過	手術後経過
25. 手術後経過	文字	1	1	手術後経過	手術後経過
26. 手術後経過	文字	1	1	手術後経過	手術後経過
27. 手術後経過	文字	1	1	手術後経過	手術後経過
28. 手術後経過	文字	1	1	手術後経過	手術後経過
29. 手術後経過	文字	1	1	手術後経過	手術後経過
30. 手術後経過	文字	1	1	手術後経過	手術後経過

レコードの結合

連結・追跡するためのID

共通の項目（キー情報）について
以下の項目を利用することで、1つのレセプトを特定する。患者を特定する事が可能です。

項目名	内訳
レセプト番号	レセプト内での並び順（発行番号）を指し示す
レセプト年度	レセプト年度で示される年度（1年度に1レセプト発行される）

格納されているデータ（記録条件仕様）

例：医科レセプト （IRレコード）

(4) 各種レコードの記録条件に関する事項
コードは入力する文字の種類、最大バイト数は項目の最大バイト数、項目形式は項目長が設定済み可変長か定数かを示す。

(7) 診療報酬コード

項目	バイト	最大バイト	項目形式	記録内容	備考
レセプト番号	数字	2	固定	"01"を記録する	
診療科目	数字	1	固定	診療科目コードを記録する	
医師	数字	2	固定	医師コードを記録する	
患者氏名	数字	1	固定	患者氏名コードを記録する	
生年月日	数字	7	固定	生年月日コードを記録する	
性別	数字	2	固定	性別コードを記録する	
手術開始日時	数字	40	可変	手術開始日時を記録する	
手術終了日時	数字	40	可変	手術終了日時を記録する	
手術時間	数字	2	固定	手術時間を記録する	
手術部位	数字	2	固定	手術部位コードを記録する	
手術器具	数字	2	固定	手術器具コードを記録する	
手術内容	数字	2	固定	手術内容コードを記録する	
手術経過	数字	2	固定	手術経過コードを記録する	
手術結果	数字	2	固定	手術結果コードを記録する	
手術費用	数字	2	固定	手術費用コードを記録する	
手術リスク	数字	2	固定	手術リスクコードを記録する	
手術後経過	数字	2	固定	手術後経過コードを記録する	

https://www.ssk.or.jp/seikyushihara/rezept/iryokikan/iryokikan_02.html

提供されるデータ

- 特別抽出：個票
- サンプリングデータセット：
 - 医科入院・DPCは10%、医科入院外・調剤は1%抽出
単月（2011年～1・4・7・10月）
 - 探索的研究に対応
 - IDは格納されていないため、経時的な追跡は不能
- 集計表情報：
 - 複雑な集計（IDで追跡した後の集計等）は原則お断り

提供されるデータ(ID)

- 匿名化IDの作成はハッシュ関数を使用

	作成の元となる情報	特徴
ID1	保険者番号、被保険者証の記号・番号、生年月日、性別	保険者の変更や誤記により紐づけができなくなる可能性がある
ID2	漢字氏名・生年月日・性別	氏名の変更や誤記により紐づけができなくなる可能性がある
ID4	カナ氏名・生年月日・性別	上記の課題に対応
ID5	2020年10月時点の被保険者証の記号・番号・枝番	上記の課題に対応

<https://www.mhlw.go.jp/content/12401000/000557476.pdf>

オンサイトリサーチセンター

- 厚生労働省・東大・京大に設置
- NDBデータのほぼ全データを閲覧できる
- 利用形態は2種類
 - 研究者がオンサイトリサーチセンターにいき、全解析を完了する
中間生成物の持ち出しを行い、自施設で解析する
- セキュリティ要件はかなり厳格
 - スマホ、PC等電子媒体の持ち込み禁止
 - 監視カメラの設置がある 等
- 外部委託はできない
- PostgreSQL・R・SPSS・STATAが利用可能

死亡情報の収載

- 死亡情報は転帰区分等にしかなく、自殺や事故等で死亡した場合、死亡したかどうかは不明
- 人口動態調査票（≒死亡診断書（死体検案書）+死亡届）と同一の収集ルートで市町村に提供を求める
- 下記の内容を検討のうえ、収載
 - 死亡年月日
 - 死亡したところの種別
 - 死亡の原因
 - 死亡の種類 等
- 2023年度分からの死亡情報を2024年以降に収載予定
- DPCDB、介護DBはNDBと連結申出のみ死亡情報が提供される

厚生労働省が第三者提供する公的データベース

- NDB（高齢者の医療の確保に関する法律）
 - DPCDB（健康保険法）
 - 介護DB（介護保険法）
 - 難病・小児慢性疾患DB
 - がん登録DB（がん登録推進法）
 - 人口動態死亡票（統計法）等
- } 連結申出が可能

今後、提供もしくは連結を予定しているデータベース

- 難病・小児慢性疾患DB
 - がん登録DB（がん登録推進法）
 - 感染症DB
 - 予防接種DB
 - 障害福祉DB
 - 次世代DB（次世代医療基盤法）
- ※次世代DBは認定事業者がDBを保有する民間データベース

DPCDB

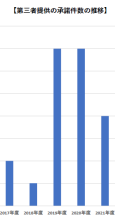
- 第三者提供は2017年から開始（集計表情報のみ）
- 2022年度から個票の提供を開始



https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryuuhoken/dpc/index.html

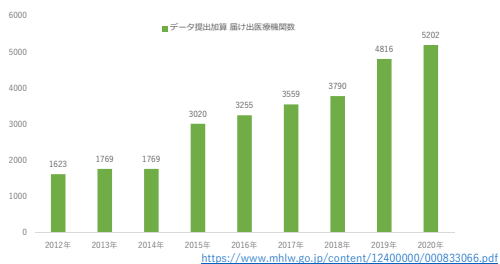
特徴

- DPC病院等から提出されたデータは年1回DBへ格納（2011年度実施分～最新のデータ：2年前）
- 第三者提供の実績はNDBと比較すると少ない
- オンサイトリサーチセンターの設置やオープンデータの公表はない
- データ構造はNDBほど複雑ではない
- DPC病院等は年々増加傾向

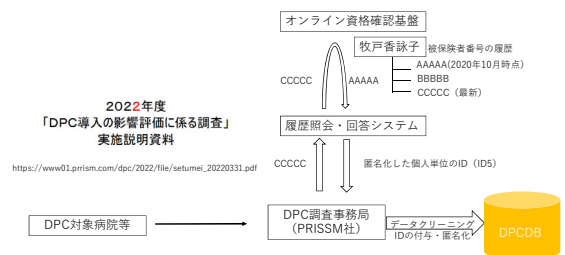


<https://www.mhlw.go.jp/content/12301000/000952170.pdf>

DPC病院数の推移



DPCデータ格納の流れ



格納されているデータ

様式名	内容	入力される情報
様式1	患者属性や病態等の情報	性別、生年月日、病名、病期分類など
様式4	医科保険診療以外の診療情報	保険診療以外(公費、先進医療等)の実施状況
Dファイル	診断群分類点数表に基づく診療報酬算定情報	包括レセプトの情報
入院E結合ファイル	医科点数表に基づく診療報酬算定情報	入院診療患者の医科点数表に基づく出来高情報
外来E結合ファイル	外来患者の医科点数表に基づく診療報酬算定情報	外来診療患者の医科点数表に基づく出来高情報
Hファイル	日ごとの患者情報	重症度、医療・看護必要度
様式3	施設情報(施設ごとで作成)	入院基本料等の届け出状況
Kファイル	3情報から生成した一次共通IDに関する情報	生年月日、カナ氏名、性別から生成した一次共通ID

詳細な項目の内容は「DPC導入の影響評価に係る調査」実施説明資料を参考
<https://www.mhlw.go.jp/content/12400000/000863369.pdf>

提供されるデータ

- 留意が必要とされるデータ項目は下記のように変換
 例：生年月日→生年月へ変換
 患者居住地情報→2次医療圏単位への変換
- 施設コード、保険者番号、医師コードに関しては専門医委員会
 で特に認める場合を除き、原則として提供対象外
 ※施設コード = 都道府県番号 + 医療機関コード

DPCデータ内での連結

ファイル名	内容	入力される内容	施設コード	データ識別番号	入院年月日
様式1	患者属性や病態等の情報	性別、生年月日、病名、病期分類など	○	○	○
様式3	施設情報	入院基本料等の届け出状況	○	×	×
様式4	医科保険診療以外の診療情報	医科保険診療以外(公費、先進医療等)の実施状況	○	○	○
入院E結合ファイル	医科点数表に基づく診療報酬算定情報	入院の出来高レセプト	○	○	○
外来E結合ファイル	外来患者の医科点数表に基づく診療報酬算定情報	外来の出来高レセプト	○	○	×
Dファイル	診断群分類点数表に基づく診療報酬算定情報	DPCレセプト	○	○	○
Hファイル	日ごとの患者情報	重症度、医療・看護必要度	○	○	○
Kファイル	3情報から生成した一次共通IDに関する情報	患者の生年月日、カナ氏名及び性別から生成した一次共通ID及び保険診療年度等	○	○	○

施設コードとデータ識別番号から各ファイルを連結するキー情報を申出ごとに作成し、提供もしくは、結合して提供も可能
<https://www.mhlw.go.jp/content/12301000/000995129.pdf>

介護DB

- 第三者提供は2018年から開始
- 提供するデータは特別抽出、サンプリングデータセット、集計表情報



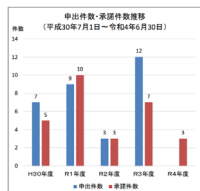
匿名介護情報等の提供について

※本ページは令和2年10月1日以後の手続きに関するページです。令和2年9月30日以前の匿名介護認定情報・介護レセプト等情報の提供に係る手続きについては、ご注意ください。

https://www.mhlw.go.jp/stf/shing2/0000198094_00033.html

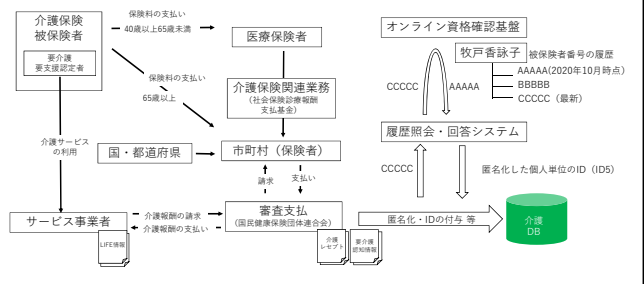
特徴

- 介護事業所から収集された介護レセプト情報・匿名LIFE情報および市町村から収集された要介護認定情報
- 介護保険受給者台帳情報(ID等格納)されているため、経時的に経過を追うことが可能
- 保険者等から提出されたデータは月1回、DBへの格納
 要介護認定情報：2009年4月分～
 介護レセプト情報：2012年4月分～
 LIFE情報：2021年4月～
- 第三者提供の実績はNDBと比較すると少ない
- オンサイトリサーチセンターの設置はない
- 2022年11月からオープンデータの公表を開始
 (2018年度実施分～年1回秋ごろ)
- テーブル定義書がある



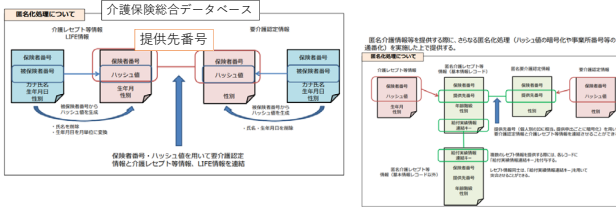
<https://www.mhlw.go.jp/content/12301000/000995136.pdf>

介護レセプト等が介護DBに格納されるまでの流れ



格納されているデータ

- ・介護保険総合データベースは2013年から運用開始



格納されているデータ (テーブル定義書)

レコード識別名	データの概要
要介護認定情報	D_NINTEI 認定調査項目、主治医等受療の項目等
給付実績情報 (基本情報レコード)	DT1111_041 要介護者の基本項目、施設、要介護、介護支援センター及び自給型の合計給率
給付実績情報 (詳細情報レコード)	DT1111_D11 介護サービスの種類、介護給付費率施設等サービスコード等
給付実績情報 (要介護認定番号・要介護施設情報情報レコード)	DT1111_D21 介護施設種別、施設名称、サービスコード等、設置形態、主要点検等の情報
給付実績情報 (特別介護・特別介護情報レコード)	DT1111_D31 介護療養型医療施設・介護老人保健施設の介護給付率、介護施設に認定されている特定診療費及び特別療養費について、療養回費、介護サービス提供回数等
給付実績情報 (要介護費用情報レコード)	DT1111_D41 介護保険制度の介護給付費用に記録されている介護費用に関するデータ項目
給付実績情報 (在宅介護費用情報レコード)	DT1111_D51 在宅介護費用・介護予防支援の介護給付費用に記録されている介護サービスの種類、介護給付費率施設等サービスコード等
給付実績情報 (福祉用具販売費用情報レコード)	DT1111_D61 福祉用具販売費における福祉用具の購入日、購入金額、販売店、福祉用具の品目コード、製造した事業者名、販売した事業者名等のデータ
給付実績情報 (在宅改修費情報レコード)	DT1111_D71 在宅改修費に関するデータ項目
給付実績情報 (介護サービス提供情報レコード)	DT1111_D81 介護サービス提供に関する情報として、利用者情報、支給額等のデータ
受給者台帳情報	D75341 被保険者の属性に関する情報として、資格取得年月日、資格喪失年月日、要介護度、認定有効期間、住所地特別適用開始年月日、住所地特別適用終了年月日等のデータ、IQ4、IQ5

https://www.mhlw.go.jp/content/12301000/000343808.pdf

格納されているデータ (匿名LIFE情報)

○科学的介護情報システム (Long term care Information system For Evidence : LIFE) の略

○格納されている主なデータ

利用者の状態・ケアの内容等の情報

- ・利用者情報
- ・科学的介護推進情報 (アセスメント結果、既往歴情報等 等)

提供されるデータ

- ・特別抽出：個票
- ・サンプリングデータセット：
 - 介護レセプト等情報のうち、基本情報・明細情報・居宅サービス計画費情報レコード
 - 2012年度以降の各年4月、7月、10月、1月サービス提供分のデータから一定の割合で抽出
 - 利用の少ないサービス項目コード等について匿名化処理
 - IDは格納されていないため、経時的な追跡は不能
- ・集計表情報

NDB・DPC・介護DBの共通の特徴

- ・申請書類やガイドラインは類似している
- ・連結申出の場合は申請書類等は、連結するデータが必要
- ・専門委員会 (年4回開催：6・9・12・3月) での審査にて承認が必要 ※連結の場合は合同委員会
- ・連結・追跡するための個人単位のIDがある
- ・手数料がかかる ※公的機関や厚労科研究費等の取得があれば免除

DPC	4,250円/時間
NDB	7,700円/時間
介護DB	5,900円/時間

NDB・DPC・介護DBの共通の特徴

- ・ガイドライン記載の情報セキュリティ要件を遵守する必要
- 物理的・技術的安全管理措置

入退室管理	利用場所の施設、取次者の名札等の着用、台帳等による入退室管理及び入退室の記録を定期的にチェックし、その妥当性の確認、入退室管理の保管 (1年)
データ保護	クリアスクリーン等による窃聴防止
盗難・紛失	匿名レセプト情報等、中間生成物等が格納された記録媒体の管理
認証・識別	専用端末の窃盗防止用ワイヤレス等設置による盗難防止
ウイルス対策	MFA (二要素認証) による利用者識別の確保
消去	外部からの情報受領時には、不正なウイルスの混入を防止
ログ管理	利用終了後は専用ソフトウェア等を利用して、復元不可能な形で消去
	アクセスの記録及び定期的なログの確認・保管

- ・実地監査 (対象は主に特別抽出の申出)



https://www.mhlw.go.jp/content/12301000/000952167.pdf

NDB・DPC・介護DBの共通の特徴

- 組織的安全管理措置
 - 利用者及び取扱者の権限、責務及び業務を明確にすること。
 - 運用管理規程を定めること。
- 人的安全管理措置
 - 取扱者は、高齢者の医療の確保法等又はこれらの法律に基づく命令の規定に違反し、罰金以上の刑に処せられ、その執行を終わり、又は執行を受けることがなくなった日から起算して5年を経過しないこと
 - 提供申出者（もしくは利用者）は取扱者に対し、匿名レセプト情報等を取り扱う上で必要な教育及び訓練を行うこと。

NDB・DPC・介護DBの共通の特徴

不適切利用の疑い

利用制限・詳細調査・措置の検討

専門委員会での審議

措置の決定

不適切利用の内容
他の情報と照合を行う
利用期間終了以降にレセプト情報等の返却・消去を行わない
情報セキュリティ上の危険にさらした
提供を受けた匿名レセプト情報等を紛失・漏洩した
申出書に記載された目的以外の利用
公表物確認をせずに公表した
※公表物確認：学会発表や論文投稿等をする前の事前の厚生労働省の確認
その他

高齢者の医療の確保に関する法律等により、罰則規定があり、また提供申出者（法人等）に罰則がかかるため、充分注意が必要です！

NDB・DPC・介護DBの追跡・連結

【連結・追跡するためのID】

ID4	3情報（カナ氏名、生年月日、性別）から作成し、匿名化したID
ID5	被保険者番号（2020年10月時点）を匿名化したID

【各データおよびIDの格納時期】

	2008/4	2009/4	2011/4	2012/4	2016/4	2020/3	2020/4	2022/2	2022/3	2022/4
NDB										
ID1-2										
ID1N										
ID4										
ID5										
DPC										
ID4										
ID5										
介護DB										
ID4										
ID5										

専門委員会での審査ポイント

- 相当の公益性があることが必要
- 実現可能性があるかどうか
- 研究を実施する上で「必要最小限の範囲」となっているか
 - 探索的研究となっていないか
 - 複数の研究が1申出に盛り込まれていないか
 - 研究内容を鑑みて、不必要な項目を抽出していないか
- 個人の識別可能性が上がる場合は慎重な審査を行う
- 研究者や所属施設、研究施設が複数にまたがる申出は、セキュリティ対策実践の難易度が上がるため、その対応方法について慎重な審査を行う

NDBの今後の動向

- 医療介護データ等の解析基盤：HIC（Healthcare Intelligence Cloud）
研究者がクラウド上でNDB等を解析できる基盤の開発

	HICの機能	想定される研究者
ポータル機能	NDB・介護DB等の提供申出、利用及び終了に至る一連の手続き教育・啓発のためのコンテンツの形成や各種マスターの共有等	全研究者
探索的利用環境	ダミーデータを用いて探索・試行的に分析するための環境	これからNDB研究を始めようと考えている申出者
HIC解析環境	専門委員会の審査にて承諾された提供申出ごとに、利用者に対して解析環境	承諾を受けた申出者

- 施行的利用後、本格運用予定

他の公的データベース

難病・小児慢性疾患DB	がん登録DB
2019年度～患者の同意に基づく提供	2018年度～
<ul style="list-style-type: none"> 臨床調査個人票（難病）・医療意見書（小児）の情報 研究用ID（確率的ID）の提供により経時的に追跡が可能 連結用ID（ID4、ID5）の取得予定 指定難病患者データ及び小児慢性特定疾病児童等データ 	<ul style="list-style-type: none"> 原発癌の初回の診断・診療および死亡情報 利用形態はリンクージュ利用（調査研究対象者への説明と同意の取得が必要）および集計統計利用 リンクージュ利用は厚生科学審議会にて審査 集計登録利用は国立がん研究センター全国がん登録情報提供等審議委員会にて審査 ※都道府県がん情報は都道府県が開催する審議会等にて審査
<ul style="list-style-type: none"> 提供依頼申出者の範囲は、厚労省・文科省が補助を行う研究事業 都道府県・指定都市・中核市 その他 	<ul style="list-style-type: none"> 提供依頼申出者の範囲は、研究者 国・都道府県の関係者 ※民間事業者は除外されていない
<ul style="list-style-type: none"> 申請書、ガイドラインはNDB（DPCDB・介護DB）をもとに作成されており、類似点が多い 	<ul style="list-style-type: none"> 遵守すべき安全管理措置等はガイドラインにて規定
手数料の徴収はない	手数料の徴収あり

他の公的・民間データベース

予防接種DB	感染症DB	障害福祉DB	次世代DB (民間)
予防接種記録 ・年齢、性別 ・接種日 ・接種場所 ・接種したワクチン ・電子予防接種情報等 副反応疑い情報 ・年齢、性別 ・接種日 ・接種場所 ・接種したワクチン等 医学的所見 (副反応時) ・発症日 ・診断名 (疑い含む) ・重症度 ・転帰 ・因果関係の推定等	感染症発生届の内容 ・氏名・年齢・性別 ・診断病名 ・症状 ・診断方法 ・診断年月日 ・検体採取日 ・感染推定年月日 ・感染原因等 ※感染症によって記載事項は異なる	事業所情報 ・事業所台帳情報 ・障害児施設台帳情報 受給者情報 ・受給者台帳情報 ・障害児支援受給者台帳情報 給付費等明細情報 ・給付費明細データ (障害福祉サービス・障害児支援) ・計画相談給付費明細データ ・障害児相談給付費明細データ 障害支援区分認定情報 ・障害支援区分認定データ	医療機関の診療情報 ・レセプト ・電子カルテ ・健診情報等

連結用ID (ID4、ID5)

連結用ID (ID4)

連結用ID (ID4、ID5)

※連結の方法や連結して提供する情報の範囲等については、各DBが法制化され、DBの情報の詳細が確定した後、検討予定

NDB eLearning 教材 1

NDBガイドラインの理解

次の各問に当てはまる選択肢を1つ選べ。

1. NDB 活用の本来目的を以下から1つ選べ。

- ① 医薬品・医療機器の市販後調査
- ② 臨床研究・疫学研究の推進
- ③ 医療安全管理
- ④ 医療費適正化
- ⑤ 医療サービスの情報公開

2. 匿名レセプト情報・匿名特定健診等情報の提供は、以下のどの法律に基づいているか？

- ① 医療法
- ② 医師法
- ③ 地域における医療及び介護の総合的な確保を推進するための関係法律の整備に関する法律
- ④ 高齢者の医療の確保に関する法律
- ⑤ 医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律

3. 匿名レセプト情報・匿名特定健診等情報のデータ保有者はだれか？

- ① 総務大臣
- ② 厚生労働大臣
- ③ 都道府県知事
- ④ 保険者
- ⑤ 支払基金

4. 有識者から構成される「匿名医療情報等の提供に関する専門委員会」は以下のいずれの下に設けられているか？

- ① 社会保障審議会

- ② 厚生科学審議会
- ③ 社会保険審査会
- ④ 中央社会保険医療協議会
- ⑤ 医薬品等行政評価・監視委員会

5. 第三者申出件数は 2018 年度以降、年間 X 件を超えている。X に入る適当な数値を以下から 1 つ選べ。

- ① 50
- ② 100
- ③ 300
- ④ 500
- ⑤ 1000

6. サンプルングデータセットを除き、提供申出が承諾された日から NDB データが提供されるまでに要した平均所要日数は、2018-2019 年度においては X 日以内であった。X に入る適当な数値を以下から 1 つ選べ。

- ① 30
- ② 90
- ③ 180
- ④ 300

7. ①—④誤っているものを 1 つ選べ。すべて正しい場合は⑤を選べ。

- ① 「中間生成物」とは、匿名レセプト情報等を提供したのち取扱者が生成したものであって、最終生成物と成果物以外のものをいう。
- ② 「中間生成物」については、厚生労働省による公表前の事前の承認を得て成果物となったものを除き、取扱者以外に公表することは禁じられている。
- ③ 本ガイドラインにおいて「最終生成物」とは、匿名レセプト情報等を提供したのち取扱者が最終的に生成したものであって、厚生労働省による公表前の事前の承認を得ていないものをいう。
- ④ 「最終生成物」については、厚生労働省による公表前の事前の承認を得て成果物となったものも含めて、取扱者以外に公表することができる。

8. ①—④誤っているものを 1 つ選べ。すべて正しい場合は⑤を選べ。

- ① 提供申出者及び取扱者は、提供された匿名レセプト情報等について、個人情報の保護に関する法律に規定する個人情報に準じた取扱いを行うこととする。
- ② 提供申出者及び取扱者は、医療情報システムの安全管理に関するガイドラインに定められた措置に準じた措置を匿名レセプト情報等の利用形態を勘案した上で適切に講ずるものとする。
- ③ 提供申出者は、外部委託を行う必要性が、研究の目的及び内容に照らして合理的である場合、匿名レセプト情報等を用いた研究を外部委託することができる。
- ④ オンサイトリサーチセンター内での作業については外部委託することは認められない。

9. 厚生労働省は、匿名レセプト情報等の提供により、提供申出者、取扱者及び第三者に患者等の情報が特定されることがないように、各提供申出書の内容に応じて、専門委員会における議論及び技術的な問題等を勘案し、提供する匿名レセプト情報等に適切な処理を施すものとし、処理を講じた場合には、その内容を提供申出者及び取扱者に明示するものとする、とされている。「適切な処理」に該当しないものを以下から1つ選べ。

- ① 第三者に情報が特定される可能性のある特定個人の全データ削除
- ② データの再ソート（配列順の並べ替え）
- ③ 特定個人又は特定機関の識別情報のトップ（ボトム）・コーディング
- ④ 特定個人又は特定機関の識別情報のグルーピング（リコーディング）
- ⑤ リサンプリング

10. ①—④誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 医療機関・薬局コードについては、専門委員会が特に認める場合を除き、原則として提供しないこととされている。
- ② 保険者番号については、専門委員会が特に認める場合を除き、原則として提供しないこととされている。
- ③ 技術的な問題等により適切な処理が行い難い場合には、専門委員会の議論を経て、匿名レセプト情報等の提供を行わない場合もあり得る。
- ④ 厚生労働省は、提供する匿名レセプト情報等について利用方法や情報の範囲等を勘案し、専門委員会の意見を聴取した上で適切な処理を行うこととする。

11. ①—④誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 提供申出書は、匿名レセプト情報等の提供の判断要件となる「利用目的」ごとに作成するものとする。

- ② 提供申出者が実施する複数の研究に用いる匿名レセプト情報等について併せて提供申出を行って差し支えない。
- ③ 提供された匿名レセプト情報等1ファイルについて、当該ファイルを別の記憶装置に複写・保存する行為は2回に限定する。
- ④ 当該記憶装置の保存・複製ファイルが消去されない限り、別の記憶装置への保存・複写は原則として認めない。

12. 匿名レセプト情報等の提供申出者の範囲について、①—④のうち該当しないものはどれか？すべて該当する場合は⑤を選べ。

- ① 公的機関（国の行政機関、都道府県、市区町村）
- ② 大学その他の研究機関
- ③ 研究開発独立行政法人等
- ④ 民間事業者等

13. ①—④誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 法人（公的機関を除く法人その他の団体で代表者又は管理人の定めがあるもの）は、原則として登記された法人等を単位として提供申出を行う。
- ② 公的機関が開設する医療機関の場合、当該医療機関を開設する公的機関を単位として提供申出を行う。
- ③ 国立病院機構及び労働者健康安全機構が開設する医療機関の場合、当該医療機関を単位として提供申出を行う。
- ④ 大学病院の場合、当該大学病院を単位として提供申出を行う。

14. ①—③のうち、誤っているものを1つ選べ。すべて正しい場合は④を選べ。

- ① 提供申出者が公的機関の場合、担当者の身分証明書等の写しを提出する。
- ② 提供申出者が法人等の場合、担当者の身分証明書等の写しを提出する。
- ③ 提供申出者が個人の場合、提供申出者の身分証明書等の写しを提出する。

15. ①—④のうち、誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 利用目的が特定の商品又は役務の広告又は宣伝に直接利用する又は利用されると推測される場合、提供は認められない。
- ② 匿名レセプト情報等の直接的な利用目的が、企業等の組織内部に企業等の組織内部における業務上の資料として利用される場合、提供は認められない。
- ③ 匿名レセプト情報等の直接的な利用目的が、企業等の特定の顧客に対するレポート作成の基礎資料とされるような場合、提供は認められない。

- ④ 他の研究や政策利用等を阻害するような場合でも、特許の取得は認められる。

16. ①-④のうち、ガイドラインが定める研究の内容に該当するものを1つ選べ。すべて該当場合は⑤を選べ。

- ① 医療分野の研究開発に資する分析
- ② 適正な保健医療サービスの提供に資する施策の企画及び立案に関する調査
- ③ 疾病の原因並びに疾病の予防、診断及び治療の方法に関する研究
- ④ 保健医療の経済性、効率性及び有効性に関する研究

17. ①-⑤のうち、誤っているものを1つ選べ。すべて正しい場合は⑥を選べ。

- ① 匿名レセプト情報等を他の情報と照合してはならない。
- ② 提供申出を行う匿名レセプト情報等が研究内容に鑑みて最小限であるとする根拠を記入する。
- ③ 匿名レセプト情報等を実際に利用する場所は日本国内に限る。
- ④ 匿名レセプト情報等の利用期間の上限は、原則として、1年間とする。
- ⑤ 匿名レセプト情報等の提供に必要な媒体（CD-R、DVD、外付けハードディスク等）は、匿名レセプト情報等の情報量等を勘案し、厚生労働省において用意する。

18. ①-④誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 提供申出に係る手数料は、人件費等を踏まえた時間単位の金額（1時間までごとに6100円）に、作業に要した時間を乗じて得た額とする。
- ② 作業に要した時間のうち申出処理業務には、申出書類確認・専門委員会への諮問手続・データの抽出条件の精査等が含まれる。
- ③ 作業に要した時間のうちデータ抽出業務とは、SQL作成・テスト実施・結果の検証等が含まれる。
- ④ 補助金は免除されない。

19. ①-④のうち誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 提供申出書等の受付窓口は、厚生労働省保険局医療介護連携政策課保険データ企画室である。
- ② 事務処理を円滑に行うため受付窓口を外部委託する場合がある。
- ③ 厚生労働省は、担当者及び代理人に対して、氏名、生年月日及び住所を確認できる書類のコピーを求める。

- ④ 受付窓口へ郵送により提出する書類は、原則として直筆の必要がある書類のみとし、その他についてはEメールでの送付を可とする。

20. 匿名レセプト情報等を利用する必要性等の基準のうち、該当しないものはどれか？①―④から1つ選べ。すべて正しい場合は⑤を選べ。

- ① 利用する匿名レセプト情報等の範囲及び匿名レセプト情報等から分析する事項が研究内容から判断して必要最小限であること。
- ② 匿名レセプト情報等の性格に鑑みて、その利用に合理性があり、他の情報では研究目的が達成できないこと。
- ③ 匿名レセプト情報等の利用期間と研究の計画・公表時期が整合的であること。
- ④ 匿名レセプト情報等の利用について、申し出られている研究内容を現時点で行うことについて合理的な理由があること。

21. ①―④のうち誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 匿名レセプト情報等を複製した情報システムの利用場所、保管場所及び管理方法は、あらかじめ申し出られた施錠可能な物理的なスペースに限定されており、原則として持ち出してはならない。
- ② 匿名レセプト情報等を複製した情報システムは、インターネット等の外部ネットワークに接続してはならない。
- ③ 提供された匿名レセプト情報等は、あらかじめ申し出られた取扱者のみが利用することとし、その他の者へ譲渡、貸与又は他の情報との交換等を行ってはならない。
- ④ 現に匿名レセプト情報等の提供を承諾された提供申出における担当者が、当該匿名レセプト情報等の利用を終了していない場合でも、新たな提供申出を行うことができる。

22. ①―④のうち誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 情報システム運用責任者の設置及び担当者（システム管理者を含む。）の限定を行わなければならない。所属機関が小規模な場合においても明確な規程を定めなければならない。
- ② 個人情報参照可能な場所においては、来訪者の記録・識別、入退を制限する等の入退管理を定めること。
- ③ 情報システムへのアクセス制限、記録、点検等を定めたアクセス管理規程を作成すること。

- ④ 個人情報の取扱いを委託する場合、委託契約において安全管理に関する条項を含めること。

23. ①—④のうち誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 匿名レセプト情報等が保存されている機器の設置場所及び記録媒体の保存場所には施錠しなければならない。
- ② 匿名レセプト情報等を物理的に保存している区画への入退管理を実施しなければならない。
- ③ 匿名レセプト情報等の消去にあたっては、専用ソフトウェア等を用い、復元不可能な形で行わなければならない。
- ④ 匿名レセプト情報等を利用する情報システムへのアクセスにおける取扱者の識別と認証を行わなければならない。

24. 結果の公表基準について、①—④のうち誤っているものを1つ選べ。すべて正しい場合は⑤を選べ。

- ① 原則として、公表される研究の成果物において患者等の数が10未満になる集計単位が含まれていないこと（ただし患者等の数が「0」の場合を除く。）
- ② 集計単位が市区町村の場合、人口2,000人未満の市区町村では、患者等の数が20未満になる集計単位が含まれないこと。
- ③ 原則として、公表される研究の成果物において年齢区分が、5歳毎にグルーピングして集計されていること。
- ④ 原則として、特定健診等情報にかかる受診者の住所地については、公表される研究の成果物における最も狭い地域区分の集計単位は2次医療圏又は市区町村とすること。

<解答と解説>

1. 正解④

NDB 活用の本来目的は医療費適用であり、それ以外は目的外利用という位置づけである。

2. 正解④

匿名レセプト情報・匿名特定健診等情報の提供は「高齢者の医療の確保に関する法律」に基づいている。

「高齢者の医療の確保に関する法律」

(医療費適正化計画の作成等のための調査及び分析等)

第十六条 厚生労働大臣は、全国医療費適正化計画及び都道府県医療費適正化計画の作成、実施及び評価に資するため、次に掲げる事項に関する情報（以下「医療保険等関連情報」という。）について調査及び分析を行い、その結果を公表するものとする。

一 医療に要する費用に関する地域別、年齢別又は疾病別の状況その他の厚生労働省令で定める事項

二 医療の提供に関する地域別の病床数の推移の状況その他の厚生労働省令で定める事項

2 保険者及び後期高齢者医療広域連合は、厚生労働大臣に対し、医療保険等関連情報を、厚生労働省令で定める方法により提供しなければならない。

3 厚生労働大臣は、必要があると認めるときは、都道府県及び市町村に対し、医療保険等関連情報を、厚生労働省令で定める方法により提供するよう求めることができる。

(国民保健の向上のための匿名医療保険等関連情報の利用又は提供)

第十六条の二 厚生労働大臣は、国民保健の向上に資するため、匿名医療保険等関連情報

(医療保険等関連情報に係る特定の被保険者その他の厚生労働省令で定める者(次条において「本人」という。))を識別すること及びその作成に用いる医療保険等関連情報を復元することができないようにするために厚生労働省令で定める基準に従い加工した医療保険等関連情報をいう。以下同じ。)を利用し、又は厚生労働省令で定めるところにより、次の各号に掲げる者であつて、匿名医療保険等関連情報の提供を受けて行うことについて相当の公益性を有すると認められる業務としてそれぞれ当該各号に定めるものを行うものに提供することができる。

一 国の他の行政機関及び地方公共団体 適正な保健医療サービスの提供に資する施策の企画及び立案に関する調査

二 大学その他の研究機関 疾病の原因並びに疾病の予防、診断及び治療の方法に関する研究その他の公衆衛生の向上及び増進に関する研究

三 民間事業者その他の厚生労働省令で定める者 医療分野の研究開発に資する分析その他の厚生労働省令で定める業務(特定の商品又は役務の広告又は宣伝に利用するために行

うものを除く。)

2 厚生労働大臣は、前項の規定による利用又は提供を行う場合には、当該匿名医療保険等関連情報を健康保険法第五十条の二第一項に規定する匿名診療等関連情報及び介護保険法第一百八条の三第一項に規定する匿名介護保険等関連情報その他の厚生労働省令で定めるものと連結して利用し、又は連結して利用することができる状態で提供することができる。

3 厚生労働大臣は、第一項の規定により匿名医療保険等関連情報を提供しようとする場合には、あらかじめ、社会保障審議会の意見を聴かなければならない。

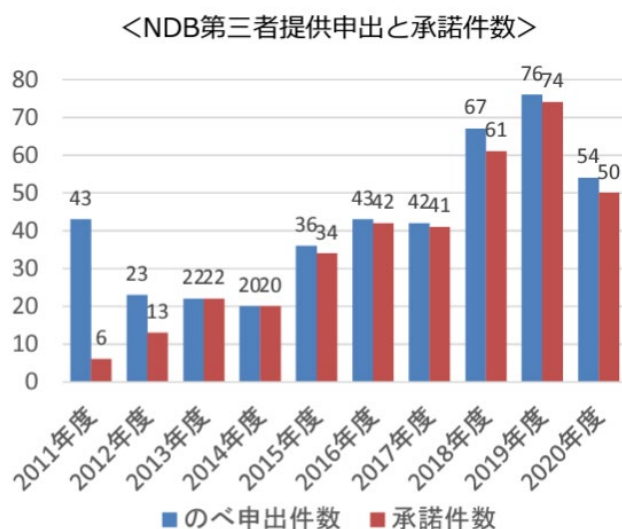
3. 正解②

匿名レセプト情報・匿名特定健診等情報のデータ保有者は厚生労働大臣である。

4. 正解①

「匿名医療情報等の提供に関する専門委員会」は社会保障審議会の下に設けられている。

5. 正解①



6. 正解④

平均所要日数は、2018年度は281.2日、2019年度は203.1日であった。

7. 正解④

「最終生成物」については、厚生労働省による公表前の事前の承認を得て成果物となったものを除き、取扱者以外に公表することを禁ずる。

8. 正解⑤

すべて正しい。

9. 正解①

特定個人又は特定機関の識別情報の削除

10. 正解⑤

すべて正しい。

11. 正解③

提供された匿名レセプト情報等1ファイルについて、当該ファイルを別の記憶装置に複写・保存する行為は1回に限定する。

12. 正解⑤

NDBは2009年に稼働開始し、2011年から「有識者会議」（現在は「専門家委員会」）の審査を経て、第三者の研究者等への提供が開始された。2020年10月の法改正により、民間事業者を含めた幅広い主体への提供が可能となった。

13. 正解④

大学の学長

14. 正解②

提供申出者が法人等の場合、提供申出書の提出日前6ヶ月以内に作成された登記事項証明書等を提出する。

15. 正解④

推測されるものは認めない。また、匿名レセプト情報等の提供の制度趣旨は国民保健の向上に資するといった相当の公益性を有することを求めるものであることを考慮し、他の研究や政策利用等を阻害するような特許の取得を禁止する。

16. 正解⑤

すべて正しい。

17. 正解④

匿名レセプト情報等の利用期間の上限は、原則として、2年間とする。

18. 正解④

ガイドラインに掲げる補助金等を充てて匿名レセプト情報等を用いて研究又は業務を行う者は手数料を免除される。手数料の免除を希望する場合は、補助金等の交付決定通知の写し及び研究計画書又は交付申請書を添付する。免除申請は、提供申出時から、厚生労働省が提供申出者に手数料額を通知する時までとする。

19. 正解⑤

すべて正しい。

20. 正解⑤

すべて正しい。

【審査基準】

①利用目的

レセプト情報等の利用目的は、医療サービスの質の向上等を目指した施策の推進や、学術の発展に資する研究に資するものであるか

②利用の必要性

利用するレセプト情報の範囲が利用目的に照らして必要最小限であるか、レセプト情報の性格に鑑みて情報の利用が合理的か

③研究内容の実行可能性

研究計画の内容は、申出者の過去の研究実績や人的体制に照らして実行可能であるか

④セキュリティ

適切な措置（レセプト情報等を複製した情報システムを外部ネットワークに接続しない、個人情報保護に関する方針の策定・公表、外部委託契約における安全管理条項の有無等）を講じているか

⑤結果公表等

学術論文等の形で研究成果が公表される予定か、施策の推進に適切に反映されるか等

21. 正解④

現に匿名レセプト情報等の提供を承諾された提供申出における担当者が、当該匿名レセプト情報等の利用を終了していない場合については、新たな提供申出を行うことは原則認められない。

22. 正解①

所属機関が小規模な場合において役割が自明の場合は、明確な規程を定めなくとも良い。

23. 正解⑤

すべて正しい。

24. 正解②

人口2,000人未満の市区町村では、患者等の数を表示しないこと。

人口2,000人以上25,000人未満の市区町村では、患者等の数が20未満になる集計単位が含まれないこと。

人口25,000人以上の市区町村では、患者等の数が10未満になる集計単位が含まれないこと。

SQL入門

東京大学大学院医学系研究科糖尿病・生活習慣予防講座
岡田 啓

SQL概論

SQLをなぜ使うか？①

○データテーブルを自分で構築する

- ・テーブルが複数あるようなデータの場合、自分で解析用のテーブルを作る必要がある
- ・商用データベース（DeSCデータベースなど）、ナショナルレセプトデータベース(NDB)などで研究を行う際には自分で、解析のためのテーブルを作成する必要がある。

○統計ソフトではなぜ不十分か？

SQLをなぜ使うか？②

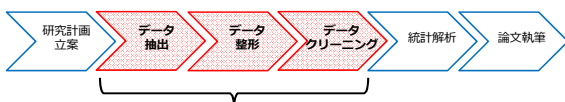
○データベースサイズの拡大

- ・ビッグデータが更にhugeに変遷：MB単位→GB単位のデータベースへ

○統計ソフトではデータクリーニングが難しい

- ・統計ソフトはメモリ上にデータを載せるので数十GBのデータだと載せるのが困難
- ・統計ソフトでは、結合などに時間がかかり、そもそもhugeデータには不向き

データ抽出・クリーニングでSQLは有用



SQLが必須の場合もそうではない場合もSQLを使った方が圧倒的に処理が早い

← どちらにしてもSQLを身につけるメリットは大きい

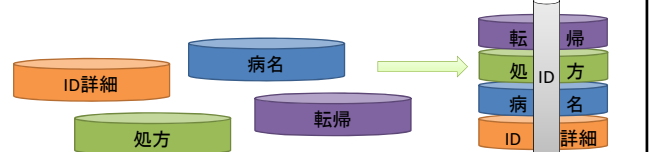
リレーショナルデータベース(Relational DataBase)

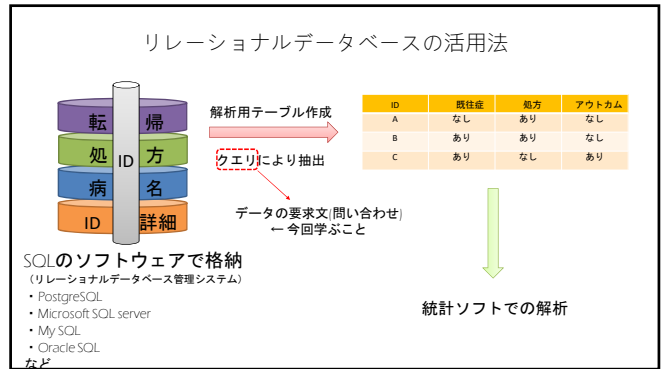
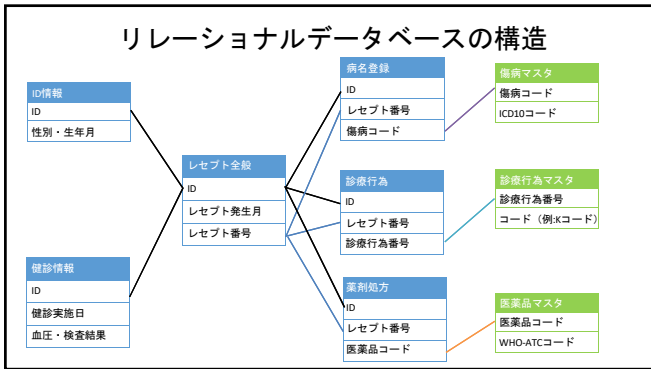
- ・データベース：Data + Base なので実際扱うデータの源

ID	既往症	処方	アウトカム
A	なし	あり	なし
B	あり	あり	なし
C	あり	なし	あり

- ・リレーショナルデータベース：

ID/コードに関連 (relational) させて複数のデータベースに情報を格納





ここまでのまとめ：SQL導入

①SQLをなぜ使うか？
→ データサイズの大きいデータベースは統計ソフトでは扱うのが難しい

②リレーショナルデータベース
データテーブル同士が、キーIDや病名番号などで結合されたデータベース

SQL入門①

データ外観、SELECT文、条件設定まで

注意：
・これから挙げるクエリ例は、すべてMicrosoft SQL Serverの記述方法であり、ほかのソフトウェアでは方言が異なることにご注意ください。
・また、同じ結果を得るためのクエリの書き方は、必ずしも1通りではなく、わかりやすさを優先して例を載せているので、各自でもっとも適したと思われるクエリを探してみてください。

それぞれのデータテーブルを眺める

データベース内のテーブル名を右クリック
「上位1000行の選択」をクリック

出現したクエリ
SELECT TOP 1000 *
FROM [データシート名]

→ 右下に結果が表示される

SELECT文①基本、一般

基本クエリ

```
SELECT [変数名1], [変数名2], ..., [変数名n]
FROM [データシート名]
```

ポイント

- * SELECTの後の変数名は、半角のカンマ「,」で繋ぐ
- * SELECT A FROM Bという語順に慣れる

参考：

- ①上から数えてN行だけ表示したいときはTOP 1000などと変数名の前に置く
- ②全ての変数を表現したいときは半角のアスタリスク「*」で可能

SELECT文②：基本、具体

(1) IDと性別だけのテーブルが欲しい

```
SELECT [変数名1], [変数名2], ..., [変数名n]
FROM [データシート名]
```

ID	年齢	性別
A	30	0
B	50	1
C	87	0
D	4	1

```
SELECT [ID], [性別]
FROM [TEST_TABLE]
```

ID	性別
A	0
B	1
C	0
D	1

SELECT文③：条件

(1) 条件を満たす (例:40歳以上) テーブルが欲しい

```
SELECT [変数名1], [変数名2], ..., [変数名n]
FROM [データシート名]
```

ID	年齢	性別
A	30	0
B	50	1
C	87	0
D	4	1

```
SELECT [ID], [性別]
FROM [TEST_TABLE]
WHERE [年齢]>=40
```

ID	年齢	性別
B	50	1
C	87	0

ポイント

* 条件はWHEREでFROMよりも後ろに記述

SELECT文④：並べ替え

(2) 並べ替えた (例:年齢で降順) テーブルが欲しい

```
SELECT [変数名1], [変数名2], ..., [変数名n]
FROM [データシート名]
```

ID	年齢	性別
A	30	1
B	50	1
C	87	0
D	4	1

```
SELECT [ID], [年齢], [性別]
FROM [TEST_TABLE]
ORDER BY [年齢] DESC
```

ID	年齢	性別
C	87	0
B	50	1
A	30	1
D	4	1

ポイント

* 並べ替え指定はORDER BYでFROMよりも後ろに記述
* 降順はDESC、昇順はASCで変数の後ろに記載

SELECT文⑤：重複の削除

同じ情報を削除して表示する

```
SELECT DISTINCT * FROM [データシート名]
```

ID	年齢	性別
A	30	1
B	50	1
C	87	0
D	4	1
D	4	1

ID	年齢	性別
A	30	1
B	50	1
C	87	0
D	4	1

SELECT DISTINCT ID FROM [データシート名]だとどうなる？

ID
A
B
C
D

SELECT DISTINCT 性別 FROM [データシート名]だとどうなる？

性別
1
0

条件の指定①LIKE/IN

☆データ抽出条件に用いる

```
SELECT * FROM [tablex]
WHERE [職業] LIKE '基礎研究者' OR [職業] LIKE '疫学研究者' /*☆☆/
```

ID	年齢	職業
A	30	基礎研究者
B	50	疫学研究者
C	87	遊び人
D	4	無職



ID	年齢	職業
A	30	基礎研究者
B	50	疫学研究者

ポイント

・複数の条件はきちんと変数名から書く (同じものでも省略不可)
☆LIKEの具体例を並べるときは下記のようにINを使っても書ける
WHERE [職業] IN ('基礎研究者', '疫学研究者')
☆LIKEの後に部分一致を使いたいときは%を用いて%研究者%とすればよい (後方一致なら2回目の%は不要)

条件の指定②BETWEEN

☆データ抽出条件に用いる

```
SELECT * FROM [tablex]
WHERE [年齢] BETWEEN 30 AND 50 /*☆☆/
```

ID	年齢	職業
A	30	基礎研究者
B	50	疫学研究者
C	87	遊び人
D	4	無職



ID	年齢	職業
A	30	基礎研究者
B	50	疫学研究者

ポイント

・BETWEEN [下限値] AND [上限値]で2つの数字の間という条件を作る
← 上限と下限を逆にすると誤った条件と考えられ不可
(クエリは流れ、赤字のエラーはでないが、うまくいかない)
☆本表記は下記のように2条件の組み合わせでも書くことができる
WHERE [年齢] >=30 AND [年齢] <=50

集計値表示①MAX/MIN/AVG/SUM

☆データの集計値を出す (ほかに標準偏差, STDEV)

SELECT MAX(BMI) as BMI最大値,

MIN(BMI) as BMI最小値,

AVG(BMI) as BMI平均値

SUM(BMI) as BMI合計値

FROM [tablex]

ID	BMI	性別
A	30	男
B	23	男
C	22	男
D	18	女
E	25	女
F	21	女

→

BMI最大値	BMI最小値	BMI平均値	BMI合計値
30	18	22.5	139

集計値表示②GROUP BY

☆層ごと (ここでは性別) のデータの集計値を出す

SELECT MAX(BMI) as BMI最大値,

MIN(BMI) as BMI最小値,

AVG(BMI) as BMI平均値

FROM [tablex]

GROUP BY 性別

ID	BMI	性別
A	30	男
B	23	男
C	22	男
D	18	女
E	25	女
F	21	女

→

性別	BMI最大値	BMI最小値	BMI平均値
男	30	22	25
女	25	18	22

テーブルの保存と削除

・保存

SELECT *

INTO [保存したいテーブル名] FROM [既存のTable]

注) 基本的には絶対パスで記述

注) テーブル名を#で始めると一時的なテーブルとして保存される
一時的なテーブルとはタブを閉じると消去されるテーブル

・削除

DROP TABLE [削除したいテーブル名]

注) 一旦落とすと戻せないから注意が必要

ここまでのまとめ: SQL入門①

①基本構文を理解

SELECT [変数名1], [変数名2], [変数名3], ..., [変数名N]

INTO [保存したいテーブル名] FROM [既存のTable]

②条件や並べ替えの書き方の理解

条件はWHEREを使って記述

並べ替えはORDER BYを使って記述

} 共に FROM [テーブル名]の後ろに記述

②集計値記述

GROUP BYを使った記述方法

SQL入門②

変数作成と日付関数

変数作成①: 新変数の作成(1)

新変数 (alive: 全て1) を追加する

SELECT *, 1 as alive

INTO #1 FROM [データシート名]

ID	年齢	性別	alive
A	30	1	1
B	50	1	1
C	37	0	1
D	4	1	1

STATA:

generate alive = 1

R:

tablex %>% mutate(alive = 1)

変数作成②：新変数の作成(2)

新変数 (adult: 20歳以上が1、それ以外は0) を追加する
 SELECT *, case when 年齢 >= 20 then 1 else 0 end as adult
 INTO #1 FROM [データシート名]

ID	年齢	性別	adult
A	30	1	1
B	50	1	1
C	87	0	1
D	4	1	0

STATA:
 generate adult = (年齢 >= 20)
 もしくは
 generate adult = 1
 replace adult = 0 if 年齢 < 20

R:
 tablex %>% mutate(adult = (年齢 >= 20)) %>%
 as.integer()
 もしくは
 tablex %>% mutate(adult = ifelse((年齢
 >= 20), 1, 0))

変数作成③：新変数の作成(3)

新変数 (generation: 18未満をyoung、18-64をmiddle、65以上をolder) を追加する

SELECT *, case when 年齢 < 18 then 'young' when 年齢 >= 18 and 年齢 < 65 then 'middle' when 年齢 >= 65 then 'older' else null end as generation
 INTO #1 FROM [データシート名]

ID	年齢	性別	generation
A	30	1	middle
B	50	1	middle
C	87	0	older
D	4	1	young

ポイント
 * 複数条件はwhenを繰り返す
 * 最後のendを忘れないように!

変数作成④：nullの扱い

☆欠測はNULLで表される (数字変数でも文字列扱いだが'は付けない)
 SELECT *, INTO #1 FROM [データシート名]

WHERE 年齢 is null /*年齢=NULLは不可*/

ID	年齢	性別
A	30	1
B	50	1
C	87	0
D	NULL	1

SELECT *, INTO #1 FROM [データシート名]

WHERE 年齢 is not null /*年齢=NULLは不可*/

ID	年齢	性別
A	30	1
B	50	1
C	87	0
D	NULL	1

変数作成⑤：nullを置き換える

(1) ISNULLを使う場合

SELECT *, ISNULL(年齢, 0) as c年齢 INTO #1 FROM [データシート名]

ID	年齢	性別	c年齢
A	30	1	30
B	50	1	50
C	87	0	87
D	NULL	1	0

(2) CASE WHENを使う場合

SELECT *, CASE WHEN 年齢 is null then 0 else 年齢 end as c年齢 INTO #1 FROM [データシート名]

ID	年齢	性別	c年齢
A	30	1	30
B	50	1	50
C	87	0	87
D	NULL	1	0

変数作成⑥：DATEDIFF関数を用いた年齢作成

☆DATE関数を用いて日付と認識させるところから (StataやRと同様)

ID	誕生日	入院年月日	年齢
A	19750105	20191001	44
B	19611221	20200502	58
C	19871029	20220819	35

☆SQL自身が日付と変数を認識している場合
 SELECT *, DATEDIFF(YEAR, [誕生日], [入院年月日]) as 年齢 into #tableX1
 FROM [tableX]

(注) 月齢ならYEAR → MONTH、日齢ならYEAR → DAYにすればよい

☆ [年月日] が日付認識されずエラーが生じる場合
 日付認識されていない場合は[誕生日]のところに
 CONVERT(date, [誕生日], 112)
 とすると日付に一旦変換される

この「112」はYYYYMMDDで表記された数字を日付として読み込むための番号
 ちなみに、YYYY/MM/DDを読み込むなら111 (詳細は成書参照)

ここまでのまとめ：SQL入門②

①変数作成を理解

(1) 新変数の場合

SELECT 変数の値 as [新変数名] FROM [既存のTable]

(2) 既存値を用いた変数の場合

SELECT [旧変数名] as [新変数名] FROM [既存のTable]

②変数作成に役立つ関数

case when [条件] then [値1] else [値2] end as [新変数名]

②日付計算

DATEDIFFを使った計算方法

SQL入門③

データ結合

テーブル結合①：導入

☆SQLを用いる一番のメリット→ データ結合の容易さ
例：年齢しかないテーブルに性別という情報を追加したい

ID	年齢
A	87
B	50
D	4

Table X

ID	性別
A	0
B	1
C	0

Table Y

行うこと
IDをキーにして情報を結合
基本クエリ

```
SELECT X.*, Y.性別
FROM [データシート1] X
LEFT JOIN
[データシート2] Y
ON X.ID=Y.ID
注) 元々対応する
```

ID	年齢	性別
A	87	0
B	50	1
D	4	NULL

Table XにYの性別を加える

テーブルの結合②：概略

内部結合INNER JOIN

ID	年齢
A	87
B	50
D	4

Table X

ID	性別
A	0
B	1
C	0

Table Y

ID	年齢	性別
A	87	0
B	50	1

両方のテーブルに存在するIDのみ結合

外部結合OUTER JOIN

ID	年齢
A	87
B	50
D	4

Table X

ID	性別
A	0
B	1
C	0

Table Y

左外部結合 LEFT JOIN

基準が左

ID	年齢	性別
A	87	0
B	50	1
D	4	NULL

左にあるIDのみ情報追加

右外部結合 RIGHT JOIN

基準が右

ID	性別	年齢
A	0	87
B	1	50
C	0	NULL
D	4	NULL

右にあるIDのみ情報追加

全外部結合 FULL JOIN

基準は両方

ID	年齢	性別
A	87	0
B	50	1
C	NULL	0
D	4	NULL

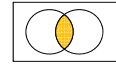
どちらかにあるIDに情報追加

テーブルの結合③：概念

内部結合INNER JOIN

ID	年齢	性別
A	87	0
B	50	1

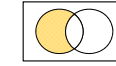
両方のテーブルに存在するIDのみ結合



左外部結合 LEFT JOIN

ID	年齢	性別
A	87	0
B	50	1
D	4	NULL

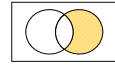
左にあるIDのみ情報追加



右外部結合 RIGHT JOIN

ID	性別	年齢
A	0	87
B	1	50
C	0	NULL

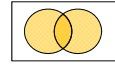
右にあるIDのみ情報追加



全外部結合 FULL JOIN

ID	年齢	性別
A	87	0
B	50	1
C	NULL	0
D	4	NULL

どちらかにあるIDに情報追加



テーブル結合④：具体例

ID	年齢
A	87
B	50
D	4

Table X

ID	性別
A	0
B	1
C	0

Table Y

```
実際のクエリ
SELECT X.*, Y.性別 /*Xの全てを、Yから性別だけを取る①*/
FROM [データシート1] X /*シートをXと名付ける②*/
INTO [データシート3] /*intoの位置はここ③*/
LEFT JOIN
[データシート2] Y /*Yと名付ける④*/
ON X.ID=Y.ID /*XとYのIDをキー変数と宣言④*/
```

注) ① XとYの両方にある変数を両方選ぶとエラーが出る
(1) どちらから選んでくるのか不明
(2) 同じ変数名は同じテーブルには存在できない
② 名付けるのはYではなくてもX/Bなどとしてもよい
③ INTOの位置に注意
④ キー変数はIDのみならず、二つ以上も可能
その場合はON X.ID=Y.ID AND X.AGE=Y.AGE
などとす

ID	年齢	性別
A	87	0
B	50	1
D	4	NULL

Table XにYの性別を加える

テーブル結合⑤：注意点！

☆結合されるキーを意識しないと、データ量が膨大になる

ID	性別	職業	処方月	薬剤
A	男	タマネギ剣士	201401	グラルギン
A	男	タマネギ剣士	201401	リスプロ
A	男	黒魔道士	201402	カンデサルタン
A
A	男	運び人	202006	エンハグワロジソ
B	女	語り子	201401	アスピリン
...
C	女	賢者	202206	クロビドグレル

Table X

ID	病名登録月	病名
A	201401	糖尿病
A	201402	高血圧
A
A	202006	糖尿病
B	201401	脳梗塞
...
C	202206	心筋梗塞

Table Y

```
SELECT X.*, Y.病名登録月, Y.病名
FROM [データシート1] X
INTO [データシート3]
LEFT JOIN
[データシート2] Y
ON X.ID=Y.ID
```

結果の行数は何行でしょうか？
Aさんだけでも、

A	男	タマネギ剣士	201401	グラルギン
...

に対して、Table Yの2014~2020までの全病名登録月・全病名が結合されるのでデータ量は少なくとも数十倍以上

テーブル結合⑥：問題解決へ(1)

☆結合されるキーを意識してクエリを作る！

ID	性別	職業	処方月	薬剤
A	男	タマネギ剣士	201401	グラルギン
A	男	タマネギ剣士	201401	リスプロ
A	男	黒龍道士	201402	カンデサルタン
A
A	男	迷ひ人	202006	エンハグリアロジン
B	女	踊り子	201401	アスピリン
...
C	女	賢者	202206	クロビドグレル

Table X

ID	病名登録月	病名
A	201401	糖尿病
A	201402	高血圧
A
A	202006	糖尿病
B	201401	脳梗塞
...
C	202206	心筋梗塞

Table Y

☆解決策②
結合されるキーを複数にする
SELECT distinct X.ID, X.性別, X.処方月,
Y.病名登録月, Y.病名
FROM #X X
LEFT JOIN
#Y Y
ON X.ID=Y.ID AND X.処方月=Y.病名登録月

☆解決策①

Table X, Yともに繋げる前にシンプルにしておく
例：select distinct ID, 性別 into #X1 from #X
select distinct ID, 病名 into #Y1 from #Y
それぞれ作ってから、結合する

テーブル結合⑦：問題解決へ(2)

☆よく生じる問題 → 結合で同一のものが複数行生成

ID	性別	処方月	薬剤	病名
A	男	201401	グラルギン	糖尿病
A	男	201401	グラルギン	高血圧
A	男	201401	グラルギン	心筋梗塞
A	男	201401	リスプロ	糖尿病
A	男	201401	リスプロ	高血圧
A	男	201401	リスプロ	心筋梗塞
...

☆事前の解決策

- ① 結合するテーブルをそれぞれ、繋げる前にシンプルにしておく（既出）
- ② 上位XX行（例えば1000行）だけで実行してみる

☆事後の解決策：統計ソフトと同じく結果は必ず確認すること

- ① 結果の (XXXX行 処理されました)の「XXXX」が欲しい行数か確認
- ② 同じIDが複数あるかを確認、distinctなどを使い重複のIDの有無を確認
SELECT count([ID]) from TableZ
SELECT count(distinct [ID]) from TableZ

テーブル結合⑧：データのappend（参考）

☆データのappendはUNIONを使う

ID	病名登録月	病名
A	201401	糖尿病
A	201402	高血圧
A
A	202006	糖尿病
B	201401	脳梗塞
...
C	202206	心筋梗塞

Table X

ID	病名登録月	病名
D	201401	回廊リウマチ
D	201402	ステロイド糖尿病
D
D	202006	IgA腎症
E	201401	真性多血症
...
F	202206	多発性囊胎腎

Table Y

SELECT ID, 病名登録月, 病名
FROM TableX
UNION
SELECT ID, 病名登録月, 病名
FROM TableY

ここまでのまとめ：SQL入門③

☆データ結合、特に左外部結合LEFT (OUTER) JOINを理解

SELECT X.*, Y.性別 /*Xの全てを、Yから性別だけを取る①*/
FROM [データシート1] X /*シート1をXと名付ける②*/
INTO [データシート3] /*intoの位置はここ③*/
LEFT JOIN
[データシート2] Y /*Yと名付ける②*/
ON X.ID=Y.ID /*XとYのIDをキー変数と宣言④*/

ID	年齢
A	87
B	50
D	4

Table X

ID	性別
A	0
B	1
C	0

Table Y



ID	年齢	性別
A	87	0
B	50	1
D	4	NULL

Table XにYの性別を加える

本日のまとめ：SQL入門

- ①ビッグデータ時代におけるSQL活用の必要性を説明出来る
- ②SELECT文を用いて、データテーブルの内容や要約量計算を表示できる
- ③外部結合を用いて、データテーブルの結合出来る

ご清聴ありがとうございました

NDB・DPCデータベース研究人材育成Webinar

SQLを用いた レセプトデータのハンドリング

東京大学大学院医学系研究科臨床疫学・経済学
松居宏樹

この授業では

我が国における代表的かつ重要な大規模医療リアルワールドデータベースである、**レセプト情報・特定健診等情報データベース(NDB)**の成り立ち、構造、利活用事例について学ぶ。そのうえで、データ（特にレセプト情報部分）について、その**データハンドリング手法を実際にデータを触りながら学ぶ。**

学習サイト

学習を円滑に進めるため、一時的なWebサイトを公開しました。

<https://sites.google.com/m.u-tokyo.ac.jp/ndb-study/>



はじめに

NDB とは何か

- 正式名称：レセプト情報・特定健診等情報 データベース
- 利用目的は全国医療費適正化計画及び北海道府県医療費適正化計画の作成、実施及び評価に資するため（高齢者の医療の確保に関する法律 第16条）
- 保有：厚生労働大臣
- 収集元：審査支払機関
- 内容
 - レセプトデータ 年間約18億7000万件（H28年度）
 - 特定健診・保健指導データ 年間約2,730万件（H28年度）
 - Administrative Claims Database の一種
- 窓口など：
https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryuuhoken/reseputo/index.html

ここ最近の研究成果

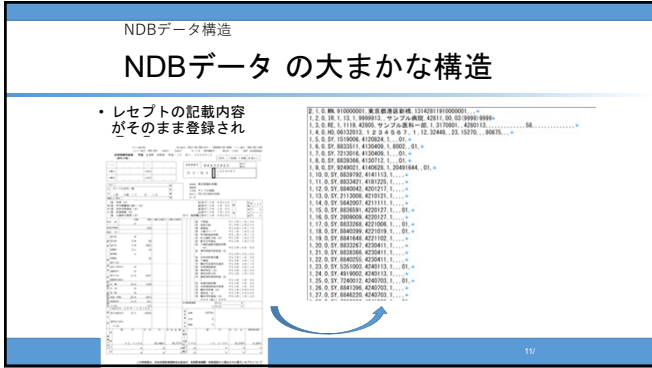
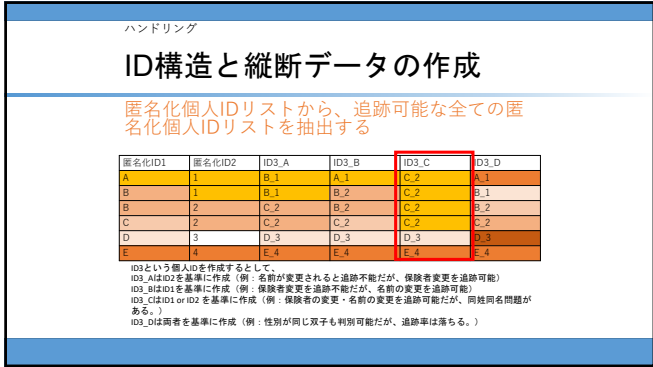
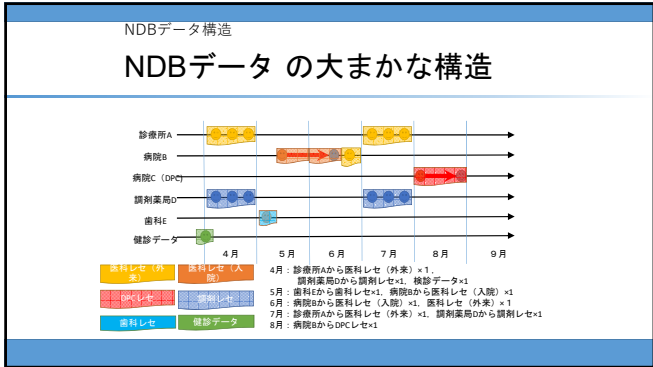
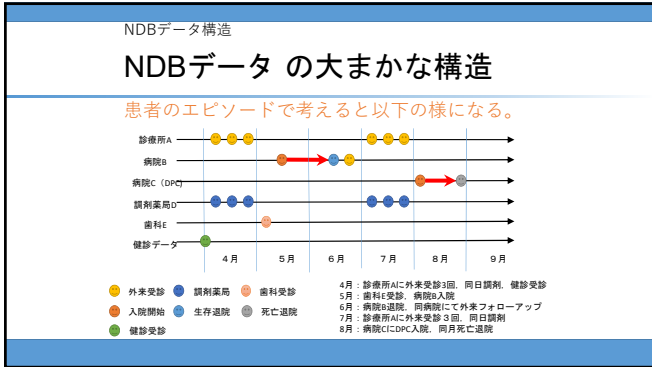
- Ishimaru M, Matsui H, Ono S, Hagiwara Y, Morita K, Yasunaga H. Preoperative oral care and effect on postoperative complications after major cancer surgery. *British Journal of Surgery*. 2018 Oct 12;105(12):1868-96.
- Takeuchi Y, Kumamaru H, Hagiwara Y, Matsui H, Yasunaga H, Miyata H, et al. Sodium-glucose cotransporter-2 inhibitors and the risk of urinary tract infection among diabetic patients in Japan: Target trial emulation using a nationwide administrative claims database. *Diabetes Obes Metab*. 2021 Jun;23(6):1379-88.
- Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H. Prevalence, Incidence Rate, and Risk Factors of Medication-Related Osteonecrosis of the Jaw in Patients With Osteoporosis and Cancer: A Nationwide Population-Based Study in Japan. *Journal of Oral and Maxillofacial Surgery*. 2022 Apr 1;80(4):714-27.
- Hasegawa Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, et al. Incidence of sympathetic ophthalmia after inciting events: a national database study in Japan. *Ophthalmology*. 2021 Sep 21;50161-6420(21)00719-3.
- Kasajima M, Eggleston K, Kusaka S, Matsui H, Tanaka T, Son BK, et al. Prevalence of frailty and dementia and the economic cost of care in Japan from 2016 to 2043: a microsimulation modelling study. *The Lancet Public Health*. 2022;7(5):e458-68.

NDBデータ構造

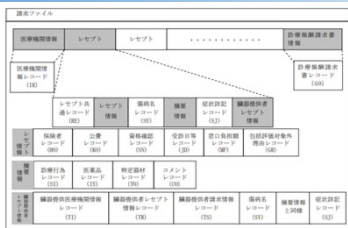
NDBデータの大まかな構造

レセプト部分と健診情報部分がある。また、介護情報との接続も可能に

- 医科(MED)
 - DPC(DPC)
 - 調剤(PHA)
 - 歯科(DEN)
 - 特定健診データ
 - 保健指導データ
 - 介護レセデータ
- レセプト情報
- 健診情報
- 介護情報
- 今回の演習では主として医療レセプトのデータハンドリングを学ぶ



レセプトデータの構造化



レセプトデータの構造化

演習：データを検索してみる

社会保険診療報酬支払基金よりサンプル医科レセプトをダウンロード

https://www.ssk.or.jp/kyouyaku/kyouyaku_03/index.html

.medsample/サンプル1 /0_COMMON001_MED/ 11_REC0DINFO_MED.CSV

【サンプルレセプトCSV】
 → 処方済薬 (E01)
 → 処方済薬 (E02)
 → 処方済薬 (E03)
 → 処方済薬 (E04)
 → 処方済薬 (E05)
 → 処方済薬 (E06)

※ (サンプルデータ) 処方済薬 (E01) 処方済薬 (E02) 処方済薬 (E03) 処方済薬 (E04) 処方済薬 (E05) 処方済薬 (E06)

レセプトデータの構造化

演習：データを検索してみる

- CSV形式のレセプトをエクセルで開く
 *.medsample\0_COMMON001_MED\11_REC0DINFO_MED.CSV
- サンプルデータをレセプト単位に切り分ける
MNレコードが切れ目
- サンプルデータから、医科レセプトの医薬品処方履歴を集める。
- "サンプル "さんの"尿素クリーム10% 「フジナガ」"の処方日は？

レセプトデータの構造化

- レセプトデータは生のままでは扱いにくい。
- 行ごとに書いてある情報が違う
- **行の種類ごとに分けてデータベース化しよう。**

実際の研究とデータ構造

構造化されたデータの研究利用

NDBのテーブル構造はあまり大げらにされていない。

[NDBの提供について\(厚労省\)](#)のサイト内にある、抽出依頼テンプレートに詳細がある。

- 抽出依頼テンプレート：当方と相談の上、必要に応じて作成してください
 - ▶ 抽出依頼テンプレート(の注意) [XLS形式:1,210KB]
 - ▶ 抽出依頼テンプレート(抽出) [XLS形式:1,720KB]
 - ▶ 抽出依頼テンプレート(集計) [XLS形式:1,530KB]

DB仕様は機械判読可能な形で

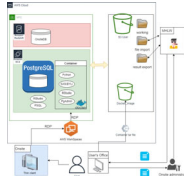
エクセル内に提供ファイルのカラム型情報あり

項目名	データ項目名(日本語)	項目名(英語)	型	長さ	出力	項目仕様	案件毎記入欄
1	レセプト番号	RES_NO	文字	8			
2	レセプト日	RES_DT	文字	10			
3	処方済薬	RES_DRUG	文字	15			
4	処方量	RES_QTY	数字	1			
5	処方単位	RES_UNIT	文字	2			
6	処方済薬	RES_DRUG	文字	15			
7	処方量	RES_QTY	数字	1			
8	処方単位	RES_UNIT	文字	2			
9	処方済薬	RES_DRUG	文字	15			
10	処方量	RES_QTY	数字	1			
11	処方単位	RES_UNIT	文字	2			
12	処方済薬	RES_DRUG	文字	15			

特別抽出仕様NDB データベース

postgresql 上にサンプルレセプトを基にレセプトデータベースを作成

- 特別抽出のデータ仕様に合わせたデータベースを構築
 - テーブル名・カラム名は先述の仕様を基に決定
 - サンプルレセプトをデータベースに格納
- データベースソフトウェアはpostgresql ver.14.6を用いた。
 - オンサイトセンターでの採用
 - Redshiftなどの種々のクラウドマネージドデータベースと構文が似ている。
 - オープンソースソフトウェアで導入しやすい。



NDBデータ構造

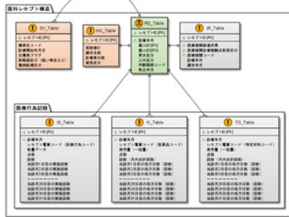
NDBデータ の大まかな構造

- 複数のテーブル(右図)で1レセの情報がなる。
- 匿名化個人IDは各REに含まれる。匿名化ID1と匿名化ID2の2つ
- 他のレセプト(MED, DPC, PHA, DEN)情報は(や健診情報)とは匿名化個人IDで接続可能
- 月単位で登録されるため、月またぎ入院や月内外来受診の処理が煩雑



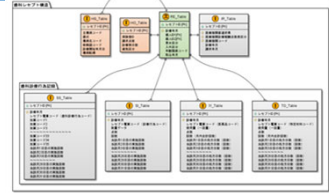
内科レセプトの情報

- 医療機関毎・診療年月毎・入外毎にRE(レセプトヘッダ)が発生
- REに紐づく形で全18テーブル
- 研究に主として用いるテーブルは概ね7種類



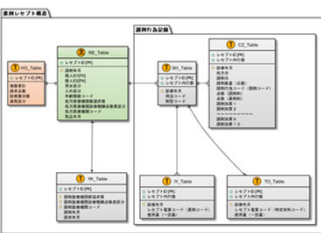
歯科レセプトの情報

- 医療機関毎・診療年月毎・入外毎にRE(レセプトヘッダ)が発生
- REに紐づく形で全15テーブル
- 研究に主として用いるテーブルは概ね8種類



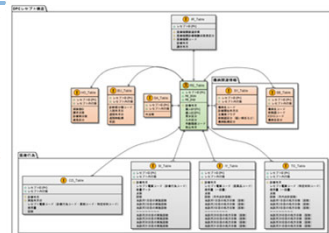
調剤レセプトの情報

- 調剤機関毎・調剤年月毎にREが発生
- REに紐づく形で全14テーブル
- 研究に主として用いるテーブルは概ね7種類
- 薬剤の処方状況の確認のため、処方レセプトの行数を基に接続が必要



DPCレセプトの情報

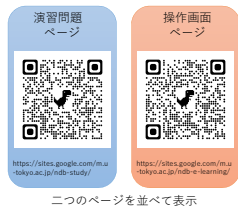
- 医療機関毎・診療年月毎に複数のREが発生
- REとその出現行数に紐づく形で全26テーブル
- 研究に主として用いるテーブルは概ね11種類



演習環境の説明

以降は構築したデータベースを用いた演習を行う

- この演習では、オープンソース PostgreSQL Webassemblyである、`postgres-wasm` (<https://github.com/snaple/postgres-wasm>) を用いて、参加者のブラウザ上に小規模な PostgreSQL サーバーを立ち上げています。
- ブラウザを閉じると、構築された環境が消去されます。
- 演習問題ページとデータベース操作画面を並べて表示してください。



演習終了後のまとめ

今回触れなかった内容

- 環境の整備
 - postgresql サーバーのインストール (<https://www.postgresql.jp/>)
 - DB構築用DDL
 - CSVデータのデータベースへの取り込み
 - データの圧縮やインデックスを用いた効率的なクエリに書き方
- レセプトデータを整理した後の疫学研究
- アンケートにご協力ください。

NDBを用いた研究

東京大学大学院医学系研究科臨床疫学・経済学
佐藤壮

目次

1. リアルワールドデータ研究のレビュー
2. 研究紹介

1

目次

1. リアルワールドデータ研究のレビュー
2. 研究紹介

2

背景

- リアルワールドデータを二次利用した研究は近年増えているが、その具体的な程度については不明である。
- リアルワールドデータを使用した研究を包括的にレビューした論文はHiroseのNDB研究(ACE 2020;2:13-26)に関するもののみである。
- リアルワールドデータを使用した研究の全体としての傾向や、それぞれのデータベース研究での研究分野等を知ることは今後の研究を行う上で有用である。

3

【参考】先行NDB研究レビュー¹の要約

【方法】

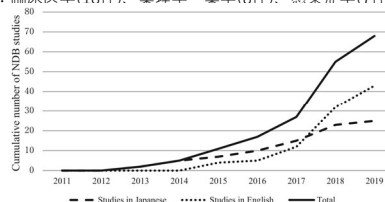
- 情報源：PubMed, 医中誌Web, 厚生労働省社会保障審議会(医療保険部会 匿名医療情報等の提供に関する専門委員会)資料
- 検索期間：2011/1 - 2019/6
- 検索対象：NDB, NDB Open Dataを使用した研究の原著論文
 1. 研究デザイン
 2. 研究分野
 3. 設定、サンプル数
 4. 研究結果
 5. 研究のstrength, limitation

4

【参考】先行NDB研究レビューの要約

【結果】

- 最終的なレビュー対象：英語論文43件、日本語論文25件→合計**68件**
- 研究デザイン：記述研究42件、横断研究14件、後ろ向きコホート研究12件
- 研究分野：臨床医学(18件)、薬理学・薬学(8件)、感染症学(7件)を含む36分野



5

今回レビュー対象としたリアルワールドデータ

- NDB (National Database of Health Insurance Claims and Specific Health Checkups of Japan)
- NDB Open Data
- JMDC Claims Database
- DPC (Diagnosis Procedure Combination) database

6

レビュー方法 (NDB, NDB Open Data)

- 検索エンジン：PubMed
- 検索期間：2013/1/1 – 2022/7/31
- 検索ワード：(((claims and NDB) OR (National Database of Health Insurance Claims) OR (National Database of Health Insurance Claim) OR (National Database of Japanese Health Insurance Claims) OR (nationwide administrative claims database)) AND (Japan OR Japanese)) AND (2013:2022[pdat]) NOT (Korea) NOT (Taiwan)

7

レビュー方法 (JMDC Claims Database)

- 検索エンジン：PubMed
- 検索期間：2010/1/1 – 2022/7/31
- 検索ワード：(("JMDC") OR ("jmdc database") OR ("Japan Medical Data Center")) AND (Japan OR Japanese)

8

レビュー方法 (DPC database)

- 検索エンジン：PubMed
- 検索期間：2010/1/1 – 2022/7/31
- 検索ワード：
(("Diagnosis Procedure Combination database"[Title/Abstract]) OR ("nationwide database"[Title/Abstract]) OR ("administrative database"[Title/Abstract]) OR ("inpatient database"[Title/Abstract]) OR ("discharge database"[Title/Abstract])) AND (Japan OR Japanese)

9

レビュー方法 (全体)

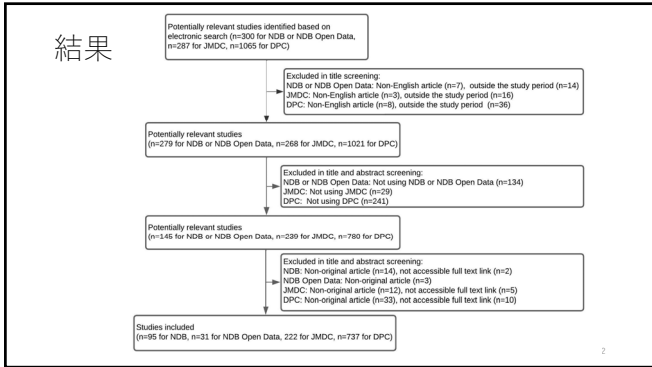
- 組み入れ基準
 1. 原著論文かつNDB, NDB Open Data, the JMDC Claims Database, the DPC databaseのいずれかを使用した研究であること
 2. 英語論文であること
- 除外基準
 1. letters, notes等のoriginal article以外のもの
 2. PubMedからFull textにアクセスできない場合

10

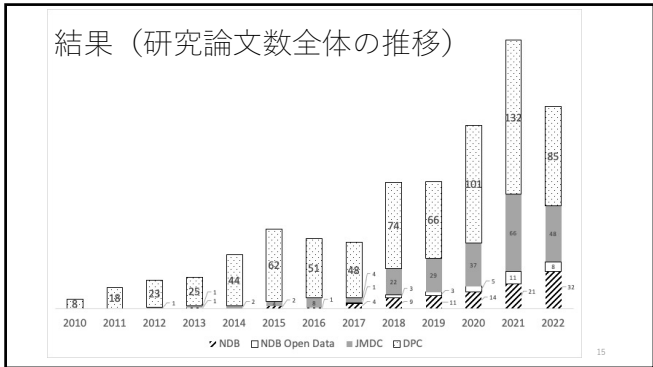
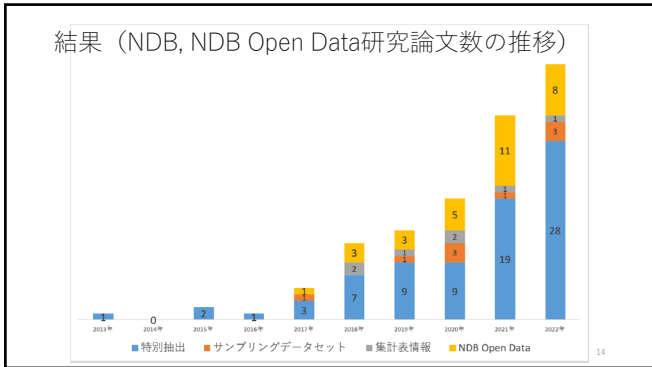
レビュー方法 (全体)

- 抽出したデータ
 1. 主に使用されたデータベース
 2. 研究論文が雑誌に掲載された時期 (年)
 3. 研究分野 ^(注1)

11



- ### 結果 (前ページ要約)
- NDB研究：95件
 - 特別抽出79件、サンプリングデータセット9件、集計表情報7件
 - NDB Open Data研究：31件
 - the JMDC Claims Database研究：222件
 - the DPC database研究：737件
- 合計 **1085件**



結果 (研究分野)

	NDB	NDB Open Data	JMDC Claims Database	DPC Database
1位	内科	内科	内科	内科
2位	精神科	整形外科	精神科	外科
3位	整形外科	その他	眼科	救急科
4位	産婦人科	産婦人科 リハビリテーション科	小児科	整形外科
5位	リハビリテーション科 総合診療科	-	産婦人科	小児科

研究分野	NDB, n(%)	NDB Open Data, n(%)	JMDC, n(%)	DPC, n(%)	All, n(%)
合計	95 (51.6)	31 (38.7)	222 (64.9)	737 (29.3)	1085 (38.8)
内科	3 (3.2)	1 (3.2)	12 (5.4)	44 (6.0)	60 (5.5)
小児科	1 (1.1)	1 (3.2)	4 (1.8)	3 (0.4)	9 (0.8)
皮膚科	12 (12.6)	1 (3.2)	17 (7.7)	13 (1.8)	43 (4.0)
外科	3 (3.2)	1 (3.2)	0 (0.0)	139 (18.9)	143 (13.2)
整形外科	7 (7.4)	5 (16.1)	6 (2.7)	57 (7.7)	75 (6.9)
産婦人科	5 (5.3)	2 (6.5)	7 (3.2)	14 (1.9)	28 (2.6)
眼科	3 (3.2)	1 (3.2)	14 (6.3)	4 (0.5)	22 (2.0)
耳鼻咽喉科	0 (0.0)	0 (0.0)	3 (1.4)	15 (2.0)	18 (1.7)
泌尿器科	0 (0.0)	1 (3.2)	2 (0.9)	22 (3.0)	25 (2.3)
脳神経外科	1 (1.1)	0 (0.0)	2 (0.9)	43 (5.8)	46 (4.2)
放射線科	0 (0.0)	1 (3.2)	0 (0.0)	3 (0.4)	4 (0.4)
麻酔科	0 (0.0)	0 (0.0)	0 (0.0)	15 (2.0)	15 (1.4)
病理	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
臨床検査	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
救急科	0 (0.0)	0 (0.0)	2 (0.9)	100 (13.6)	102 (9.4)
形成外科	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.4)	3 (0.3)
リハビリテーション科	4 (4.2)	2 (6.5)	2 (0.9)	24 (3.3)	32 (2.9)
総合診療科	4 (4.2)	0 (0.0)	2 (0.9)	15 (2.0)	21 (1.9)
その他	3 (3.2)	3 (9.7)	5 (2.3)	7 (0.9)	18 (1.7)

考察

- (NDBを含む)リアルワールドデータを使用した研究は全体かつ個別で増加傾向にある。
 - 海外のデータベース研究数と比較しても遜色ない。
- JMDC Claims Database, DPC databaseを使用した研究が多い。
- NDB, NDB Open Data, JMDC Claims Databaseでは記述研究が多く、DPC databaseでは比較研究が多い。
- NDB, NDB Open Data, JMDC Claims Databaseは健診情報が含まれることから生活習慣病関連の研究も多い。

18

補足 (Limitation等)

- NDBを使用した研究に関しては、日本語論文も数多く出版されている。
- 研究分野の選択は、議論の余地がある可能性。
- 今回調査したデータベース以外にも日本のリアルワールドデータが存在する。

19

【小括】リアルワールドデータ研究レビュー

- リアルワールドデータを使用した研究は増加傾向 (平均増加率約41%)。
- それぞれのデータベースごとの特色があるため、研究目的に適したデータベース選択を行う必要がある。

20

目次

1. リアルワールドデータ研究のレビュー
2. 研究紹介

21

当研究室の先行NDB研究紹介

1. Ishimaru M, Matsui H, Ono S, Hagiwara Y, Morita K, Yasunaga H. Preoperative oral care and effect on postoperative complications after major cancer surgery. *Br J Surg*. 2018 Nov;105(12):1688-1696.
2. Takeuchi Y, Kumamaru H, Hagiwara Y, Matsui H, Yasunaga H, Miyata H, Matsuyama Y. Sodium-glucose cotransporter-2 inhibitors and the risk of urinary tract infection among diabetic patients in Japan: Target trial emulation using a nationwide administrative claims database. *Diabetes Obes Metab*. 2021 Jun;23(6):1379-1388.
3. Hashimoto Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, Kaburaki T. Incidence of Sympathetic Ophthalmia after Inciting Events: A National Database Study in Japan. *Ophthalmology*. 2022 Mar;129(3):344-352.
4. Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H. Prevalence, Incidence Rate, and Risk Factors of Medication-Related Osteonecrosis of the Jaw in Patients With Osteoporosis and Cancer: A Nationwide Population-Based Study in Japan. *J Oral Maxillofac Surg*. 2022 Apr;80(4):714-727.

22

当研究室の先行NDB研究紹介①

歯科医による術前口腔ケア介入はがん手術後合併症を減少させるか

Preoperative oral care and effect on postoperative complications after major cancer surgery. *Br J Surg*. 2018;105:1688-96

- 主ながん手術の術後合併症として誤嚥性肺炎は2.6 – 3.5%発生するとされており、入院日数の延長や死亡率上昇と関連している。
- 口腔ケアによる、口腔環境改善等を通じて術後肺炎を予防できる可能性が示唆されている。
- 先行研究は、サンプルサイズや参加施設数が小さいというLimitationがあり、術前口腔ケアと術後肺炎の関連は不明である。

23

対象者	頭頸部がん、食道がん、肺がん、肝がん、消化器系がん手術を行なった患者
曝露・対照	周術期口腔機能管理料の算定の有無
アウトカム	術後30日以内（入院中）の肺炎発症、全死亡率
デザイン	コホート研究、傾向スコアによる重みづけ

- 509,179名中、81,632名(16%)が術前口腔ケアを受けた
- 傾向スコアによる重み付け解析で、術前口腔ケア群は有意にアウトカム発生割合が少なかった
(30日肺炎発症 3.28% vs. 3.76%; 30日全死亡 0.30% vs. 0.42%)

24

当研究室の先行NDB研究紹介②

糖尿病患者におけるSGLT2阻害薬と尿路感染症の関連

Sodium-glucose cotransporter-2 inhibitors and the risk of urinary tract infection among diabetic patients in Japan: Target trial emulation using a nationwide administrative claims database. *Diabetes Obes Metab.* 2021;23:1379-88.

- 2型糖尿病治療の第一選択薬としてSGLT2阻害薬はDPP-4阻害薬等と同様に使用されているが、尿路感染症（UTI）との関連が示唆されている。
- リアルワールドでは、治療中に治療薬の変更が起こるが先行研究はその影響を考慮しておらず、妥当性が低い可能性がある。

25

対象者	40歳以上の2型糖尿病患者で新規に薬物療法を開始した者
曝露・対照	SGLT2阻害薬 or DPP-4阻害薬 vs. ビグアナイド薬
アウトカム	新規薬物治療開始後1080日以内の尿路感染症発症
デザイン	コホート研究、傾向スコアによる重み付けCox回帰

- 11,364名のSGLT2阻害薬群、9,035名のDPP-4阻害薬群、10,359名のビグアナイド薬群を解析。
- 傾向スコアによる重み付け解析によるハザード比は以下のとおりで、UTIのリスク増加を認めなかった。
SGLT2阻害薬 vs. ビグアナイド
ITT: 0.94 (95% CI 0.86-1.03), PP: 0.90 (95% CI 0.78-1.03)
DPP-4阻害薬 vs. ビグアナイド
ITT: 0.85 (95% CI 0.77-0.94), PP: 0.83 (95% CI 0.71-0.95)
(ITT, intention-to-treat; PP, per-protocol)

26

当研究室の先行NDB研究紹介③

外傷・手術後の交感性眼炎の発生率

Incidence of Sympathetic Ophthalmia after Inciting Events: A National Database Study in Japan. *Ophthalmology.* 2022;129:344-52

- 交感性眼炎(Sympathetic ophthalmia)は、一方の目への外傷や手術の曝露で他方の目に生じる汎ぶどう膜炎であり、発症率0.01%程度と稀な疾患である。
- 先行研究では、外傷後の交感性眼炎発症率の報告や、特に硝子体切除術(vitrectomy)でリスク増加の可能性等が示唆されているものの、単施設研究であったり、現在の医療技術水準にはそぐわない時代のものである等のLimitationが挙げられる。

27

対象者	眼への外傷、特定の眼手術を受けた患者
曝露・対照	眼への外傷、特定の眼手術
アウトカム	交感性眼炎、Vogt-Koyanagi-Harada病、(汎ぶどう膜炎)の診断
デザイン	コホート、記述研究

- 888,041の曝露発生(704,717名)が解析され、60ヶ月での累積交感性眼炎発生率は0.044%であった。
- 交感性眼炎は硝子体切除術群 0.016% vs. 外傷群 0.073%であり、外傷の方が4 - 5倍誘引することが示唆された。

28

当研究室の先行NDB研究紹介④

骨粗鬆症およびがん患者における薬剤関連顎骨壊死の有病率・発生率・リスク因子

Prevalence, Incidence Rate, and Risk Factors of Medication-Related Osteonecrosis of the Jaw in Patients With Osteoporosis and Cancer: A Nationwide Population-Based Study in Japan. *J Oral Maxillofac Surg.* 2022;80:714-27.

- 薬剤関連顎骨壊死 (Medication-Related Osteonecrosis of the Jaw, MRONJ) は、骨粗鬆症や、がん患者への治療薬であるビスホスホネート薬 (BP) や RANKL阻害薬に関連する発症率0.001 - 0.1%の稀な疾患である。
- 先行研究は、症例対照研究デザインである、発症イベントは自己報告に基づく等のLimitationもあり、MRONJの有病率、発症率、リスク因子は不明である。

29

対象者	骨粗鬆症、または（かつ）がんで骨吸収抑制薬を使用している患者
曝露・対照	新規の骨吸収抑制薬の開始（BP薬、RANKL阻害薬）
アウトカム	MRONJの発症
デザイン	コホート研究、Cox回帰

- 2,819,310名が新規骨吸収抑制薬を開始されており、2,664,104 (94.5%)が骨粗鬆症、155,206ががん患者だった。
 - 骨粗鬆症：1,603名 (0.06%)がMRONJを発症、100,000人年あたり22.9の発症率
 - がん：2,274名 (1.47%)がMRONJを発症、100,000人年あたり1,231.7の発症率
- MRONJの発症は、口腔環境の悪さ（骨粗鬆症患者の歯周病でHR 1.30 (95%CI 1.10 - 1.52)）等の因子と関連していることが判明した。

30

【総括】 NDBを用いた研究

- NDBを用いた研究は増加傾向
- 国民全体（約98%）のレセプトデータを含む悉皆性
 - 稀なアウトカムの研究
 - 1入院中のアウトカムに限らず追跡可能
- 歯科や健診のデータも含んでいる
 - 特に歯科関連の研究を行うのに有用な可能性

31

ご清聴ありがとうございました

32

DPCデータを用いた研究1 -概要と研究紹介-

東京大学大学院医学系研究科ヘルスサービスリサーチ講座 特任准教授
城 大祐

DPCデータベース

DPCデータは、一次的には各DPC病院が厚生労働省に提出する目的で作成される。様々な団体がDPCデータを二次的に収集し、臨床研究などに活用している。

(1) 厚生労働省が提供するDPCデータ
2022年から個票データの提供を開始

https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryouhoken/dpc/index.html

(2) 学会や学術団体等が独自に収集
日本循環器学会JROAD-DPC、厚生労働科学研究DPCデータ調査研究班、など

(3) 商用のDPCデータ
Medical Data Vision、など

COI

本発表に関連するCOIはございません

DPCデータベース

様式1の項目

病院属性等 施設コード、診療科コード

データ属性等 データ識別番号、性別、年齢

入退院情報 予定・救急入院、救急車による搬送、
退院時転帰、在院日数

診断情報 主傷病名、入院の契機となった傷病名、
医療資源を最も投入した傷病名、
二番目に医療資源を投入した傷病名
入院時併存症名(10)、入院後発症疾患名(10)

手術情報 手術術式、麻酔

5

□ DPCデータベースとは

□ DPC研究の紹介

□ まとめ

DPCデータベース

様式1の項目

診療情報

身長・体重、喫煙指数、入院時・退院時JCS、
入院時・退院時ADLスコア、
がんUICC病期分類・Stage分類、
入院時・退院時mRS、脳卒中の発症時期、
Hugh-Jones分類、NYHA心機能分類、狭心症CCS分類、
急性心筋梗塞Killip分類、肺炎の重症度、肝硬変Child-Pugh分類、
急性膵炎の重症度、
精神保健福祉法における入院形態・隔離日数・身体拘束日数、
入院時GAF尺度、SOFAスコア

6

DPCデータベース

EFファイルから得られるデータ

薬剤・特定保険医療材料の
名称・使用日・使用量

検査・処置の実施

医療費

など

7

DPCデータベース研究の限界

- ▶ 患者が病院を変えると追跡できない
⇒ 長期生存の追跡には向いていない
(5年生存率など)
- ▶ 検査結果データが無いので、リスク調整がいつも十分にできるわけではない
⇒ 未測定の変数がある

10

DPCデータベース

2020年3月30日時点
2019年4月2日版からの変更は赤字
2020年2月28日版からの変更は青字

行数	機能	項目
1	呼吸	PwO ₂ /FiO ₂ (mmHg)
2	凝固	血小板数 (x 10 ⁹ /mm ³)
3	肝	総ビリルビン値 (mg/dL)
4	循環	平均血圧/循環 作動薬投与
5	中枢神経	Glasgow Coma Scale 総点
6	腎	クレアチニン値 (mg/dL)

2020年度「DPC導入の影響評価に係る調査」実施説明資料

2020年2月30日

<https://www.01.prrism.com/dpc/2020/top.html>

- DPCデータベースとは
- DPC研究の紹介
- まとめ

DPCデータベース研究

前向き研究が行い難い研究テーマ
対象症例（疾患や病態）が少ない
対象が小児、高齢者や合併症のある症例
すでに広く普及している介入
重篤な病態への介入

前向き研究の制約があるテーマ
外的妥当性（選抜された症例のため、結果が外挿しにくい）
時間がかかる（サロゲートアウトカムのことが少なくない）
費用がかかる

9

DPC研究の疾患分野 (2014-2022 域共著)

呼吸器内科	循環器内科
COPD	血液内科
肺がん	感染症内科
喘息	集中治療
びまん性肺疾患	救急
希少疾患	麻酔・周術期
肺炎・感染症	小児科
診断および治療の手法	消化器外科
呼吸器外科	整形外科
リハビリ	眼科
薬剤	
看護	

DPC研究の分類 (2014-2022 城共著)

記述的研究 descriptive study

予測モデル prediction model
ノモグラム nomogram
機械学習 machine learning

分析的研究 analytic study

傾向スコア分析 propensity score analysis

生存分析 survival analysis

目的

薬剤カテゴリー毎に薬剤性肺障害発症のリスクを評価する。

研究事例

Risk of drug-induced interstitial lung disease in hospitalised patients: a nested case-control study

Jo T, et al. *Thorax*. 2021

方法

研究デザイン：コホート内**症例対照研究**
データベース：DPCデータベース
期間：2010年7月から2016年3月

対象：全入院患者
症例：入院中に薬剤性肺障害と診断されステロイド治療を要した症例
対照：それ以外の入院症例からマッチングにより選択

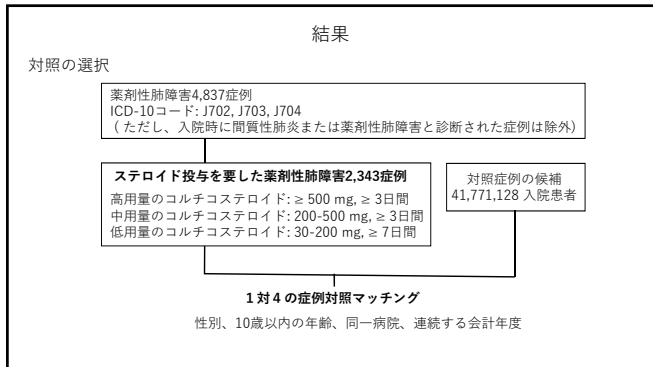
薬剤：
➢ 添付文書の有害事象に間質性肺炎の記載がある、約6000の薬剤を薬効により75のカテゴリーに分類
➢ そのうち、過去の大規模コホート研究やシステマティックレビューなど英文誌に記載のある42の薬剤カテゴリーを評価

背景

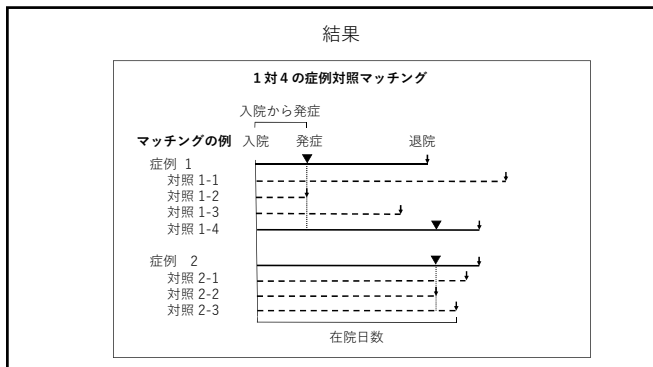
- 薬剤性肺障害は、発症頻度が低く研究が難しい
- 症例報告やケースシリーズでは、因果推論が困難
- 小規模研究の結果は、一般化が難しい
- リンパ球刺激試験などの臨床検査の診断における有効性は高くない
- 原因薬剤の推定が難しい場合がある

方法

- 統計解析：
- 対照の選択
症例と対照を、
性別
年齢（10歳以内）、
施設
前後1年の会計年度
で1対4マッチング
 - 条件付き単変量ロジスティック回帰分析
 - 多重代入法による欠損値補完後に
条件付き多変量ロジスティック回帰分析



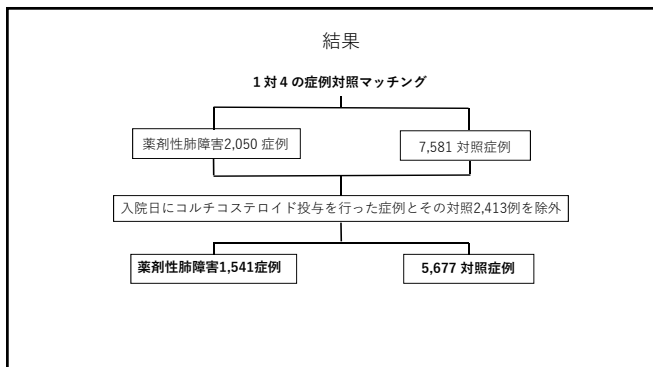
- 結果**
- 薬剤の選択
- 75薬剤カテゴリー
6つの大規模コホート研究やシステマティックレビューに含まれるものを選択
⇒33薬剤カテゴリーを除外
 - 42薬剤カテゴリー
カテゴリーあたり5症例未満を除外
⇒23薬剤カテゴリーを除外
 - 19薬剤カテゴリー
退院3日前に処方された薬剤を除外後にカテゴリーあたり5症例未満を除外
⇒2薬剤カテゴリーを除外
 - 17薬剤カテゴリー
⇒単変量の条件付きロジスティック回帰分析



結果

単変量ロジスティック回帰分析による症例と対照のステロイド投与量および在院死亡の比較

	症例		対照		オッズ比	95%信頼区間	P値
	n	(%)	n	(%)			
高用量ステロイド	1246	(80.9)	156	(2.8)	185.92	123.0-280.9	<0.001
中用量ステロイド	119	(7.7)	28	(0.5)	19.30	12.10-30.80	<0.001
低用量ステロイド	671	(43.5)	263	(4.6)	18.47	15.17-22.48	<0.001
在院死亡	235	(34.7)	610	(10.8)	5.24	4.49-6.10	<0.001



結果

17薬剤カテゴリーと薬剤性肺障害発症の条件付き単変量ロジスティック回帰分析

	オッズ比	95%信頼区間	P値
Antiplatelets	1.01	0.69-1.46	0.977
Anticoagulants	2.11	1.16-3.85	0.024
Statins	0.68	0.50-0.92	0.012
Sodium channel blockers	1.28	0.61-2.67	0.511
Class III antiarrhythmic drugs	6.89	3.96-11.66	<0.001
Angiotensin/converting enzyme inhibitor	1.83	0.80-4.20	0.156
Thiazides	0.59	0.31-1.14	0.116
NSAIDs	2.43	2.04-2.88	<0.001
Antiepileptics	1.07	0.59-1.94	0.836
Sulfamethoxazole/trimethoprim	2.89	1.36-6.13	0.006
Quinolones	2.93	2.35-3.65	<0.001
Tetracyclines	2.30	1.50-3.52	<0.001
Beta-lactams	1.62	1.39-1.89	<0.001
Anti-tuberculosis drugs	2.14	0.94-4.89	0.071
EGFR inhibitors	16.48	9.57-28.39	<0.001
Pyrimidine	1.18	0.63-2.21	0.599
Anthracyclines	3.09	1.24-7.70	0.015

結果

多重代入法後に単変量で有意差のあった10薬剤カテゴリーを含めた条件付き多変量ロジスティック回帰分析

	オッズ比	95%信頼区間	P値
Anticoagulants	2.58	0.76-8.81	0.13
Statins	0.53	0.37-0.75	<0.001
Class III antiarrhythmic drugs	7.01	3.86-12.73	<0.001
NSAIDs	1.9	1.56-2.31	<0.001
Sulfamethoxazole/trimethoprim	2.54	1.04-6.24	0.042
Quinolones	3.1	2.41-3.99	<0.001
Tetracyclines	1.6	0.97-2.66	0.067
Beta-lactams	1.54	1.29-1.84	<0.001
EGFR inhibitors	16.84	9.32-30.41	<0.001
Anthracyclines	1.89	0.68-5.23	0.223

調整項目：パーセルインデックス、ブリンクマンインデックス、チャールソン併存疾患インデックス、肺癌、その他の癌、入院2日以内のICU入室、入院2日以内の挿管人工呼吸器管理

結語

コホート内症例対照研究により、薬剤カテゴリーを薬剤性肺障害発症の関連を評価し、薬剤カテゴリー毎の薬剤性肺障害の発症リスクを定量化して示した。

考察

- 本研究では、初めて大規模データベースを用いて薬剤性肺障害の発症リスクを評価した。
- 薬剤性肺障害のリスクが高いと考えられる42カテゴリーの薬剤を評価し、そのうち6つの薬剤が薬剤性肺障害の高リスクであることが示された。
- 薬剤以外では既報と同様に、男性、60歳以上、多い併存疾患、痩せ、肺癌が薬剤性肺障害の発症リスクであることが示された。
- 薬剤性肺障害発症のリスクの高い薬剤のなかでも、EGFR阻害剤とIII群の抗不整脈薬が、とりわけ高いオッズ比を示し、これは小規模な既報と合致した。

- DPCデータベースとは
- DPC研究の紹介
- まとめ

限界

- 薬剤性肺障害の正確な発症日は不明。
- シスプラチンのようにステロイドと一緒に投与される薬剤の評価が出来ていない。
- 薬剤間の交互作用の探索的評価をしていない。
- 画像や採血などの検査結果は評価が出来ていない。
- 処方が少ない薬剤は評価が出来ていない。
- 未測定の間接が存在する可能性を排除できない。

まとめ

- DPCデータベースについて概説した
- DPCデータベースを用いた薬剤性肺障害の解析事例を紹介した
- 入院診療にかかわる様々な職種・診療科の方々が、種々の解析法を駆使した研究を行っている
- 柔軟な発想により、前向き研究を補完するような研究が可能である
- データの蓄積により今後も多くの研究が行われていくと考える

DPCデータを用いた研究2

東京大学大学院医学系研究科臨床疫学・経済学
東京大学大学院医学系研究科乳腺・内分泌外科学
小西孝明

DPCデータに適した研究

▶DPCデータは以下の研究に適している

- 症例数が膨大な点を活かす
 - ✓ 希少な病態・介入の記述疫学研究
 - ✓ (倫理的・経済的・実務的に) RCTが困難な比較研究
- 入院データである点を活かす
 - ✓ 入院を要する病態・介入の研究
 - ✓ 入院中に重要なアウトカムが発生する研究

▶救急領域や外科領域などの急性疾患を扱いやすい

2

臨床疫学研究の内容による分類

1. 治療・予防などの効果判定
2. 診断法や患者評価法
3. 疾病のリスク因子の同定
4. 疾病の予後予測
5. 希少疾患の予後予測
6. 診療実態の分析 (practice pattern analysis)
7. 医療の質研究、ヘルス・サービス・リサーチ
などなど

ランダム化比較試験がゴールドスタンダードだが、
後ろ向きデータベース研究でも実施可能

3

外科領域のDPCデータ研究例

1. 統合失調症患者の乳癌手術成績
2. Body mass indexが食道癌手術に及ぼす影響
3. 尾側臍切除における開腹 vs. 鏡視下手術

4

外科領域のDPCデータ研究例

1. 統合失調症患者の乳癌手術成績
2. Body mass indexが食道癌手術に及ぼす影響
3. 尾側臍切除における開腹 vs. 鏡視下手術

5

統合失調症患者の乳癌手術成績

- "Breast cancer surgery in patients with schizophrenia: short-term outcomes from a nationwide cohort"
- Konishi T, Fujiogi M, Michihata N, Tanaka-Mizutani H, Morita K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H
- *British Journal of Surgery* 2021;108:168-173

6

Introduction

- ▶統合失調症は人口の1%が罹患する慢性精神疾患である
 - ・ 乳癌の罹患リスクが高い
 - ・ 治療後の長期予後も不良である
 - ・ **診断の遅れ**や**不十分な治療**の影響があると考えられている
- ▶乳癌術後短期成績への影響はいまだ不明である
 - ・ 股関節手術50万例では、**統合失調症は高リスク**との報告あり
 - ・ 術後ICUに入室した約9千例では、**統合失調症は有意なリスクでない**との報告あり
- ▶外科医と精神科医の**リエゾン**が世界的に重要視されつつある

7

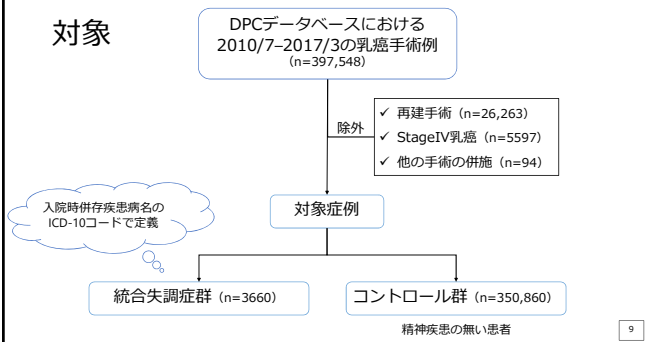
PECO

- P: 乳癌手術を受けた患者
 E: 統合失調症あり
 C: 精神疾患なし
 O: 全合併症 (術後出血・創部感染など)、入院費用
- ▶主解析 : 多変量解析 (マルチレベル分析を併用)
 ▶感度分析: 1:4 matched-pair cohort解析 など

入院後発症病名の
ICD-10コードで定義

8

対象



9

多変量解析

- ▶背景因子として下記で調整
- ・ 患者要因: 性別、年齢、Body mass index (BMI)、喫煙歴、Charlson併存疾患指数
 - ・ 治療要因: 癌病期、術式
 - ・ 施設要因: 臨床研修指定の有無、年間手術件数
- ▶マルチレベル分析を併用
- ・ 一般化推定方程式
 - ・ 患者・治療背景の病院ごとの違いを考慮する

DPCデータ様式1に
含まれる情報

マルチレベル分析について
は別講義も参照ください

10

感度分析

- 1:4 matched-pair cohort解析
 - ✓ 両群から年齢・施設・治療年が同一の患者を1:4の割合で抽出する
 - ✓ この解析でも施設間差異の影響を除去できる
- 統合失調症群を抗精神病薬の処方がある患者のみに限定した解析
 - ✓ 統合失調症群に**統合失調感情障害**の患者が含まれているかもしれないため
 - ✓ 真の統合失調症であれば継続的な抗精神病薬の内服が必要
- 統合失調症群から強制入院の患者を除外した解析
 - ✓ 自傷他害のおそれがある重症患者と考えられるため

病名と処方の
組み合わせで
精度向上

11

患者要因	統合失調症群			コントロール群			
	n=3,660	n=350,860	ASD*	n=3,660	n=350,860	ASD*	
女性	3636 (99)	348,818 (99)	0.9				
年齢カテゴリー、歳				癌病期			
18-55	403 (11)	46,835 (13)	7.2	0	230 (6.3)	38,168 (11)	16.5
55-65	753 (21)	78,997 (23)	4.7	I	929 (25)	143,511 (41)	33.4
65-75	860 (24)	80,726 (23)	1.2	II	1550 (42)	119,995 (34)	16.8
75-85	917 (25)	85,228 (24)	1.8	III	570 (16)	26,016 (7.4)	25.8
≥85	727 (20)	59,074 (17)	7.8	欠損値	611 (17)	61,338 (18)	2.1
Body mass index, kg/m ²				乳房全切除	2591 (71)	162,955 (46)	51.0
<18.5	399 (11)	31,684 (9.0)	6.3	腋窩リン結	1711 (47)	118,025 (34)	27.0
18.5-25	1792 (49)	222,319 (63)	29.3	臨床研修指定病院	3290 (90)	311,846 (89)	3.3
25-30	956 (26)	73,101 (21)	12.5	Hospital volume			
≥30	455 (12)	20,616 (5.9)	22.9	Low (<70)	1832 (50)	116,216 (33)	34.9
欠損値	58 (1.6)	3140 (0.9)	6.2	Medium (70-152)	1186 (32)	117,043 (33)	2.0
喫煙歴あり	852 (23)	70,791 (20)	7.5	High (≥152)	642 (18)	117,601 (34)	37.3
Charlson併存疾患指数							
0	2268 (62)	268,132 (76)	31.7				
1	683 (19)	38,840 (11)	21.5				
≥2	709 (19)	43,888 (13)	18.8				

*ASD(Absolute standardized difference)>10%は
差が大きいとみなす

12

主解析の結果

	統合失調症群 n=3660		コントロール群 n=350,860		多変量解析	
	n	(%)	n	(%)	OR	95%CI
在院中合併症	288	(7.9)	16,356	(4.7)	1.37	(1.21-1.55)
術後出血	85	(2.3)	4882	(1.4)	1.34	(1.05-1.71)
創部感染	107	(2.9)	7511	(2.1)	1.22	(1.04-1.43)
	中央値	四分位範囲	中央値	四分位範囲	係数	95%CI
総入院費用, Euro	7204	5941-8970	6133	5444-7259	743	(680-806)

OR, オッズ比, CI, 信頼区間

統合失調症患者で有意に合併症が多く、入院費用も高い

13

感度分析の結果

	1:4 matched-pair cohort解析		統合失調症群を精神科病 薬処方ありに限定		統合失調症群から 強制入院を除外	
	OR	95%CI	OR	95%CI	OR	95%CI
在院中合併症	1.25	(1.09-1.43)	1.54	(1.33-1.79)	1.33	(1.17-1.51)
術後出血	1.16	(0.90-1.51)	1.40	(1.04-1.88)	1.31	(1.02-1.69)
創部感染	1.08	(0.88-1.32)	1.35	(1.12-1.62)	1.23	(1.05-1.44)
	係数	95%CI	係数	95%CI	係数	95%CI
総入院費用, Euro	1233	(1047-1418)	1180	(1101-1259)	549	(487-611)

OR, オッズ比, CI, 信頼区間

感度分析でも同様の傾向

14

考察

統合失調症患者では合併症が多い

- ✓ 多数の薬剤内服による薬剤相互作用、意思疎通の困難さ
- ✓ 通常通りの医療が提供されていない可能性がある
- ✓ 外科医と精神科医の協働が必要だろう

Limitationとして以下の情報が不明

- ✓ 術前化学療法の有無
- ✓ 乳癌の罹患歴
- ✓ 統合失調症の重症度

こうしたデータベースに含まれない
情報が重大な交絡でないかは注意が必要

15

外科領域のDPCデータ研究例

1. 統合失調症患者の乳癌手術成績
2. Body mass indexが食道癌手術に及ぼす影響
3. 尾側膵切除における開腹 vs. 鏡視下手術

16

BMIが食道癌手術に及ぼす影響

- "Impact of body mass index on major complications, multiple complications, in-hospital mortality, and failure to rescue following esophagectomy for esophageal cancer: A nationwide inpatient database study in Japan"
- Hirano Y, Kaneko H, Konishi T, Itoh H, Matsuda S, Kawakubo H, Uda K, Hiroki M, Fushimi K, Itano O, Hideo Y, Kitagawa Y
- *Annals of Surgery* 2021 online ahead of print

17

Introduction

食道癌手術は高侵襲である

- 合併症や術後死亡が比較的多い
- 合併症後の死亡 (failure to rescue) も注目されている

BMIと術後短期成績との関連は未だ明らかでない

- 肥満は縫合不全や呼吸器合併症のリスクとの既報あり
- やせは術後死亡のリスクとの既報あり
- Failure to rescueなどとの関連は報告されていない

18

PECO

P: 食道癌手術を受けた患者
 E: やせ・肥満
 C: 標準BMI
 O: 院内死亡、合併症、Failure to rescue、術後在院日数、入院費用

▶主解析: Restricted cubic spline(RCS)を併用した多変量解析

病名と処置の組み合わせで定義

19

対象

DPCデータベースにおける
2010/7-2019/3の食道癌手術例
(n=40,245)

除外

- ✓ 18歳未満 (n=1)
- ✓ 咽喉頭癌手術の併施 (n=466)
- ✓ BMIの欠損値/外れ値 (n=1318)

対象症例
(n=39,406)

様式1の手術情報で定義

20

RCSを併用した多変量解析

▶背景因子として下記で調整

- 患者要因: 性別、年齢、喫煙歴、Charlson併存疾患指数
- 治療要因: 癌病期、術前治療、術式
- 施設要因: 臨床研修指定の有無、年間手術件数

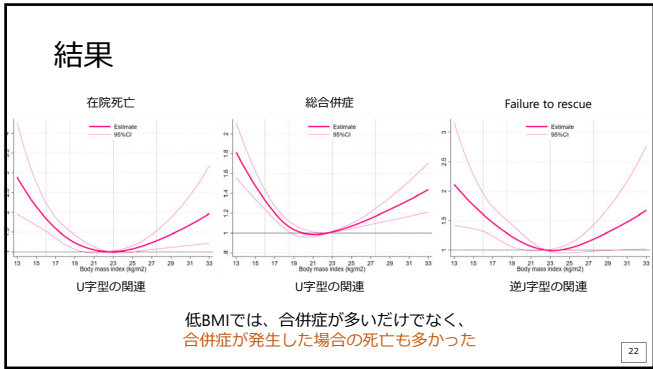
▶マルチレベル分析を併用

- 患者・治療背景の病院ごとの違いを考慮する

▶RCS解析を併用

- 連続変数 (BMI) をカテゴリー化せずに非線形の関連を图示できる

21



考察

▶BMIは食道癌の重要な術前評価項目であるだろう

- ✓ 低BMIは低い呼吸機能・サルコペニアなどの指標とされる
- ✓ 注意深い術後管理によるFailure to rescueの改善が望まれる
- ✓ 術前リハビリや栄養管理が有用かもしれない

▶Limitationとして以下の情報が不明

- ✓ 食道癌の組織型
- ✓ 鏡視下手術の保険償還開始 (2014年) 前の、鏡視下での実施の有無
- ✓ 術前体重減少の有無

癌病期はvalidateされているが、組織型まではわからない

保険償還の開始時期の知識が必要

23

外科領域のDPCデータ研究例

1. 統合失調症患者の乳癌手術成績
2. Body mass indexが食道癌手術に及ぼす影響
3. 尾側碎切除における開腹 vs. 鏡視下手術

24

尾側膵切除における開腹 vs. 鏡視下手術

- "Laparoscopic versus open distal pancreatectomy with or without splenectomy: A propensity score analysis in Japan"
- Konishi T, Takamoto T, Fujiogi M, Hashimoto Y, Hiroki M, Fushimi K, Tanabe M, Seto Y, Hideo Y
- *International Journal of Surgery* 2022;104:106765

25

Introduction

- ▶膵体尾部腫瘍に対する**腹腔鏡下切除**が普及しつつある
 - 本邦では2012年4月に保険適応になった
- ▶術後合併症や入院費用への影響はいまだ不明である
 - 術後合併症に関しては、既報内で結果が不一致
 - 高額の腹腔鏡手術の、入院費用への影響も不明
 - 施設因子などの背景因子も調整されていない

26

PICO

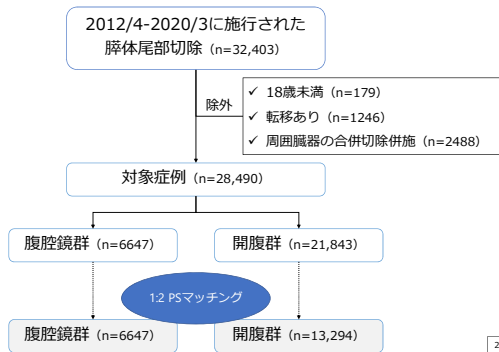
- P : 尾側膵切除を受けた患者
 I : 鏡視下手術
 C : 開腹手術
 O : 合併症、死亡、再手術、30日再入院、
 麻酔時間、ドレーン留置期間、術後在院日数、入院費用

費用はDPCデータベース
ならではアウトカム

- ▶主解析：1:2傾向スコア(PS)マッチング
- ▶感度分析：PSを用いたオーバーラップ重み付けなど

27

対象



28

1:2 PSマッチング

- ▶背景因子として下記でPSを作成
 - 患者要因：性別、年齢、BMI、喫煙歴、Charlson併存疾患指数
 - 治療要因：術直前の血糖降下薬の使用、良性/悪性、癌TN分類、脾臓温存
 - 施設要因：臨床研修指定の有無、年間手術件数
- ▶腹腔鏡群と開腹群が1:2になるようマッチング
 - PSを用いた最近傍マッチング

傾向スコア解析については
別講義も参照ください

29

感度分析

- ① PSを用いたオーバーラップ重み付け
 - 操作変数法については別講義も参照ください
 - ② 操作変数法
 - ✓ 未測定の変数 (詳細な術式など) を調整できる手法
 - ✓ 操作変数として、施設における腹腔鏡手術の施行割合 (選好率) を用いた
- ◆サブグループ解析も実施
- ✓ Low-volume施設を除いた解析 (熟練した施設に限定するため)
 - ✓ 良性腫瘍に限った解析 (リンパ(郭清)の影響を除くため)

30

Table 1. 背景

患者因子	全患者 (マッチング前)			1:2PSマッチ後の患者				
	腹腔鏡	開腹	標準偏差 (%)	腹腔鏡	開腹	標準偏差 (%)		
性別	男性	2669 (40)	12,014 (55)	30.1	2669 (40)	5200 (39)	2.1	
年齢	<50	1613 (24)	1807 (8.3)	44.4	1613 (24)	3487 (26)	4.5	
	50-59	1008 (15)	2429 (11)	12.0	1008 (15)	1826 (14)	4.1	
	60-69	1608 (24)	6164 (28)	9.2	1608 (24)	3260 (25)	0.8	
	70-79	1863 (28)	8600 (39)	24.2	1863 (28)	3692 (28)	0.6	
	>80	555 (8.3)	2843 (13)	15.1	555 (8.3)	1029 (7.7)	2.2	
Body mass index, kg/m ²	<18.5	647 (9.7)	2623 (12)	7.3	647 (9.7)	1261 (9.5)	0.8	
	18.5-21.9	2339 (35)	7955 (36)	2.6	2339 (35)	4807 (36)	2.0	
	22.0-24.9	2006 (30)	6707 (31)	1.1	2006 (30)	4078 (31)	1.1	
	25.0-29.9	1351 (20)	3851 (18)	6.9	1351 (20)	2645 (20)	1.1	
	≥30.0	282 (4.2)	571 (2.6)	9.0	282 (4.2)	469 (3.5)	3.7	
欠損値	22 (0.3)	136 (0.6)	4.2	22 (0.3)	34 (0.3)	1.4		
喫煙歴	あり	2638 (40)	3919 (18)	11.8	2638 (40)	5301 (40)	0.4	
	なし	5284 (79)	10,651 (49)	67.6	5284 (79)	10,754 (81)	3.5	
	Charlson Comorbidity index	3	820 (12)	6617 (30)	44.9	820 (12)	1559 (12)	1.9
	4	347 (5.2)	3182 (15)	31.7	347 (5.2)	691 (5.2)	0.1	
	≥5	196 (2.9)	1393 (6.4)	16.3	196 (2.9)	691 (5.2)	0.1	

Data are presented as n and (%).

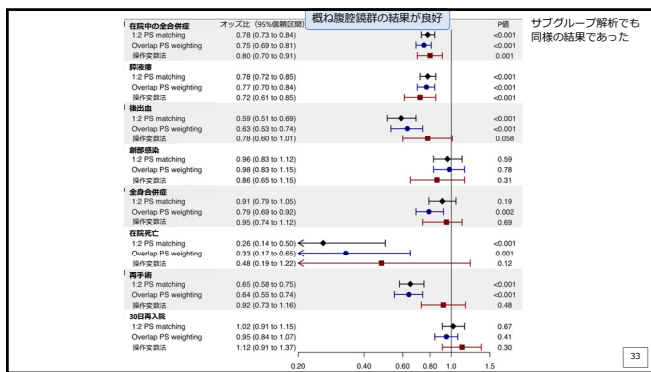
31

Table 1. 背景

治療・施設因子	全患者 (マッチング前)			1:2PSマッチ後の患者			
	腹腔鏡	開腹	標準偏差 (%)	腹腔鏡	開腹	標準偏差 (%)	
血糖降下薬の使用	インスリン	638 (10)	3894 (18)	24.1	638 (9.6)	1070 (8.0)	5.5
術前診断	悪性	1969 (30)	16,069 (74)	97.9	1969 (30)	3771 (28)	2.8
	T2以上	882 (13)	10,742 (49)	84.1	882 (13)	1783 (13)	0.4
術式	N1以上	168 (2.5)	3232 (15)	44.7	168 (2.5)	325 (2.4)	0.5
	脾温存	924 (14)	1052 (4.8)	31.6	924 (14)	1417 (11)	9.9
臨床研修認定あり	あり	4692 (71)	14,413 (66)	9.9	4692 (71)	9491 (71)	1.8
	なし	1001 (15)	7956 (36)	50.4	1001 (15)	2037 (15)	0.7
手術年度	2012-2015	2322 (35)	10,630 (49)	28.1	2322 (35)	5028 (38)	6.0
	2015-2019	4925 (65)	14,219 (64)	28.1	4925 (65)	8266 (62)	6.0

Data are presented as n and (%).

32



連続変数のアウトカム

アウトカム	PS解析			操作変数法		
	差	95%信頼区間	P値	差	95%信頼区間	P値
麻酔時間, 分	59	(56-63)	<0.001	59	(55-63)	<0.001
ドレナージ期間, 日	-4.0	(-4.5 to -3.6)	<0.001	-4.6	(-5.1 to -4.0)	<0.001
術後在院期間, 日	-4.4	(-4.9 to -3.9)	<0.001	-5	(-5.6 to -4.5)	<0.001
入院費用, 円	-138,394	(-166,750 to -110,037)	<0.001	-195,294	(-227,002 to -163,586)	<0.001

腹腔鏡群で
 ・麻酔時間は約1時間長い
 ・ドレナージ期間・在院期間は短い
 ・入院費用も安い

サブグループ解析でも同様の結果であった

34

考察

▶ 腹腔鏡では尿液瘻や後出血が少ない

- ✓ 死亡や入院期間も減少
- ✓ 高額な術式のはずだが、入院費用も減少していた
- ✓ 麻酔時間は1時間長い、それは許容可能だろう

▶ Limitationとして以下の情報が不明

- ✓ 胃排遅延 (代わりに入院期間で評価可能か)
- ✓ 術者情報 (代わりに施設因子で調整した)
- ✓ 腹腔動脈合併切除 (操作変数法で調整し、サブグループ解析も行った)

入手不能な因子はあるが、代替変数で評価したり、疫学的手法で対処したりする

35

小括

▶ 治療などの効果判定に関する外科領域のDPCデータ研究例を紹介した

▶ 多変量解析に種々の臨床疫学的手法を併用した

- ・マルチレベル分析、RCS分析、傾向スコア、操作変数法など

▶ 信頼性を高める様々な工夫を行った

- ・病名と処置・薬剤処方との組み合わせ
- ・代替変数の活用
- ・感度分析の実施

36

おわりに

- DPCデータを用いた救急・外科領域の研究例を提示した
- 病名・処置・処方データを活用した様々な研究が行えた
- DPCに含まれるデータの内容に注意して解析を行った

DPCデータを用いた研究2

東京大学大学院医学系研究科臨床疫学・経済学
大邊寛幸

DPCデータの利点

- 大規模、N数は桁外れ
 - 症例数やアウトカムが稀な研究が実施可能
- 母集団代表性に優れる
 - 日本の急性期病院の50%、ICU・HCUの70-80%、3次救急病院の90%をカバー
- 実臨床を反映している
 - ランダム化が方法論や倫理的に不可能な臨床課題を解決

DPCデータの欠点

- 患者が病院を変えともはや追跡できない
 - 長期生存の追跡には向いていない
- バイタルや、検査結果データが無い
 - Risk adjustmentがいつも十分にできるわけではない
- 診断病名の妥当性が中程度
 - 主要な病名（心不全など）の感度78.9%、特異度93.2%

Yamana H, et al. J Epidemiol. 2017;27:476-483

DPCデータに適した研究

- DPCデータは以下の研究に適している
 - 症例数が膨大な点を活かす
 - ✓ 希少な病態・介入の記述疫学研究
 - ✓ (倫理的・経済的・実務的に) RCTが困難な比較研究
 - 入院データである点を活かす
 - ✓ 入院を要する病態・介入の研究
 - ✓ 入院中に重要なアウトカムが発生する研究
- 救急領域や外科領域などの急性疾患を扱いやすい

臨床疫学研究の内容による分類

1. 治療・予防などの効果判定
⇒RCTはこれに含まれるが、RCTだけでない
2. 診断法や患者評価法
3. 疾病のリスク因子の同定
4. 疾病の予後予測
5. 稀少疾患の記述疫学
6. 診療実態の分析(practice pattern analysis)
7. 医療の質研究、ヘルス・サービス・リサーチ などなど

救急・集中治療領域のDPCデータ研究例

1. 稀少疾患の記述疫学
コリン作動性クリーゼの記述疫学
2. 診療実態の分析(practice pattern analysis)
日本のICU病床稼働率の実態
3. 医療の質研究、ヘルス・サービス・リサーチ
「重症度、医療・看護必要度」ICU入室基準の導入が与えた影響

救急・集中治療領域のDPCデータ研究例

1. 稀少疾患の記述疫学

コリン作動性クリーゼの記述疫学

2. 診療実態の分析(practice pattern analysis)

日本のICU病床稼働率の実態

3. 医療の質研究、ヘルス・サービス・リサーチ

「重症度、医療・看護必要度」ICU入室基準の導入が与えた影響

論文掲載報告

ORIGINAL ARTICLE



Cholinergic Crisis Caused by Cholinesterase Inhibitors: a Retrospective Nationwide Database Study

Hiroyuki Ohbe¹ · Taisuke Jo¹ · Hiroki Matsui¹ · Kiyohide Fushimi² · Hideo Yasunaga¹

Received: 7 January 2018 / Revised: 3 May 2018 / Accepted: 31 May 2018 / Published online: 15 June 2018
© American College of Medical Toxicology 2018

1. *Journal of Medical Toxicology* (2018) 14:237–241
<https://doi.org/10.1007/s13181-018-0669-1>

研究例1：稀少疾患の記述疫学

Clinical Question

“コリン作動性クリーゼ”という稀少疾患についてコリンエステラーゼ阻害薬による重篤な副作用で呼吸不全を来たとされている。

しかし既存の報告は10数例程度のケースシリーズが報告されているのみであり予後などは不明。

患者抽出条件

- 病名のいずれかに以下のICD10コードを含む
T44.0 (主として自律神経系に作用する薬物による中毒, コリンエステラーゼ阻害薬)
- その患者のうち“コリン作動性クリーゼ”と日本語で傷病名に記載がある

病名くん(ICD10コード)

<http://www.byomei.org/byomei-kun.2.0/>

病名くん(ICD10コード)

ICD10 国際疾病分類第10版 (2013年版)

大分類 (章) 一覧

章	ICDコード	分類見出し
17	Q00-Q99	先天奇形、変形および染色体異常
18	R00-R99	症状、徴候および異常臨床所見・異常検査所見で他に分類されないもの
19	S00-T98	損傷、中毒およびその他の外因の影響
20	V01-Y98	傷病および死亡の外因
21	Z00-Z99	健康状態に影響をおよぼす要因および保健サービスの利用
22	U00-U99	特殊目的用コード

<http://www.byomei.org/byomei-kun.2.0/>

病名くん(ICD10コード)

T44 主として自律神経系に作用する薬物による中毒

- T44.0 主として自律神経系に作用する薬物による中毒、コリンエステラーゼ阻害薬
- T44.1 主として自律神経系に作用する薬物による中毒、その他の副交感神経興奮薬[コリン作動薬]
- T44.2 主として自律神経系に作用する薬物による中毒、神経筋遮断薬、他に分類されないもの

<http://www.byomei.org/byomei-kun>

結果

- DPCデータにおいて2010年7月から2016年3月の期間中、235人のコリン作動性クリーゼの患者が同定された。
- 20%の患者が人工呼吸器を必要とし、院内死亡率は6.4%であった。
- 入院患者の半数以上がカテコラミン投与を必要とした。

Ohbe H, et al. J Med Toxicol. 2018;14:237-241

限界

- 患者抽出条件の感度・特異度が不明。
- 原因となるコリンエステラーゼ阻害薬をいつどのくらい内服したのか不明。
- コリン作動性クリーゼの診断基準に含まれる症状、バイタル、身体所見、血液検査値のデータが無い。

救急・集中治療領域のDPCデータ研究例

1. 稀少疾患の記述疫学
コリン作動性クリーゼの記述疫学
2. 診療実態の分析(practice pattern analysis)
日本のICU病床稼働率の実態
3. 医療の質研究、ヘルス・サービス・リサーチ
「重症度、医療・看護必要度」ICU入室基準の導入が与えた影響

論文掲載報告



Journal of Epidemiology



Original Article

J Epidemiol 2022

Intensive Care Unit Occupancy in Japan, 2015–2018: A Nationwide Inpatient Database Study

Hiroyuki Ohbe¹, Yusuke Sasabuchi², Ryosuke Kumazawa¹, Hiroki Matsui¹, and Hideo Yasunaga¹

¹Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo, Tokyo, Japan
²Data Science Center, Jichi Medical University, Tochigi, Japan

1. J Epidemiol 2021 Apr 10.
doi: 10.2188/jea.JE20210016.

研究例2：診療実態の分析

- 人口10万人当たりICUベッド数は既存によく使われている供給能力を示す指標の一つであり、日本は5.6床と他先進国に比べ少ないため供給が足りないと議論されている²。
- ICUの病床稼働率もその地域におけるCritical Care Systemの供給能力を示す指標の一つ¹。
- 過去に日本のICU病床稼働率は検証されたことは無い。

1. Intensive Care Med. 2019;45:1231-1240.
2. J Japanese Soc Intensive Care Med. 2010;17:227-232.

方法

- 記述疫学
- 厚生労働科学研究調査研究班のDPCデータ + 病床機能報告2017
- 観察期間：2015年1月1日～2018年12月31日
- 対象患者：ICU入室した全ての患者
 - ICU定義: 1:2看護患者比率の病床

主要評価項目

- ICU稼働率を以下のように算出

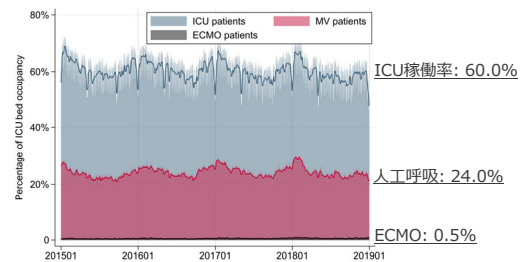
$$\text{ある日のA病院のICU稼働率(\%)} = \frac{\text{ある日のA病院のICU管理料取得人数の合計(DPC)}}{\text{A病院の2017年許可ICU病床数(病床機能報告)}} \times 100$$

- 人工呼吸器,ECMOの患者の稼働率も算出

結果

- 137万9618人のICU患者(495病院, 5341のICU病床) が同定された。
- これは2017年度の日本の全ICU病床の75%(N=5341/7109) を含んでいた。

結果 -ICU稼働率-



考察 -先行研究との比較-

	日本	米国
ICU稼働率	60%	70%
人工呼吸	24%	30%
10万人対ICU病床数	5.6床	20床

- 10万人対ICU病床数が4倍多い米国と比較しても、ICU稼働率並びに人工呼吸稼働率は日本は低い傾向にあった¹。
- これまでの欧米諸国の声明ではICU稼働率は70-75%が最適であろうと推奨されている。

1. Crit Care Med. 2013;41:2712-2719.
 2. Crit Care Med. 2009;37:2753-2758.
 3. Aust Crit Care. 2014;27:77-84.

考察 -限界-

- ICU算定を受けられないICU長期滞在患者（14日以降）を無視したため、稼働率が低く見積もられている

救急・集中治療領域のDPCデータ研究例

1. 稀少疾患の記述疫学

コリン作動性クレーゼの記述疫学

2. 診療実態の分析(practice pattern analysis)

日本のICU病床稼働率の実態

3. 医療の質研究、ヘルス・サービス・リサーチ

「重症度、医療・看護必要度」ICU入室基準の導入が与えた影響

論文掲載報告

Associations of Government-issued Intensive Care Unit Admission Criteria with Clinical Practices, Outcomes, and Intensive Care Unit Bed Occupancy

Hiroyuki Ohbe¹, Tadahiro Goto^{1,2}, Hiroki Matsui¹, Kiyohide Fushimi³, and Hideo Yasunaga¹

¹Department of Clinical Epidemiology and Health Economics, School of Public Health, University of Tokyo, Tokyo, Japan; ²TXP Medical Co. Ltd., Tokyo, Japan; and ³Department of Health Policy and Informatics, Graduate School of Medicine, Tokyo Medical and Dental University, Tokyo, Japan

ORCID IDs: 0000-0001-8544-2569 (H.O.); 0000-0002-5880-2968 (T.G.); 0000-0003-0004-4743 (H.M.); 0000-0002-1894-0290 (K.F.); 0000-0002-6017-469X (H.Y.).

1. *Ann Am Thorac Soc.* 2021 Nov 23.
doi: 10.1513/AnnalsATS.202107-844OC.

3. 医療の質研究、ヘルス・サービス・リサーチ

・日本政府は2014年4月から「**重症度・医療・看護必要度**」という新たなICU入室基準を医療政策として導入²

・各ICUは表の基準を満たす患者を**70-90%**以上受け入れなくてはならないと定められた。

モニタリング及び処置	2014-15年	2016-17年
1 心電図モニター	1点	1点
2 動脈圧測定	1点	2点
3 中心静脈圧測定	1点	2点
4 肺動脈圧測定	1点	2点
5 輸液ポンプ	1点	1点
6 シリンジポンプ	1点	1点
7 人工呼吸器	1点	2点
8 輸血・血液製剤	1点	2点
9 特殊な治療法	1点	2点
基準を満たす合計点数	≥3点以上	≥4点以上

3. 医療の質研究、ヘルス・サービス・リサーチ

・「**重症度・医療・看護必要度**」は**不必要なICU入室を減らす**目的で策定された¹。

・本基準は**医療者の診療行為に基づく指標**であり、他国の患者重症度に基づく指標(SOFAなどの臓器障害スコア)と大きく異なっていた^{2,3}。

1. 日本医療・病院管理学会誌 2008;45:37-48
2. *Ann Intensive Care* 2013;3:37.
3. *Curr Opin Crit Care* 2009;15:591-596.

方法

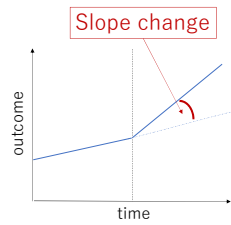
- ・分割時系列デザイン
- ・厚生労働科学研究調査研究班のDPCデータ + 病床機能報告2017
- ・観察期間：2012年4月～2018年3月

PECO

- P** 全ICU入室患者 (1:2看護患者比率)
- E** 政策導入後4年間 (2014年4月から2018年3月)
- C** 政策導入前2年間 (2012年4月から2014年3月)
- O**
 1. ICUベッド稼働率
 2. モニタリングおよび処置
 3. 臨床アウトカム(院内死亡, 合併症, 入院期間, 費用)

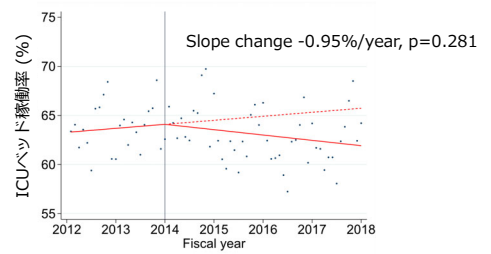
統計解析

- 分割時系列デザイン¹
- Impact modelは傾き変化のSlope changeを仮定

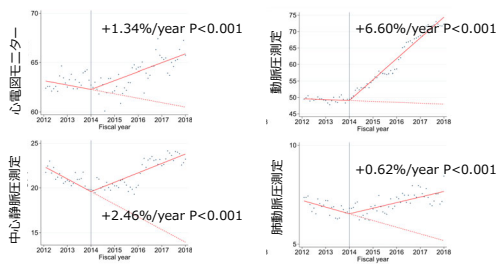


1. *Int J Epidemiol* 2017;46:348-355.

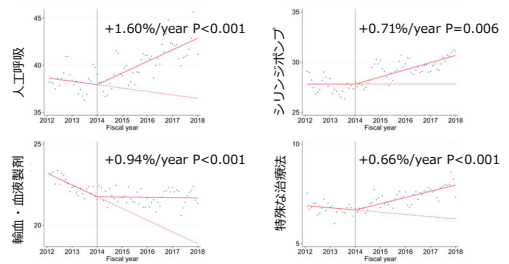
結果 -ICUベッド稼働率-



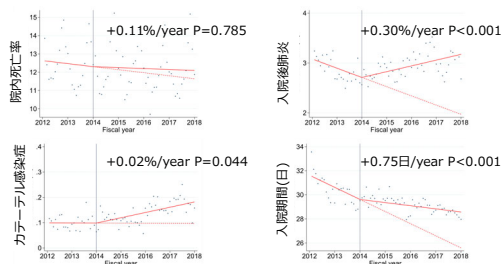
結果 -モニタリング及び処置①-



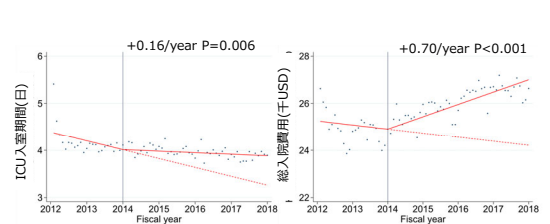
結果 -モニタリング及び処置②-



結果 -臨床アウトカム①-



結果 -臨床アウトカム②-



考察 -Study implication-

- 分割時系列デザインによる医療政策の評価を行った。
- **医療者の診療行為に基づくICU入室基準**は、ICU病床稼働率の減少とは関連せず、ICUでの診療行為・合併症・入院期間・入院費の増加と関連していた。

考察 -限界-

- 同時期に発生したイベントの影響。
- 重症度スコアなどの未測定時間依存性交絡因子。
- 他国への外的妥当性。

小括

以下の救急・集中治療領域のDPCデータ研究例を紹介した

1. 稀少疾患の記述疫学
2. 診療実態の分析
3. 医療の質研究、ヘルス・サービス・リサーチ

急性期医療に関する母集団代表性は高く、DPCデータを用いた類似の研究立案の参考になれば幸いです。

DeSCデータベースを用いた研究

東京大学大学院医学系研究科臨床疫学・経済学
康永秀生

本講義のトピック

1. はじめに
2. DeSCデータベースの概要
3. DeSCデータベースの母集団代表性
4. DeSCデータベースを用いた保険者種類別の有病率推計
5. DeSCデータを用いた研究例
6. まとめ

2

1. はじめに

リアルワールドデータ(real-world data, RWD)の種類

- 患者登録(patient registry)
- 保険データベース(administrative claims database)
- 電子カルテ情報(electronic medical records)など

3

RWD研究が隆盛するであろう理由

- (i)莫大な資金をかけて多数のRCTをやれるほどの財力が産・官・学のいずれにもない。
- (ii)RCTの対象から外される高齢患者が増加してきた。
- (iii)治療の選択肢も患者のニーズも多様化し、すべてをRCTで検証することは不可能。
- (iv)限られた答えしか得られないRCTでは、「エビデンスの隙間」を埋められない。

4

RWDの利点

- 疾患の疫学データなどがわかる。
- 薬だけでなく、手術・処置などあらゆる治療について検討可能。
- 臨床試験では分からない、実臨床における治療効果を明らかにできることもある。
- 薬剤の費用効果分析などにも応用可能。

RWDの欠点

- 適応による交絡(confounding by indication)を十分に調整しきれない。
- 入力されているデータの妥当性にやや難がある。

5

2. DeSCデータベースの概要

- DeSCデータベース：DeSCヘルスケア株式会社が、データ利活用事業の一環として、いくつかの健保・国保・高齢者医療広域連合より提供されたレセプト情報・特定健診情報の匿名加工情報。
- データベースに含まれる総人数は、2022年10月時点で約1110万人。

6

レセプト情報

被保険者台帳データ

匿名化ID、生年月、性別、保険者種別、本人または家族の別、レセプト開始・終了年月、資格取得・喪失年月など

医科・歯科・調剤の各レセプト

傷病名 (ICD-10コードおよび日本語テキスト)

診療行為とその算定日

医薬品情報 (医薬品名称とその算定日、量)

特定健診情報

身長・体重、腹囲、血圧、血液・尿検査、保健指導レベル、生活歴・現病歴に関する質問票の情報など

7

3. DeSCデータベースの母集団代表性

Okada, A, Yasunaga H. Prevalence of Non-communicable Diseases in Japan Using a Newly Developed Administrative Claims Database Covering Young, Middle-aged, and Elderly People. *JMA Journal* 2022;5(2):190-198

- 健保の被保険者はサラリーマン等、国保の被保険者は自営業者や退職者等であり、両者は年齢分布、年収を含む社会経済状況、健康状態が異なっていると考えられる。
- 健保、国保および後期高齢者のいずれか単独のデータでは、母集団から乖離した標本になる。

8

DeSCデータと他の統計情報を比較し、DeSCデータの母集団代表性を検証

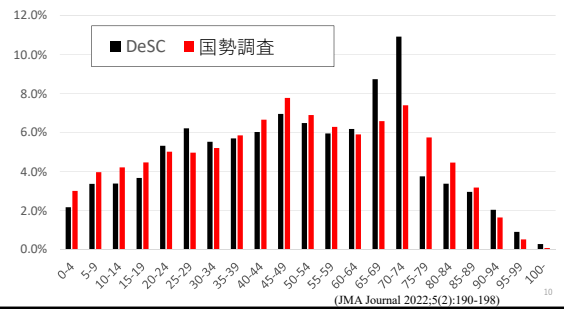
- 集団の年齢分布を、DeSCデータと国勢調査の間で比較
- 糖尿病・高血圧の有病割合を、DeSCデータと国民健康・栄養調査の間で比較
- 胃がん手術の実施割合を、DeSCデータとNDBオープンデータの間に比較

DeSCデータの中で、健保と国保を比較

- うつの年齢階級別有病割合
- 乳がんの年齢階級別有病割合

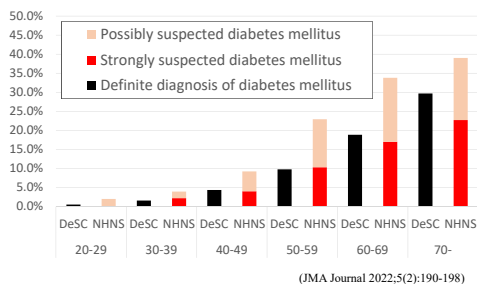
9

DeSCデータベースと国勢調査(2019年)の年齢分布比較



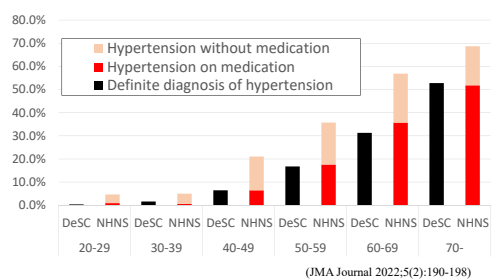
10

DeSCデータベースと国民健康・栄養調査(NHNS)の比較 ～糖尿病の有病割合



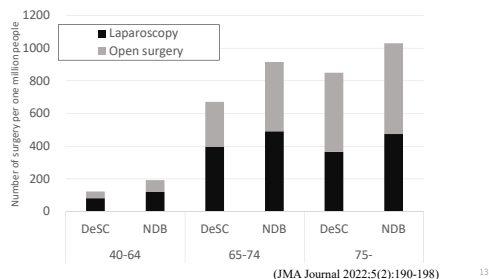
11

DeSCデータベースと国民健康・栄養調査(NHNS)の比較 ～高血圧の有病割合



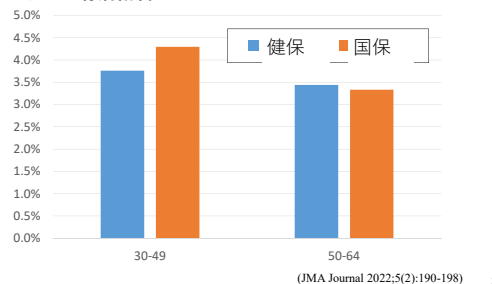
12

DeSCデータベースとNDBオープンデータの比較 ～胃がん手術の100万人当たり実施件数



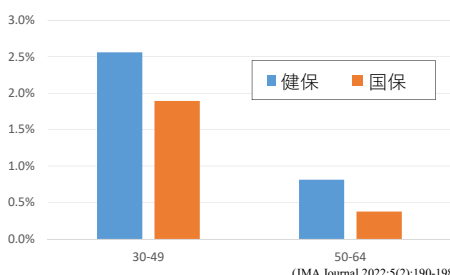
(JMA Journal 2022;5(2):190-198) 13

DeSCデータベース内での健保データと国保データの比較 ～うつの有病割合



(JMA Journal 2022;5(2):190-198) 14

DeSCデータベース内での健保データと国保データの比較 ～乳がんの有病割合



(JMA Journal 2022;5(2):190-198) 15

DeSCデータベースと国勢調査では、年齢分布がほぼ類似。
(65-74歳の年齢層のみ、DeSCデータベースにおける割合がやや高かった)

DeSCデータベースと国民健康・栄養調査では、糖尿病・高血圧の年齢階級別の有病割合がほぼ類似。

DeSCデータベースとNDBオープンデータでは、年齢階級別の胃がん手術実施件数(人口百万人当たり)がほぼ類似。

うつの有病割合は、30-49歳では健保より国保の方が高い傾向。
乳がんの有病割合は、30-64歳で健保より国保の方が低い傾向。

16

4. DeSCデータベースを用いた保険者種類別の有病率推計

岡田 啓, 康永 秀生. DeSC データベースの概要と臨床疫学・薬剤疫学研究への活用. 薬剤疫学 2022;27(1)11-18

- 健保と国保に含まれる集団は社会経済状況等が異なり、健康状態も異なると考えられる。しかし、両集団における疾病の有病率や健康指標の差異については、単一のデータベースを用いて比較した研究はほとんどない。
- 本稿では、DeSCデータベースを用いて、健保・国保・後期高齢者の保険者種類別に、糖尿病・高血圧・心血管疾患・悪性腫瘍等の有病率を推計し、さらに特定健診で得られるBody Mass Index (BMI)と収縮期血圧の分布を記述することとした。

17

方法

データソース: 2021年6月時点でのDeSCデータベース
対象: 2019年10月から2020年9月までの20歳以上の男女(N=1,685,438)
年齢層別(20歳代, 30歳代, 40歳代, 50歳代, 60歳代, 70歳以上)、
保険者種類別に、1年有病率を算出。

保険者種類別の構成:
健保669,215人(39.7%)、国保715,750人(42.5%)、後期高齢者300,473人(17.8%)

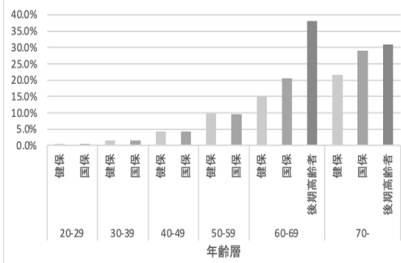
対象とする疾患:
糖尿病(ICD-10コード:E10-14)、高血圧(I10-15)、虚血性心疾患(I20-25)、心不全(I50)、脳卒中(I60-63)、大腸がん(C18-20)

BMIと血圧の値が欠測でない30歳以上の対象者(N=651,435人)について、特定健診情報を用いて、BMIと収縮期血圧の性・年齢層別の分布を保険者種類別に記述。

18

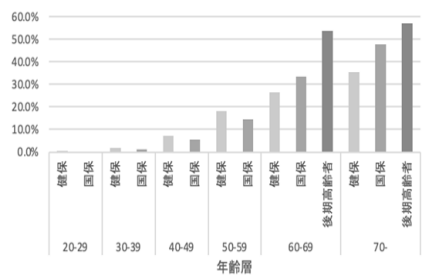
結果

糖尿病の1年有病率（保険者種別・年齢層別）



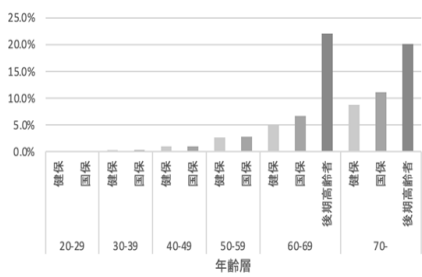
19

高血圧の1年有病率（保険者種別・年齢層別）



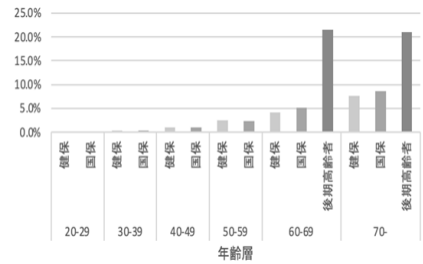
20

虚血性心疾患の1年有病率（保険者種別・年齢層別）



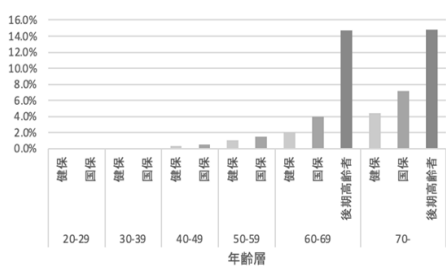
21

心不全の1年有病率（保険者種別・年齢層別）



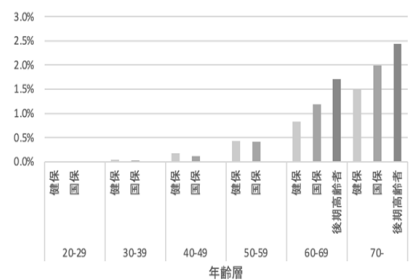
22

脳卒中の1年有病率（保険者種別・年齢層別）

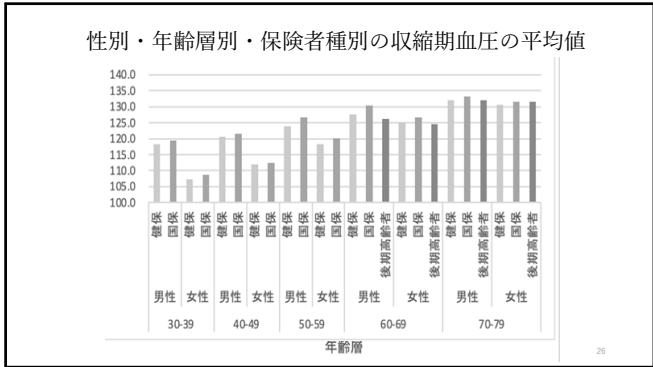
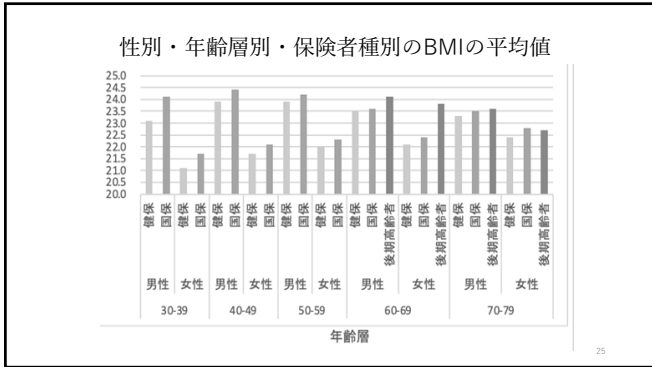


23

大腸がんの1年有病率（保険者種別・年齢層別）



24



考 察

- 糖尿病・高血圧の有病率は、59歳までの年齢層では健保よりも国保の方が低くなっており、60歳以上の層では逆に健保よりも国保の方が高くなっていった。これは、健保の方が特定健診を受診する割合が高く、59歳までの糖尿病・高血圧が受診を契機として発見されやすいことを反映しているかもしれない。
- 虚血性心疾患・心不全の有病率は、59歳までは健保と国保で同程度、60歳以上では、健保よりも国保の方が高い傾向が認められた。
- 脳卒中の有病率は、いずれの年齢層においても、健保よりも国保の方が高かった。

- 大腸がんの有病率は、59歳までは健保よりも国保の方がわずかに低く、60歳以上では健保よりも国保の方が高かった。
- 特定健診を受診した集団内の分析において、同じ性・年齢層の中でも、BMI・収縮期血圧がわずかながら健保よりも国保の方が高い傾向が示された。
- 上記のいずれも、国保の被保険者は健保の被保険者よりも健康状態が良くないことを示唆する。

5. DeSCデータを用いた研究例

骨盤臓器脱に対する腹腔鏡下およびロボット支援下仙骨腔固定術後の有害事象と再治療

Shigemitsu D, Okada A, Yasunaga H. Postoperative adverse events and re-treatment among patients who have undergone laparoscopic and robotic sacrocolpopexy for pelvic organ prolapse in Japan. *Int J Gynaecol Obstet.* 2022. PMID: 36200666

骨盤臓器脱 (pelvic organ prolapse, POP) は女性に比較的によく見られる疾患であり、近年合併症の少ない腹腔鏡下やロボット支援下仙骨腔固定術が行われている。この研究の目的は、日本人のPOP患者における腹腔鏡下およびロボット支援下仙骨腔固定術後の有害事象と再発に対する再治療を記述し比較することである。

方 法

データソース：DeSCデータベース 2014年4月から2021年5月まで
 対象：腹腔鏡下またはロボット支援下仙骨腔固定術を受けた患者

患者背景と術後アウトカムを腹腔鏡下仙骨腔固定術群とロボット支援下仙骨腔固定術群で比較。
 アウトカム：
 複合有害事象（膣びらん、術後尿失禁、術後排尿困難、尿路傷害、腹部切開ヘルニア）
 POP再発に対する再治療（ペッサリーの使用とPOPに対する手術）

結 果

POPと診断された患者28,748人
腹腔鏡下仙骨腔固定術 409人 (1.4%)
ロボット支援下仙骨腔固定術 52人 (0.2%)

術後有害事象の割合
腹腔鏡下仙骨腔固定術群 20.8%
ロボット支援下仙骨腔固定術群 13.5% (P=0.27)

ロボット支援下仙骨腔固定術の後に手術を受けた患者は1名 (1.9%) であり、両群とも術後にベッサリーを投与された患者はいなかった。

31

結 論

本邦において、腹腔鏡下およびロボット支援下仙骨腔固定術の術後アウトカムは、欧米諸国の報告と同程度であり、遜色はなかった。

32

6. まとめ

- DeSCデータベースは、健保・国保・後期高齢者のすべてのデータを含んでおり、小児から後期高齢者までの全年齢層をカバーし、母集団代表性が一定程度担保されている。
- DeSCデータベースを利用することにより、NDB以外では難しかった、全年齢層の外来・入院を含めたレセプト情報・特定健診情報を用いた臨床疫学・薬剤疫学研究が可能と考えられる。

33

ご清聴ありがとうございました

34

NDB・DPCデータベース研究人材育成Webinar

JMDCデータを用いた研究

東京大学医学部循環器内科（先進循環器病学講座）

金子 英弘

JMDC Claims Databaseとは？

- 株式会社JMDCが提供する保険者（健保）から集められた保険レセプトデータベース
- データ入手先: 保険者（主に大企業）
- 収集期間: 2005-現在（毎月更新）
- 累積母集団数: 約1400万人（2022年2月時点）
(<https://www.jmdc.co.jp/jmdc-claims-database/>)
- データ種別: 外来、入院、調剤レセプトデータ、（特定）健診データ、被保険者台帳、気象情報(集計データのみ)

Agenda

1. JMDCデータの概要
2. 解析で注意すべき点
3. 研究紹介

JMDC Claims Databaseの特徴

- 20-65歳のデータが最も多い（男性が女性より若干多い）
- 家族IDがある
- 小児のデータも含まれる
- 高齢者のデータが乏しい（最高齢75歳）
- 数年の経過を追跡可能（平均観察期間は3.5年程度）
- 保険から離脱しない限り複数の医療機関を受診しても追跡可能

Agenda

1. JMDCデータの概要
2. 解析で注意すべき点
3. 研究紹介

JMDC Claims Databaseに含まれる健診データ

- 問診（特定健診の問診票）
喫煙歴、飲酒歴、運動習慣、睡眠の質など
- BMI、腹囲（特定健診受診者） ※身長・体重データは含まれない
- 血圧 ※脈拍は含まれない
- 血液検査
血糖、HbA1c、脂質プロファイル、尿酸、クレアチニンなど
- 尿検査（尿糖、尿蛋白定性）
- 眼底検査

※項目ごとに欠損値が存在すること、また欠損値が存在する比率は項目毎に異なる点に注意が必要。

JMDC Claims Databaseの利点と限界

利点

- 外来レセプト、外来処方、調剤薬局データが存在する
- 病院を変更しても追跡可能
- 特定健診データが存在する
- 保険者台帳が存在する
- 家族IDが存在する

限界

- DPCの様式1情報がない
- 疾病の重症度指標がない
- 通常診療の検査値がない
- 高齢者の解析には向き
- SESの高い集団に偏っている
- 観察期間が比較的短い

注意すべき点（2）リスク調整

- 処方や処置、手術など保険算定される因子についてはデータが充実している
- 傷病名の妥当性については更なるValidation研究が必要
- 個々の疾病の重症度や進行度（例. がんのステージ）については情報が乏しい（JMDCが提供するDPCデータには一部存在）
- 検査値は健診時のみデータが存在する

Agenda

1. JMDCデータの概要
2. 解析で注意すべき点
3. 研究紹介

傷病名の妥当性

	n	Prevalence	感度(%)	特異度(%)
高血圧	286,139	23.7		
病名コード			80.7	95.3
薬コード			75.0	97.9
病名+薬			74.5	98.2
糖尿病	52,014	8.3		
病名コード			91.4	92.5
薬コード			78.6	99.5
病名+薬			78.6	99.6
脂質異常	277,707	38.4		
病名コード			48.9	90.4
薬コード			34.6	97.2
病名+薬			34.5	97.2

J Clin Epidemiol. 2018 Jul;99:84-95.

注意すべき点（1）データハンドリング

- リレーショナルデータベース
- 様々なデータが複数のテーブルで存在
- 必要な変数を自分で集める
- 統計ソフトだけでなくSQLのスキルが必要

加入者ID	受診日	病名コード	・・・	加入者ID	処方日	薬剤コード	・・・
0001	20151103	J09		0001	20151103	10067130	
0002	20151010	J10		0005	20151009	10002295	
0003	20141123	J11		0007	20141023	10067130	

リスク調整（対応策）

- 処方・処置・手術などの保険点数データの活用
 - 高血圧病名→降圧薬処方
 - ショック病名→カテコラミンの使用量
 - 出血量→輸血量
 - 呼吸不全→気管内挿管、人工呼吸器管理
 - 腎不全病名→透析
 - 手術時間→麻酔時間
 - リハの量→20分毎に点数加算
 - 腹腔ドレーン留置期間→ドレーン法 50点/日

リスク調整（対応策）

- ICD10コードや処置コードを用いた重症度指標
 - Charlson Comorbidity Index
 - ICD-10コードを用いた併存症指数
 - Yamana Severity Index
 - 処置データを用いた重症度指数
 - ICD-10-based trauma mortality prediction scoring system
 - ICD-10コードで外傷患者の死亡率を予測するスコア
 - ICD-10-Based Disability Predictive Index
 - ICD-10コードで外傷患者の障害を予測する指標

注意すべき点（5）Dropout

- JMDCは保険離脱によって追跡不能となる
 - 退職（病気、転職、定年、etc.）
 - 75歳
 - 被保険者である配偶者と離婚
 - 被保険者の子が就職
- 病気の重症化などによる退職が解析上は問題になる
 - Censoring Weight（打ち切りの重みづけ）などで対応

注意すべき点（3）交絡因子の調整

- 生物統計学的対応
 - 層別化
 - 多変量回帰分析
 - 傾向スコア分析（重症度のバランスング）
 - 操作変数法（疑似ランダム化）
 - 固定効果モデル（時間不変の交絡除去）
 - 自己対象研究デザイン（時間不変の交絡除去）

Agenda

1. JMDCデータの概要
2. 解析で注意すべき点
3. 研究紹介

注意すべき点（4）欠損値の取り扱い

- 欠損値のメカニズムの検討(MCAR, MAR, MNAR)
- Complete case analysis
- 単一代入法
- 多重代入法

JMDCはどんな研究に向いているか？

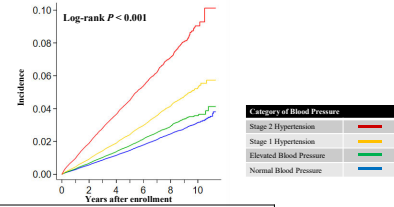
一般的に
“登録型の疫学コホートやRCTでは解決が困難な課題の検討”
にリアルワールドデータは強みを有します。

- 豊富なサンプル数・イベント数が必要な研究
- 悉皆性のあるデータが必要な研究
- 薬剤疫学研究

JMDCはどんな研究に向いているか？

- 豊富なサンプル数・イベント数が必要な研究
- 悉皆性のあるデータが必要な研究
- 薬剤疫学研究

血圧分類と共に心不全発症リスクは上昇する



Years after enrollment	Normal Blood Pressure	Elevated Blood Pressure	Stage 1 Hypertension	Stage 2 Hypertension
0	137,390	263,984	199,871	136,279
2	137,390	263,984	199,871	136,279
4	137,390	263,984	199,871	136,279
6	137,390	263,984	199,871	136,279
8	137,390	263,984	199,871	136,279
10	137,390	263,984	199,871	136,279

Circulation.
2021 Jun 8;143(23):2244-2253.

JMDCデータを用いた研究紹介（1）

背景：

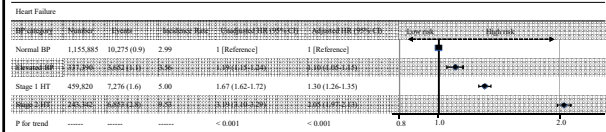
2017年に米国の血圧ガイドラインが改訂され、高血圧の定義を従来の血圧140/90 mmHg以上から、130/80 mmHgに引き下げて、血圧130-139/80-89 mmHgをステージ1高血圧、血圧140/90 mmHg以上をステージ2高血圧と定義した。しかし、ステージ1高血圧が心不全や心房細動などの心血管イベントのリスクと関連するかは明らかでなかった。

Point:

ステージ1高血圧における心血管イベント発症リスクは、ステージ2高血圧と比較して低いことが予想される。したがって、ステージ1高血圧における心血管イベント発症リスクを評価するには、豊富なサンプルサイズが必要となり、JMDCデータの強みが活かせる。

Circulation. 2021 Jun 8;143(23):2244-2253.

Elevated BP/Stage 1 高血圧から心不全リスクは上昇

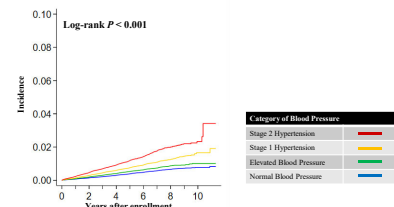


Circulation. 2021 Jun 8;143(23):2244-2253.

P	<ul style="list-style-type: none"> ・ JMDCデータベースに登録されている降圧薬を内服していない症例 ・ 心血管イベントの既往の無い症例 (n=2,196,437)
E	ステージ1高血圧 (SBP 130-139 mmHg or DBP 80-89 mmHg)
C	正常血圧 (SBP < 120 mmHg & DBP < 80 mmHg)
O	<ul style="list-style-type: none"> ・ 心不全 ・ 心房細動

Circulation. 2021 Jun 8;143(23):2244-2253.

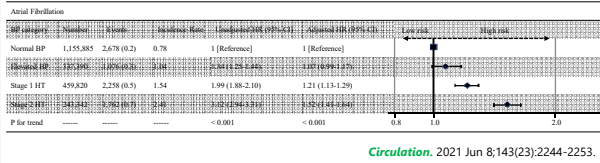
血圧分類と共に心房細動発症リスクは上昇する



Years after enrollment	Normal Blood Pressure	Elevated Blood Pressure	Stage 1 Hypertension	Stage 2 Hypertension
0	137,390	264,577	206,765	137,159
2	137,390	264,577	206,765	137,159
4	137,390	264,577	206,765	137,159
6	137,390	264,577	206,765	137,159
8	137,390	264,577	206,765	137,159
10	137,390	264,577	206,765	137,159

Circulation.
2021 Jun 8;143(23):2244-2253.

Stage 1 高血圧から心房細動リスクは上昇



P JMDCデータベースに登録されている20-49歳の症例で心血管イベントの既往の無い症例 (n=913,235)

E/C Life's Simple 7 Cardiovascular Health Metrics
※体重・喫煙・運動・食事・血圧・脂質・血糖で定義したModifiable Riskの指標

O 複合エンドポイント (心筋梗塞/脳卒中)

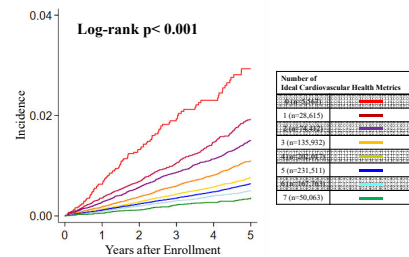
J Am Coll Cardiol. 2020 Nov 17;76(20):2414-2416.

JMDCデータを用いた研究紹介 (1)

結論:
本邦の一般人口において、正常血圧と比較してステージ1高血圧が将来の心不全や心房細動発症リスクと上昇することが示された。JMDCデータベースが有する豊富なサンプルサイズを活かして、堅牢な臨床エビデンスを創出できた一例。

Circulation. 2021 Jun 8;143(23):2244-2253.

Life's Simple 7によって Young Adultの循環器疾患発症リスクが層別化される



JMDCデータを用いた研究紹介 (2)

背景:
Young Adult世代の循環器疾患増加は先進国共通の疫学的課題であり、不健康な生活習慣などの"Modifiable Risk"が、Young Adult世代における生活習慣病や循環器疾患の増加に寄与すると考えられている。一方で、Young Adult世代を対象とした疫学研究は十分でなく、Modifiable Riskの指標である、Life's Simple 7 Cardiovascular MetricsによってYoung Adult世代の循環器疾患リスクの層別化が可能かは明らかでない。

Point:
Young Adult世代の循環器疾患発症リスクは、中高年世代と比較して低いため、Young Adult世代の循環器疫学研究には、医療ビッグデータ研究が有する大規模なサンプルサイズが適している。

J Am Coll Cardiol. 2020 Nov 17;76(20):2414-2416.

Life's Simple 7 Cardiovascular Health Metricsの各要素は Young Adultの循環器疾患発症リスクと独立して関連する

Ideal CHV metrics	Present	Adjusted Hazard ratio	95% Confidence Interval	P value
Ideal smoking Status	665,025	0.747	0.703-0.795	< 0.001
Ideal Body Mass Index	696,958	0.761	0.715-0.810	< 0.001
Ideal Physical Activity	380,741	0.948	0.881-0.999	0.044
Ideal Diet Habits	406,379	0.874	0.825-0.926	< 0.001
Ideal Blood Pressure	523,495	0.617	0.580-0.656	< 0.001
Ideal Fasting Plasma Glucose	757,115	0.806	0.753-0.861	< 0.001
Ideal Total Cholesterol	479,164	0.920	0.868-0.975	0.005

Each Cardiovascular Health Metrics Component and Cardiovascular Event

J Am Coll Cardiol. 2020 Nov 17;76(20):2414-2416.

JMDCデータを用いた研究紹介（2）

結論：

本邦のYoung Adult世代においても、Life's Simple 7 Cardiovascular Health Metricsを用いて循環器疾患（心筋梗塞/脳卒中）発症リスクが層別化されることが示された。Young Adult世代での循環器疾患の予防でModifiable Riskの重症性を示唆する研究である。Working Age Populationを豊富に有するJMDCの強みが活かされた研究の一例。

J Am Coll Cardiol. 2020 Nov 17;76(20):2414-2416.

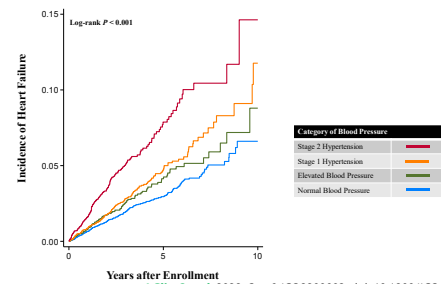
P	・JMDCデータベースに登録されている乳がん/大腸直腸がん/胃がんの既往のある症例 ・降圧薬を内服していない症例 ・心血管イベントの既往の無い症例 (n=33,991)
E	ステージ1 高血圧 (SBP 130-139 mmHg or DBP 80-89 mmHg) ステージ2 高血圧 (SBP ≥ 140 mmHg or DBP ≥ 90 mmHg)
C	正常血圧 (SBP < 120 mmHg & DBP < 80 mmHg)
O	・心不全

J Clin Oncol. 2022 Sep 8;JCO2200083. doi: 10.1200/JCO.22.00083.

JMDCはどんな研究に向いているか？

- 豊富なサンプル数・イベント数が必要な研究
- 悉皆性のあるデータが必要な研究
- 薬剤疫学研究

がんサバイバーにおいても血圧上昇と共に心不全リスクが上昇する



JMDCデータを用いた研究紹介（3）

背景：

がん治療の進歩などに伴ってがん患者の長期生存が可能となり、がんサバイバーにおける循環器疾患（とりわけ心不全）が臨床的課題となっている。高血圧は循環器疾患の最大の危険因子であるが、がんサバイバーにおける高血圧と循環器疾患発症リスクの関連についてはデータが乏しかった。

Point:

“腫瘍循環器学”と呼ばれる当該分野の疫学研究には、がんと循環器疾患と言う領域をまたいだデータベースが必要であり、JMDCデータが有する悉皆性のあるデータが応用できる。

J Clin Oncol. 2022 Sep 8;JCO2200083. doi: 10.1200/JCO.22.00083.

がんサバイバーにおいても血圧上昇と共に心不全リスクが上昇する -ステージ1 高血圧の段階から有意に心不全リスクが上昇-

	No. of Patients	No. of Events	Model 1	Model 2	Model 3	Low risk	High risk
Normal BP	37,444	3013 (7.7%)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	0.5	2.5
Elevated BP	4,733	111 (2.3)	1.31 (1.05-1.62)	1.19 (0.95-1.48)	1.15 (0.93-1.44)		
Stage 1 Hypertension	1,562	146 (9.2%)	1.46 (1.24-1.70)	1.25 (1.04-1.51)	1.24 (1.03-1.49)		
Stage 2 Hypertension	4,312	176 (4.1)	2.56 (2.13-3.09)	2.08 (1.72-2.52)	1.99 (1.63-2.43)		

Hazard Ratio (95% Confidence Interval) in Model 3

J Clin Oncol. 2022 Sep 8;JCO2200083. doi: 10.1200/JCO.22.00083.

JMDCデータを用いた研究紹介（3）

結論：

がんサバイバーを対象とした解析でも、ステージ1 高血圧の段階から心不全リスクは有意に上昇し、ステージ2 高血圧ではさらにそのリスクが上昇した。がんサバイバーにおいても血圧管理の重要性を示唆する研究である。本研究においては、JMDCデータが有する悉皆性のある健診データ・レセプトデータが大変有用であった。

J Clin Oncol. 2022 Sep 8;JCO2200083. doi: 10.1200/JCO.22.00083.

JMDCデータを用いた研究紹介（4）

背景：

SGLT2阻害薬は多くのRCTによって心血管イベント・腎イベント発症予防効果が報告されている。本邦では、6種類のSGLT2阻害薬が保険収載されており、我々は以前にこの6種類のSGLT2阻害薬間で心血管イベント発症リスクに差異が無い事を報告しているが (*Cardiovasc Diabetol.* 2022 May 18;21(1):67.)、腎イベントの発症リスクに差異があるかは明らかでなかった。

Point:

薬剤ブランド間でのイベントリスク比較はRCTの実装が困難である。このような研究においては、レセプトデータによる検討が有用である。

Kidney Int. 2022 Nov;102(5):1147-1153.

JMDCデータを用いた研究紹介（3）

結論：

本邦のYoung Adult世代においても、Life's Simple 7 Cardiovascular Health Metricsを用いて循環器疾患（心筋梗塞/脳卒中）発症リスクが層別化されることが示された。Young Adult世代での循環器疾患の予防でModifiable Riskの重症性を示唆する研究である。Working Age Populationを豊富に有するJMDCの強みが活かされた研究の一例。

J Am Coll Cardiol. 2020 Nov 17;76(20):2414-2416.

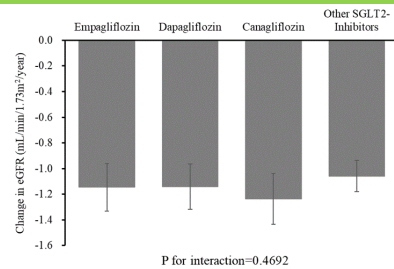
P	・JMDCデータベースに登録された糖尿病症例でSGLT2阻害薬が新たに処方された症例 (n=12,100)
E	Empagliflozin
C	Dapagliflozin、Canagliflozin、その他のSGLT2阻害薬
O	年率eGFRの低下

Kidney Int. 2022 Nov;102(5):1147-1153.

JMDCはどんな研究に向いているか？

- 豊富なサンプル数・イベント数が必要な研究
- 悉皆性のあるデータが必要な研究
- 薬剤疫学研究

年率のeGFR低下速度はSGLT2阻害薬間で同等



Kidney Int. 2022 Nov;102(5):1147-1153.

JMDCデータを用いた研究紹介（4）

結論：

本邦で保険収載されているSGLT2阻害薬のプラン間でeGFRの低下速度が同等であることが示された。SGLT2阻害薬の腎機能への影響がクラスエフェクトであることを示唆する結果である。レセプトデータを用いた薬剤疫学研究は、RCTが困難なClinical Questionに対しても応用可能である。

Kidney Int. 2022 Nov;102(5):1147-1153.

ご清聴ありがとうございました。



Agenda

1. JMDCデータの概要
2. 解析で注意すべき点
3. 研究紹介

総括

- JMDCデータは健診・レセプトデータを含む、本邦最大規模のリアルワールドデータベースであり、使用実績も豊富である。
- 若年成人や小児を対象とする研究、悉皆性のあるデータが必要な研究には強みがある一方で、高齢者を対象とする研究や、疾病の重症度、詳細な検査値が重要な研究には不向きである。
- 登録型コホートやRCTでは解決困難なClinical Questionの検証に有用となる可能性がある。
- データベースの利点・欠点を理解したうえでの研究立案、データハンドリング、統計解析、結果の解釈が重要である。

傾向スコア

東京大学大学院医学系研究科臨床疫学・経済学
東京大学大学院医学系研究科乳腺・内分泌外科学
小西孝明

目次

1. 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
2. 傾向スコアマッチング
3. 逆確率重み付け、オーバーラップ重み付け
4. 先行研究
5. 傾向スコア分析の注意点

4

Pubmedで検索した“Propensity Score”関連論文数

観察データを用いて擬似ランダム化を行い、
ランダム化比較試験に準じる結果を得られる



2

目次

1. 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
2. 傾向スコアマッチング
3. 逆確率重み付け、オーバーラップ重み付け
4. 先行研究
5. 傾向スコア分析の注意点

5

本講義の目標

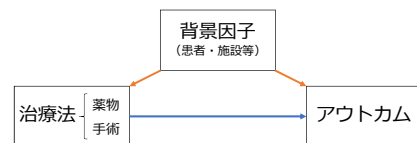
▶本講義の目標

- ✓ 傾向スコア分析の基礎を理解する
- ✓ 傾向スコアマッチングについて学ぶ
- ✓ 傾向スコアを用いた逆確率重み付け、オーバーラップ重み付けについて学ぶ
- ✓ 傾向スコア分析を用いた先行研究について学ぶ
- ✓ 傾向スコア分析の注意点について学ぶ

3

概念〈観察研究の交絡〉

治療法の選択とアウトカムの両方に影響を及ぼす因子は交絡となる



軽症例に薬物療法、重症例に手術治療
→見かけ上、薬物療法の成績が良くなる

6

概念<観察研究の交絡>

治療群間で背景因子は異なるのが自然

	治療A	治療B
人数	700	700
男性	55%	60%
高血圧	70%	40%
糖尿病	60%	30%
死亡	10%	7%

7

概念<傾向スコア>

>傾向スコアとは

- 観察された背景因子のもとで、ある治療を受ける確率
- Propensity score (PS)

>傾向スコアの特徴

- 傾向スコアが近い人は、ほぼ同じ確率でその治療を受ける
- 傾向スコアが近い人は、ほぼ同じ特性を持っている
- 傾向スコアを介して複数の交絡因子に対処し、群間のバランスをとれる

10

概念<マッチング>

治療A				治療B			
ID	性別	高血圧	糖尿病	ID	性別	高血圧	糖尿病
1	男	1	0	8	男	1	1
2	女	1	0	9	女	0	0
3	男	1	1	10	男	0	0
4	男	0	1	11	男	0	0
5	女	0	0	12	女	0	0
6	男	1	1	13	男	1	0
7	女	1	1	14	女	1	1

8

傾向スコアの推定<理論>

ある治療を受ける確率 p を、背景因子 X_n を独立変数とするロジスティック回帰で求める (β_0 は切片、 β_n は係数)

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

$$\therefore p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}$$

この確率 p が傾向スコアである

傾向スコア p は複数の背景因子が治療の割り当てに与える影響の度合いを1つの値にまとめたもので、**これがほぼ等しければその背景因子はほどほどに近い**といえる

11

概念<マッチング>

治療A				治療B			
ID	性別	高血圧	糖尿病	ID	性別	高血圧	糖尿病
1	男	1	0	13	男	1	0
3	男	1	1	8	男	1	1
5	女	0	0	12	女	0	0
7	女	1	1	14	女	1	1

似た背景因子の患者を選択 → ランダム化比較試験に類似できる
変数が多い場合には完全一致は難しいので、ほぼ同じを目指すのが傾向スコア

9

傾向スコアの推定<例>

- 従属変数: 治療A
- 独立変数: 年齢、性別、高血圧、糖尿病、...

ID	治療	年齢 ($\beta=.23$)	性別 ($\beta=-10$)	高血圧 ($\beta=.37$)	糖尿病 ($\beta=.36$)	...	予測確率 p
1	A	75	女	0	0	...	0.26
2	B	74	女	0	0	...	0.22
3	A	76	男	1	1	...	0.73
4	B	81	男	1	0	...	0.70

12

傾向スコアの推定 <実際>

>統計ソフトで計算

ロジスティック回帰を行い、予測確率 p を計算して保存
 従属変数：治療割り当て変数（治療Aを受けたかどうか）
 独立変数：背景因子（年齢、性別、高血圧、糖尿病、...）

通常のロジスティック回帰と異なり、独立変数に投入する背景因子に関して

- 投入しすぎを気にしなくてよい（傾向スコア分析の利点）
 - ✓ 過剰適合や多重共線性は考慮しなくて良い
- 治療選択より後に起きるものは投入しない（cf. 注意点）
 - ✓ 例、在院日数、合併症

13

目次

1. 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
2. 傾向スコアマッチング
3. 逆確率重み付け、オーバーラップ重み付け
4. 先行研究
5. 傾向スコア分析の注意点

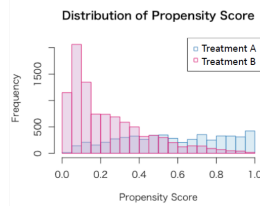
16

傾向スコアの確認 <ヒストグラム>

• 治療Aと治療Bの傾向スコアの分布をヒストグラムで確認する

- ✓ 傾向スコアが1に近いほど治療Aが選択
- ✓ 傾向スコアが0に近いほど治療Bが選択

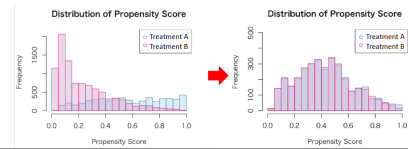
• ほどよい重なりが望ましい
 ✓ その理由は次ページに...



14

マッチングの手順

- ① 傾向スコアの推定
- ② 同じくらいの傾向スコアを持つペアを選択する
- ③ 二群のバランスの確認（後述）
- ④ 二群でアウトカムを比較する（t検定、カイ二乗検定など）

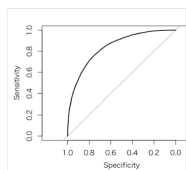


17

傾向スコアの確認 <c統計量>

• 傾向スコアが治療Aの割り当てをどの程度予測するか？

- ✓ ROC (receiver operating characteristic) 曲線を描く
- ✓ その曲線下の面積 (AUC, area under curve) を算出する
- ✓ その値をc統計量と呼ぶ
 - 1.0 : 完全識別
 - 0.9-1.0 : 治療A群と治療B群の重なりが少ない（治療適応が異なるので効果比較の意味がない）
 - 0.6-0.9 : 適切と判断されることが多い
 - 0.5-0.6 : 治療A群と治療B群の重なりが多い（通常の多変量解析でも結果が変わらない）
 - 0.5 : 識別力無し



15

ペア選択のルール設定

- A) マッチング方法
 - ✓ 最近傍、最適、など
 - ✓ 1対1、1対2、1対3、...
- B) キャリバー（閾値）
 - ✓ 最近傍マッチングの場合、ペアの傾向スコアの差の絶対値が閾値に収まるようにする
 - ✓ 傾向スコア（logit）の標準偏差を基準
 - ✓ 標準偏差×0.2に設定されることが多い
- C) 抽出方法
 - ✓ 復元、非復元

18

二群のバランスの確認

▶ 二群間のStandardized difference (SD)を各変数ごとに計算する

- ✓ マッチング前と後にそれぞれ算出する
- ✓ マッチング後にSDが0に近づいていれば二群の差異が減少している
- ✓ SDの絶対値が0.1未満でバランスが良いとみなすことが多い
- ✓ 0.1以上の場合、傾向スコア算出のモデルを修正する(変数を増やすなど)

p値の使用は推奨されない
(nの影響を受けるため)

カテゴリ変数:

$$SD = \frac{|p_A - p_B|}{\sqrt{\frac{p_A(1-p_A) + p_B(1-p_B)}{2}}}$$

各群における割合 p_A , p_B を用いる

連続変数:

$$SD = \frac{|\bar{x}_A - \bar{x}_B|}{\sqrt{\frac{s_A^2 + s_B^2}{2}}}$$

各群における平均 \bar{x}_A , \bar{x}_B と分散 s_A , s_B を用いる

19

目次

1. 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
2. 傾向スコアマッチング
3. 逆確率重み付け、オーバーラップ重み付け
4. 先行研究
5. 傾向スコア分析の注意点

22

論文中の記載例

Method

- ✓ 治療を予測する変数として X_1, X_2, \dots を用いたロジスティック回帰を行った
- ✓ 推定された傾向スコアを用い、非復元抽出による1対1の最近傍マッチングを行った
- ✓ キャリバーは傾向スコアの標準偏差の0.2倍に設定した

Result

- ✓ 傾向スコアマッチングでN人を抽出した
- ✓ そのc統計量は0.78であった(←c統計量は必ずしも記載しなくて良いとされる場合もある)
- ✓ Table1は、マッチング前後の背景因子である

20

傾向スコアを用いた重み付け

- 傾向スコアマッチングでは、マッチした患者のみを抽出して解析した
 - ✓ 傾向スコアの近い患者をマッチングさせて抽出している
 - ✓ そのためマッチする相手がいなかった患者は除外している
- 傾向スコアを用いた「重み」を掛けて、全ての患者を解析に利用する
 - ✓ 逆確率による重みづけ (Inverse probability of weighting, IPW)
 - ✓ オーバーラップ重み付け (Overlap weighting)

23

Table1の記載例

	マッチング前			マッチング後		
	治療A	治療B	SD	治療A	治療B	SD
	n = 32,630	n = 26,764		n = 23,055	n = 23,055	
男性	6,355 (20)	7,668 (29)	0.213	5,529 (24)	5,452 (24)	-0.008
年齢						
<55	12,427 (38)	7,528 (28)	-0.213	7,479 (32)	7,395 (32)	-0.008
55-65	10,969 (34)	8,568 (32)	-0.034	7,708 (33)	8,122 (35)	0.038
≥65	9,234 (28)	10,668 (40)	0.246	7,868 (34)	7,538 (33)	-0.003
高血圧	3,838 (12)	3,228 (12)	0.009	2,735 (12)	2,928 (13)	0.003
糖尿病	3,130 (9.6)	2,575 (9.6)	0.001	2,249 (9.8)	2,233 (9.7)	-0.002

*SDを絶対値(ASD, Absolute SD)や%を用いて表記することもある SD, Standardized difference

21

観察している効果の違い

- 傾向スコアマッチングでは、ATTを観察している
 - ✓ ATT: 治療群における平均処置効果 (average treatment effect on the treated)
 - ✓ 治療A群の患者が仮に治療Bを受けた時のアウトカムの差
- 傾向スコアを用いた重み付けでは、ATEを観察できる
 - ✓ ATE: 平均処置効果 (average treatment effect)
 - ✓ 全患者が治療Aと治療Bを受けた時のアウトカムの差

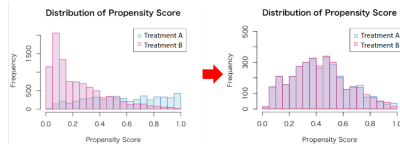
	治療A	治療B
人数	700	700
男性	55%	60%
高血圧	70%	40%
糖尿病	60%	30%
死亡	10%	7%

24

重み付けの手順

マッチングと同様の流れ

- ① 傾向スコアの推定
- ② 患者をそれぞれの傾向スコアで重み付けする
- ③ 二群のバランスの確認
- ④ 二群でアウトカムを比較する (検定、カイニ乗検定など)



25

目次

1. 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
2. 傾向スコアマッチング
3. 逆確率重み付け、オーバーラップ重み付け
4. 先行研究
5. 傾向スコア分析の注意点

28

重み付けの計算式

PSは治療Aを受ける確率
1-PSは治療Bを受ける確率

- 患者各々の傾向スコア(PS)を用いて重み付けをする

	治療A群	治療B群
逆確率による重み付け	1/PS	1/(1-PS)
オーバーラップ重み付け	(1-PS)	PS

- 逆確率による重み付けでの例

- ✓ A群に傾向スコアが0.1の人が10人いたら、 $10/0.1 = 100$ 人いるものと見なす
- ✓ B群に傾向スコアが0.6の人が20人いたら、 $20/(1-0.6) = 50$ 人いるものと見なす
- 重み付け後は各群の人数が増え、多くの場合非整数になる

26

提示する先行研究

- Mortality and morbidity after Hartmann's procedure versus primary anastomosis without a diverting stoma for colorectal perforation: a nationwide observational study
- A. Tsuchiya, H. Yasunaga, Y. Tsutsumi, H. Matsui, K. Fushimi
- World Journal of Surgery 2018;42:866-875

『DPCを用いた研究2』では、傾向スコアを用いた別の研究を例示しています

29

重み付けの利用の実際

- 主解析として利用されることはまだ少ない
 - ✓ マッチングより直感的に分かりにくい
 - ✓ 優れた解析手法ではあるが、読者に理解されにくい可能性がある
- 感度分析として利用されることが多い
 - ✓ 主解析でマッチングを行い、感度分析でいずれかの重み付けを行う
 - ✓ マッチングと重み付けを両方行うことで、ATTとATEの両方を観察できる

27

PECOと解析手法

- P: 大腸穿孔に対して緊急開腹手術を受けた患者
- E: 一時的にストマを作る手術 (Hartman's procedure)
- C: ストマを作らずに吻合手術
- O: 30日死亡、合併症、術後集中治療

➤ 傾向スコアマッチング、逆確率による重み付け、操作変数法

30

Method <傾向スコアマッチング>

- 段落の冒頭で主解析の方法を明示
 "One-to-one propensity score matching was performed..."
- 傾向スコアの作成方法
 "To estimate the propensity score, a logistic regression model was used with the baseline independent variables (details are described in Supplemental Text)."
- マッチングのルール設定 (マッチング方法、キャリパー、抽出方法)
 "Using a nearest-neighbor matching method, each patient in the PA group was matched with one patient in the HP group without replacement, with the closest estimated propensity within a caliper (0.2 standard deviations of the propensity score)."
- バランスの確認方法
 "The balance in the baseline variables between the propensity-matched HP and PA groups was examined using standardized differences, where >10% was regarded as imbalanced."

31

Table 1 <バランスの確認>

	Unmatched groups			Matched groups		
	Primary anastomosis n = 3,045	Hartmann's procedure n = 5,455	SD(%)	Primary anastomosis n = 2,800	Hartmann's procedure n = 2,800	SD(%)
Age, years						
15-59	745 (24.5)	779 (14.3)	26.0	503 (18.0)	567 (20.3)	-5.8
60-69	542 (17.8)	1,130 (20.7)	-7.4	542 (19.4)	519 (18.5)	2.3
70-79	836 (27.5)	1,646 (30.2)	-6.0	835 (29.8)	804 (28.7)	2.4
≥80	922 (30.3)	1,900 (34.8)	-9.6	920 (32.9)	910 (32.5)	0.9
Sex male	1,680 (55.2)	2,538 (46.5)	17.5	1,460 (52.1)	2,928 (52.3)	-0.4
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Data are presented as n and (%).

SD, Standardized difference

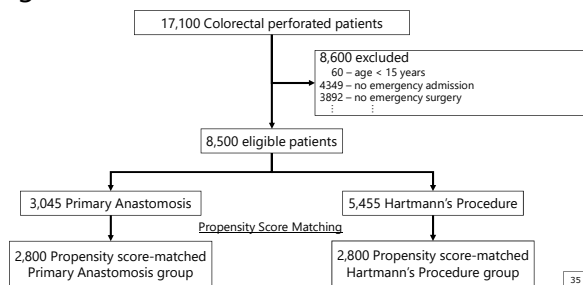
34

Method <感度分析>

- 感度分析として逆確率による重み付けを実施
 "We also used a propensity score method for inverse probability of treatment weighting (IPTW) using the same population as that in the propensity score matching analysis."
 "Each patient was weighted by the inverse probability of being in the observed group."
- 本研究では操作変数法も実施
 ✓ 操作変数法については別講義を参照ください

32

Figure 1 <患者抽出の流れ>



35

Result <マッチングの結果>

- 最初の段落 (抽出された人数、統計量)
 "A total of 8500 eligible patients with colorectal perforation were treated during the study period."
 "There were 5455 HP patients and 3045 PA without diverting stoma patients, from which 2800 propensity score-matched pairs were generated."
 "The C-statistic was 0.62 in the model for calculating propensity scores."
- 次の段落でTable1の説明 (バランスの確認)
 "Table 1 shows the baseline characteristics of the unmatched and propensity score-matched groups..."
 "The distributions of the variables in the propensity score-matched groups were well balanced."

33

Table 2 <マッチング後のアウトカム>

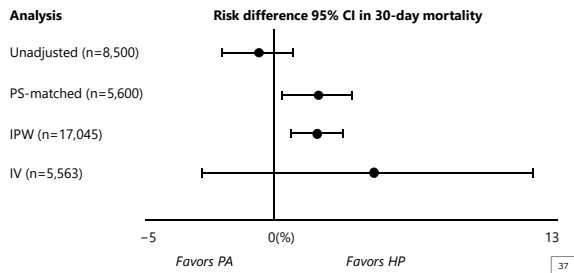
	Matched groups		RD	95% CI
	Primary anastomosis n = 2,800	Hartmann's procedure n = 2,800		
Overall postoperative complications	699 (25.0)	641 (22.9)	2.1	-0.2 to 4.3
Surgical complications				
Overall surgical interventions under GA	235 (8.4)	129 (4.6)	5.8	2.5 to 5.1
⋮	⋮	⋮	⋮	⋮

Data are presented as n and (%).

RD, Risk difference

36

Figure 2 <各解析の結果>



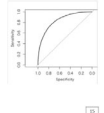
適度な群間差

>治療適応の適度な一致が必要

- c統計量が大きい(0.9-1.0)
 - ✓ 2群で背景が大きく異なる
 - ✓ 一般化可能性が低い
 - ✓ 比較することにあまり意味がない
- c統計量が小さい(0.5-0.6)
 - ✓ 2群で背景が類似している
 - ✓ 一般的な回帰分析で十分比較可能
 - ✓ 傾向スコアを用いることを否定はされない

傾向スコアの確認<c統計量>

- 傾向スコアが治療Aの割り当てをどの程度予測するか?
 - ✓ ROC (receiver operating characteristic) 曲線を描く
 - ✓ その曲線下の面積 (AUC, area under curve) を算出する
 - ✓ その値も、傾向スコアと等しい
- 1.0 : 完全予測
- 0.9-1.0 : 治療A群と治療B群の差が大きい(予測精度が高い)
- 0.6-0.9 : 適度に予測される(予測精度が低い)
- 0.5-0.6 : 治療A群と治療B群の差が小さい(予測精度が低い)
- 0.5 : 予測が難しく、傾向スコアが治療A群と治療B群の差を予測できていない



目次

- 傾向スコアの基礎
 - ① 概念
 - ② 傾向スコアの推定
 - ③ 傾向スコアの確認
- 傾向スコアマッチング
- 逆確率重み付け、オーバーラップ重み付け
- 先行研究
- 傾向スコア分析の注意点

二群のバランス

>傾向スコアによる調整後にバランス不良であった際の対処

- SD>0.1の際にはバランスがとれていない
- 傾向スコアによる調整後に以下の方法を行うとバイアスを生み出す
 - ✓ SD>0.1となった(バランス不良の)変数でさらに調整する
 - ✓ 傾向スコアの算出に用いなかった別の交絡因子でさらに調整する

>傾向スコア作成の際に投入する変数を工夫して対処する

- カテゴリー変数の区分を変更する
- 二乗項・交互作用項(例、年齢×BMI)を投入する

未測定の変数

>未測定の変数には対処できない

- あくまで測定された変数を制御しているに過ぎない
- 未測定の変数によるバイアスにはさらされている
 - ✓ 糖尿病発症の研究で、HbA1cが入手できていない場合、など
- 未測定の変数によって結果が反転する場合もある

>事後対処法はない

- 事前の研究計画・情報収集が肝要
- 感度分析(操作変数法など)やE-valueによって頑健性(robustness)を示す
- 高次元傾向スコアが有用かもしれない

操作変数法や高次元傾向スコアは別講義を参照ください

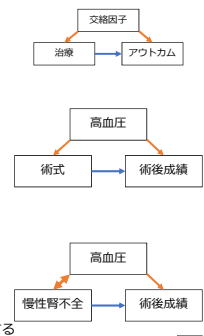
傾向スコア解析の対象

>治療介入(や短期の曝露)は評価可能

- 治療介入の評価のために傾向スコアは開発された
- 「治療介入より前に測定された因子」を明確に定義可能
 - ✓ それらの因子によって治療介入を受ける確率を算出できる
 - ✓ 術式が原因で(避けて)高血圧になることはない

>患者の状態(や長期の曝露)は評価不能?

- 「患者の状態より前に測定された因子」を明確に定義不能
 - ✓ 患者の状態より後に交絡因子が発生している可能性
 - ✓ 慢性腎不全によって高血圧になる場合がある
- 通常の変数解析やMatched-pair cohort解析などを検討する



おわりに

- 傾向スコアを用いた解析では、複数の交絡因子に対処し、**群間のバランスをとることができる**
- 傾向スコアマッチングでは、**c統計量**などを確認して記載する
- 傾向スコアを用いた重み付けでは、**全患者におけるATE**を見られる
- 未測定 of 交絡因子・適度な群間差・バランスに注意して使用する

▼ 高次元傾向スコア high-dimensional propensity score, hd-PS

石丸美穂
東京医科歯科大学医歯学総合研究科健康推進歯学分野
プロジェクト助教

1

目標

- 傾向スコアを大規模データの性質に合わせて拡張した高次元傾向スコアの基礎を理解する。
- 高次元傾向スコアの利点を理解する。
- 高次元傾向スコア分析を用いた先行研究について学ぶ。

2

目次

- 高次元傾向スコアとは何か
- 高次元傾向スコアの活用方法
- 高次元傾向スコアの先行研究
- 高次元傾向スコアの利点・限界
- 高次元傾向スコア研究のチェックリスト

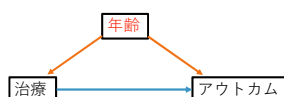
3

▼ 高次元傾向スコアとは？

4

交絡

治療法の選択とアウトカムの両方に影響を与える要因



例
・年齢若いほど治療A、高齢なほど治療Bを選択
・高齢なほどアウトカムが悪い
→ 治療Aの方が成績がよく見える

5

観察研究の限界

- レセプトデータ研究などの後ろ向き観察研究の場合、交絡因子が常に測定されて手に入るとは限らない
 - 「未測定 of 交絡因子 (unmeasured confounder)」の存在は後ろ向き観察研究の最大の課題
- 例 疾患の重症度、検査値データ、全身状態など

6

未測定交絡の調整

いままでは測定できる因子で未測定交絡因子の代理変数(proxy)とみなしてきた

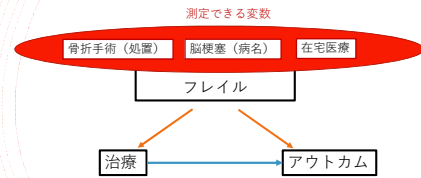
未測定交絡因子	測定できる因子
全身状態が悪い	酸素吸入
ヘルスリテラシーの度合い	定期的な健診受診
ショック状態	カテコラミンの使用



7

高次元傾向スコアのコンセプト

レセプトデータなどの大規模データベースを利用した研究で
たくさんある患者情報から未測定交絡を間接的に調整する



フレイルそのものはデータベースで測定できないが、フレイルに関連する変数は測定可能

8

傾向スコアからの拡張部分

- 傾向スコア推定モデルに投入する従属変数の選択方法
→ 今までは先行研究から重要だとされる変数を研究者が決定
- 高次元傾向スコアでは一定のアルゴリズムに従って変数選択が自動で行われる
- 多数の変数 (50~1000変数) を複数の次元から選択し、傾向スコア推定モデルに投入する

次元 = データの種類 (同じコーディングがされる同じ性質のもの)
病名、医薬品、診療行為、ラボデータなど

9

高次元傾向スコアのまとめ

- レセプトデータや電子カルテ等のリアルワールドデータを用いる研究で有用な方法
- 傾向スコアを拡張し、推定モデルに投入する変数を自動アルゴリズムで選択
- 膨大なデータを用いることで、未測定交絡を間接的に調整できる

10

▼ 高次元傾向スコアの活用

高次元傾向スコアの作成

- (1) 次元の決定
- (2) 変数コードの抽出
- (3) 変数コードの出現回数の評価
- (4) バイアス評価
- (5) 変数の決定
- (6) 傾向スコアの推定

- Schneeweiss S, et al. "High-Dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data." 2009. *Epidemiology* 20 (4): 512-22.
- Schneeweiss, S. "Automated Data-Adaptive Analytics for Electronic Healthcare Data to Study Causal Treatment Effects." 2018. *Clinical Epidemiology* 10: 771-88.

11

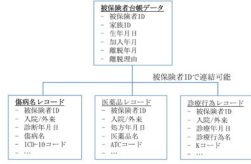
12

高次元傾向スコアの作成(1)

次元 (dimension) の決定

- 次元 = データの種類
(同じ構造、同じコーディング、同じ性質のもの)

傷病名、医薬品、診療行為
入院、外来
検査値データ NDB等レセプトデータ



6次元

- 次元の決め方に制限はない (遺伝子データ、バイオマーカー、フリーテキスト、なんでもよい)

13

高次元傾向スコアの作成(2)

変数コードの抽出

- 各次元ごとに変数コードの桁数を指定
傷病名 ICD-10コード 4桁 **E11.92**
薬剤 ATC分類 5桁 **C08CA01**
*桁数が大きいと出現割合が小さくなりすぎる

保険者ID	入院/外来	処方年月	医薬品名	ATCコード
A	外来	20150124	アムロジピン錠5mg	C08CA01
A	外来	20150124	オルメテックOD錠5mg	C09CA08
B	入院	20150130	ティーエスワシ配合顆粒	L01BC53
...

- 出現割合の確認
基準期間内に対象患者の何%で出現したか
基準期間→治療A開始前の12カ月間 (6カ月間)
- 各次元の上位n個の変数を候補として選択 n=100~200

6次元 × 200変数
= 1200変数

14

高次元傾向スコアの作成(3)

変数コードの出現回数の評価

- 患者ごとに各変数コードが何回出現したか数える (ICD10コード 110(高血圧)の発症数の例)

- 3つの二値変数を作成
- ① 1回以上出現
- ② 散発的出現 (例: 中央値以上)
- ③ 頻回出現 (例: 75%タイル値以上)

1200変数 × 3変数
= 3600変数

保険者ID	110の発症数	(1)	(2)	(3)
A	5	1	1	1
B	3	1	1	0
C	1	1	0	0
D	1	1	0	0
E	0	0	0	0
F	0	0	0	0
G	0	0	0	0
H	0	0	0	0
I	0	0	0	0
J	0	0	0	0

*発生している人中
②中央値は2、③75%タイル値は3.5

15

高次元傾向スコアの作成(4)

共変量のバイアス評価

各変数と治療及びアウトカムとの関連を用いて、バイアスの程度を評価

$$\text{Bias}_M = \frac{P_{c1}(\text{RR}_{cd}-1) + 1}{P_{c0}(\text{RR}_{cd}-1) + 1}$$

P_{c1}, P_{c0} : 変数の治療群と対照群における割合
 RR_{cd} : 変数とアウトカムについての未調整リスク比

16

高次元傾向スコアの作成(4)

$$\text{Bias}_M = \frac{P_{c1}(\text{RR}_{cd}-1) + 1}{P_{c0}(\text{RR}_{cd}-1) + 1}$$

P_{c1}, P_{c0} : 変数の治療群と対照群における割合
 RR_{cd} : 変数とアウトカムについての未調整リスク比

保険者ID	110_1	110_2	110_3	治療	死亡	変数名	P_{c1}	P_{c0}	RR_{cd}	Bias_M	$ \log(\text{Bias}_M) $
A	1	1	1	1	1	110_1	0.6	0.2	1.50	1.181	0.167
B	1	1	0	0	0	110_2	0.2	0.2	1.33	1	0
C	1	0	0	1	1	110_3	0.2	0	3.00	1.4	0.336
D	1	0	0	1	0						
E	0	0	0	1	1						
F	0	0	0	1	0						
G	0	0	0	0	1						
H	0	0	0	0	0						
I	0	0	0	0	0						
J	0	0	0	0	0						

この3変数でバイアスが高い順は 110_3 → 110_1 → 110_2

17

高次元傾向スコアの作成(5)

変数の決定

$|\log(\text{Bias}_M)|$ を高い順から k 個選択
→ k=50~1000 今回は 500個選択

最終的に傾向スコア推定モデルに投入する変数

- 既知の交絡因子を傾向スコア推定モデルに追加
→ 年齢、性別、人種、医療サービスの利用 (外来受診回数、入院回数など) は投入することを強く推奨
- その他に先行研究から交絡因子だと言われているものはモデルに投入する

3600変数
↓
500 + α

18

高次元傾向スコアの作成(6)

傾向スコアの推定

治療選択を従属変数としたロジスティック回帰により傾向スコアを推定
 →通常の傾向スコア分析と同様に傾向スコアの分布確認
 従来の傾向スコアを用いたアウトカム比較と同様に
 マッチング、逆確率重み付け分析等を行う

500+ α 変数
 ↓
 1

19

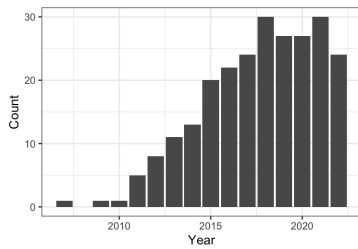
▼ 高次元傾向スコアを用いた先行研究

20

先行研究のレビュー

Pubmedで("high-dimensional propensity score" OR "high dimensional propensity score")で検索

検索結果：209編
 全体では2018年ごろから横ばい



21

先行研究で使われているデータソース例

Country	US	UK	Canada	Denmark	France	Japan
データのタイプ	診療報酬データ	電子カルテデータ	診療報酬データ	診療報酬とヘルスレジストリ	ヘルスケアデータベース	診療報酬データ
データベース例	MarketScan, Optum, Medicare	CPRD, THIN	RAMQ database, MED-ECHO database	Danish National patient register	French National Healthcare System	JMDC
設定した次元	入院病名 (ICD) 外来病名 (ICD)	入院病名 (THIN) 外来病名 (THIN)	退院データから入院病名 医師の診療報酬の病名コード	入院病名 (ICD) 外来病名 (ICD)	入院病名 (ICD) 慢性疾患レジストリデータ (ICD)	入院病名 (ICD) 外来病名 (ICD)
	入院処置 (ICD, CPTコード)	専門医への照会 (READ codes) 入院処置 (THIN)	退院データから処置コード 医師の診療報酬の処置コード		退院データから処置コード	入院処置 (日本のオリジナルコード)
	外来処置 (ICD, CPTコード)	外来処置 (THIN)	医師の診療報酬の処置コード 診療を行う医師の専門性		医師の診療報酬の処置コード 診療を行う医師の専門性	外来処置 (日本のオリジナルコード)
	外来の薬剤処方 (NDC)	処方 (BNFコード) 外来の薬剤使用 (THIN)	外来での薬剤処方	薬剤 (ATCコード)	外来の処方 (ATCコード)/入院中の高価薬剤 (out of DRG cost coding system)	入院・外来での処方 (日本のオリジナルコード, ATCコード)

(Rassen JA, et al. 2022. Pharmacoepidemiol Drug Saf.)

先行研究①：薬剤疫学

2型糖尿病患者にSU剤をセカンドラインで投与した場合の心臓血管疾患と低血糖のリスク

Douros A, et al. Sulfonylureas as second line drugs in type 2 diabetes and the risk of cardiovascular and hypoglycaemic events: population based cohort study. *BMJ* 2018 ;362:k2693.

Patient :	メトホルミンで治療開始した2型糖尿病患者
Exposure:	SU剤を追加・SU剤に変更
Control :	メトホルミン単剤継続
Outcome :	心筋梗塞・全死亡・重症低血糖

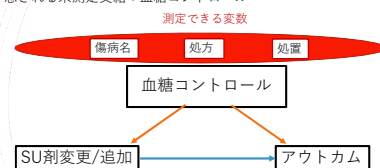
Data source: UK Clinical Practice Research Datalink (CPRD) + Hospital Episode Statistics (HES)

23

先行研究①：薬剤疫学

2型糖尿病患者にSU剤をセカンドラインで投与した場合の心臓血管疾患と低血糖のリスク

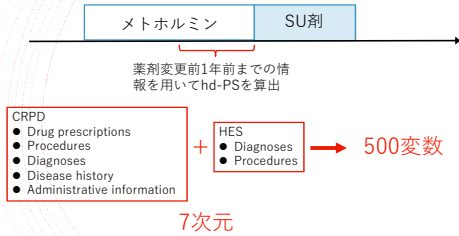
予想される未測定交絡：血糖コントロール



24

先行研究①：薬剤疫学

2型糖尿病患者にSU剤をセカンドラインで投与した場合の心臓血管疾患と低血糖のリスク



先行研究②：診療行為による臨床疫学

周術期口腔機能管理とがん患者の術後合併症との関連

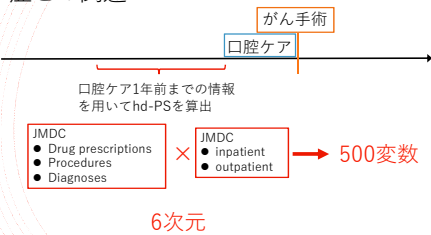
Ishimaru M, et al. Association between perioperative oral care and postoperative pneumonia after cancer resection. *Clin Oral Invest* 2019 ; 23:3581-3588

Patient :	主要ながん切除術施行患者
Exposure:	周術期口腔機能管理
Control :	周術期口腔機能管理なし
Outcome :	術後肺炎

Data source: JMDC database

先行研究②：診療行為による臨床疫学

周術期口腔機能管理とがん患者の術後合併症との関連



先行研究②：診療行為による臨床疫学

周術期口腔機能管理とがん患者の術後合併症との関連

	口腔ケアあり (%)	口腔ケアなし (%)	ASD
BMI			
<18.5	7.2	6.7	2.0
18.5-24.9	54.4	44.5	19.9
>=25.0	14.5	12.2	6.7
missing	23.1	36.5	27.9
喫煙	21.3	14.7	17.1
飲酒	41.1	32.1	18.6

hd-PS マッチング

	口腔ケアあり (%)	口腔ケアなし (%)	ASD
BMI			
<18.5	7.2	5.0	9.4
18.5-24.9	54.6	52.7	0.3
>=25.0	13.9	13.8	0.5
missing	24.3	26.5	5.1
喫煙	21.0	19.8	3.0
飲酒	40.4	40.3	0.4

*特定検診からのデータは欠測が多いためPS推定に用いなかった

推定に用いていない変数についてもバランスが取れていた

先行研究③：ヘルスサービスリサーチ研究

予定PCI手術のケアの質に関する研究

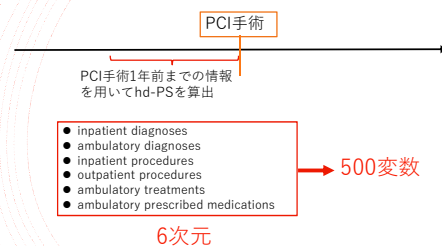
Enders D, et al. The Potential of High-Dimensional Propensity Scores in Health Services Research: An Exemplary Study on the Quality of Care for Elective Percutaneous Coronary Interventions. *Health Serv Res*. 2018

Patient :	待機的PCI(経皮的冠動脈インターベンション) 手術施行患者
Exposure:	外来で施行
Control :	入院下で施行
Outcome :	死亡

Data source:
German Pharmacoepidemiological Research Database

先行研究③：ヘルスサービスリサーチ研究

予定PCI手術のケアの質に関する研究



先行研究③：ヘルスサービスリサーチ研究 予定PCI手術のケアの質に関する研究

		HR
Adjusted analysis	■	0.58
PS matched	■	0.58
PS stabilized IPTW	■	0.45
hd-PS matched	■	1.18
hd-PS stabilized IPTW	■	1.20

先行研究やPSでは外来手術の方が健康な人が受けているという選択バイアスを調整しきれなかった
 外来手術の方が健康な人が受けているという選択バイアスを調整しきれなかった

31

高次元傾向スコアの 利点と限界

32

高次元傾向スコアの利点

- ◆ 未測定交絡による残差交絡を減少させる可能性がある
- ◆ レセプト・カルテデータなどに膨大にあるデータを無駄にしないで有効活用できる
- ◆ 研究者の変数選択時の恣意性が減少する
- ◆ 直感的には関係がない変数が重要な交絡因子である場合に調整できる

33

高次元傾向スコアの利点

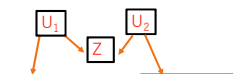
シミュレーション研究では傾向スコアと比較して

- ① 推定モデルに入っていない患者背景因子もバランスが取れていた (Guertin JR, et al. *BMC Med Res Methodol* 2016)
- ② 効果量がRCTのものに近似していた (Garbe E, et al. *Eur J Clin Pharmacol* 2013)

34

高次元傾向スコアの限界

- 独立な未測定交絡因子については調整できない (医療診療情報データベースには得られるデータの系統が偏っている。SES(社会経済的要因)が医療情報からproxyが得られるかは不明)
- 選択された変数が臨床的に解釈しづらいことがある
- 本来調整しない方が良い変数(操作変数、中間因子、M-bias、Z-biasなど)を調整している可能性がある (over-adjustment 過剰調整の可能性が示唆されるが、もしそれらをhdPSで調整してしまっても、影響は小さいだろうとは報告されている Schneeweiss, S. 2018. *Clinical Epidemiology* :771-88.)
- さまざまな高次元データを用いた傾向スコア推定法 (機械学習, Lasso回帰等)が提唱されており、今回の手法も含め、どの方法がより適しているのかについてはまだ明らかではない



35

高次元傾向スコア研究のチェックリスト

36

高次元傾向スコア研究のチェックリスト

重要な考慮事項を記載したチェックリストを提供

高次元傾向スコアに馴染みのない研究者や読者が高次元傾向スコアを用いた研究を理解し、解釈するための支援ツールとして利用できる

Rassen, JA, et al. 2022. "High-Dimensional Propensity Scores for Empirical Covariate Selection in Secondary Database Studies: Planning, Implementation, and Reporting." *Pharmacoepidemiology and Drug Safety*, November.

37

高次元傾向スコア研究のチェックリスト (発表者訳)

Guidance	Key concepts & considerations	Check
研究デザイン、プロトコル、統計解析プラン、hdPSの実行		
統計解析計画で共変量を特定・選択するためのパラメータを指定する。	●分析の前に、共変量をどのように同定し、ランク付けし、選択するかを事前に特定・決定する	●データソースごとの特徴を考慮する。
●次元の設定 ●変数と代理変数の同定 ●変数の優先順位付	研究者が事前に決めた変数の同定	研究プロトコルには事前に記載する。また、age, sex, race, health service utilization の変数は含めることを推奨
IV(操作変数)やコライダー(合流点)となる変数は除外する	PSの計算をするソフトウェア環境と、PSをどの手法で治療効果を推定するのかを記載	●ベースラインの患者特性や、絶対標準化平均差などのサマリー指標に許容できる閾値を「表1」として作成し、選択した変数の検査を行う ●PS分布図や変数追加によるプロットなどを出力する
計画された診断名ともしも異常が感知された時の手段を記載		

Rassen, JA, et al. 2022. *Pharmacoepidemiology and Drug Safety*.

38

高次元傾向スコア研究のチェックリスト (発表者訳)

Guidance	Key concepts & considerations	Check
報告および透明性		
交絡調整の成功とhdPSの性能を示す診断表とグラフを提示する	●PSの分布や標準化差をプロットしたり、優先される共変量が順次追加されていく様子を示すことは、hdPSの性能を示すのに有効な視覚化手法である ●2つの治療群間の患者のベースライン特性を示す「表1」を作成する。	
STaRT-RWEフレームワークの付録表3Fを完成させ、再現性と透明性を確保するために、共変量を同定・選択するために使用するパラメータを指定することを検討する。	●STaRT-RWEの構造化テンプレートは、研究手法の全体的な計画や報告を支援する。 ●STaRT-RWEの補足表3Fは、共変量定義のアルゴリズム、共変量評価期間、コードタイプ、診断位置を含む主要なパラメータを指定することを推奨する。 ●可能であれば、hdPS法の透明性を高めるために、変数の詳細なリストと解釈可能な説明を補足付録表で提供すること。	
*STaRT-RWE: Structured template and reporting tool for real world evidence (再現性のあるRWE研究をデザインし実施するためのガイドツール) Wang SV, et al. BMJ. 2021		

Rassen, Jeremy A, et al. 2022. *Pharmacoepidemiology and Drug Safety*.

39

まとめ

- 高次元傾向スコアは大規模データの性質に合わせて傾向スコア推定の変数選択を自動化した方法である。
- 高次元傾向スコアは未測定交絡を間接的に調整できる可能性がある。
- 高次元傾向スコアは主に薬剤の効果比較研究に用いられており、様々な国レセプトデータベースで活用されている

40

操作変数法

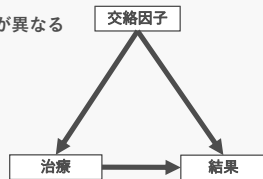
東京大学大学院医学系研究科イートロス医学
大野幸子

本日の内容

- 交絡
- 傾向スコア
- ランダム化比較試験
- 操作変数法
- 操作変数の例
- 操作変数法の仮定
- 単調性
- 操作変数法の推定の実際
- 操作変数法の実例
- 操作変数法checklist
- 操作変数法の限界

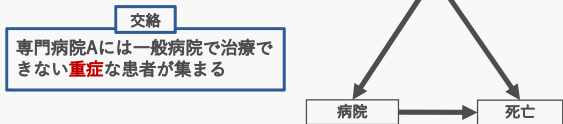
治療と結果の両方に影響

- 結果に影響を与える
- 曝露/治療の有無によって分布が異なる
- 曝露/治療から影響を受けない



専門病院Aを受診すると死にやすい？

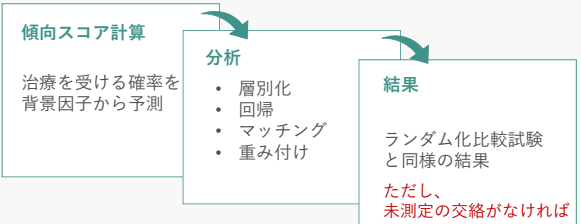
- 専門病院A 20%が死亡
- 一般病院B 10%が死亡



未測定の交絡因子が存在しないとき

- 層別解析
- 回帰モデル
- **傾向スコア**
- etc.

未測定の交絡因子が存在しないという仮定で、



因果推論いろいろ 7

未測定の変数が存在するとき

- ランダム化比較試験(RCT)
- 操作変数
- 差の差の分析
- 回帰不連続
- etc.

ランダム化比較試験 8

未測定の変数に影響を受けない

- くじ引きで治療を割り当てる
- 2群間で背景因子が揃っている

ランダム化比較試験 (割付を完全に遵守した場合) 9

アウトカムの差は治療の効果による

割付	治療群	対照群
治療を受けた割合	100%	0%
年齢	73	73
女性	21%	20%
重症度スコア	9	9
未測定の変数	?	?
アウトカムの割合	20%	50%

両群で揃う

治療の効果

ランダム化比較試験 10

いつでもRCTができるわけではない

- 費用 (高)
- 人数 (たくさん集めるのは大変)
- 倫理 (ランダムに割り付けられない治療も存在)

未測定の変数があるとき 11

RCTができないときの選択肢

- ランダム化比較試験(RCT)
- 操作変数法
- 差の差の分析
- 回帰不連続
- etc.

操作変数法 12

くじ引きの働きをする変数で未測定の変数に対処

- 操作変数がRCTのくじ引きの役割を果たす
- 操作変数は治療を確率的に割り付ける
- 未測定の変数因子も調整できる

操作変数法 13

くじ引きとなるような変数を使う

- RCTと同じイメージ図

操作変数 14

アウトカムの差は治療の割合の差に起因

操作変数の割付	治療されやすい群	治療されにくい群
治療を受けた割合	75%	25%
年齢	73	73
女性	21%	20%
重症度スコア	9	9
未測定 of 交絡	?	?
アウトカムの割合	20%	30%

両群で揃う (referring to age, gender, and severity score)

治療の効果 (referring to the difference in outcome rates)

操作変数法 15

治療の受けやすさに影響するくじ引き変数の例

- ある治療を積極的に行う病院と消極的に行う病院までの距離の差 (Differential distance)
- 医師/病院の治療の好み
- 地域の治療実施割合
- 曜日、日付
- 遺伝的変異 (Mendelian randomization)

操作変数法 16

Differential distance

病院が備える特性を見るための操作変数

- (自宅から専門病院までの距離(d1)) - (自宅から一般病院までの距離(d2))

d1-d2 小 → 専門治療を受けやすい
d1-d2 大 → 専門治療を受けにくい

操作変数法 17

曜日 (Friday admission)

早期介入の効果を見るための操作変数

- 金曜：土日を含むため、3日目以降に手術/検査が遅れやすい
- 翌日か翌々日が平日なので2日以内に手術/検査を受けやすい

月 → 火 → 水 → 木 → 金 → 土 → 日 → 月

操作変数の例 JAMA 2015;314:1272-9 18

高齢者の肺炎はICUで管理するべきか？

- P：肺炎を発症した高齢者
- E：ICU入室
- C：一般病室
- O：30日死亡

高齢者の肺炎はICUで管理すべきか？

- P：肺炎を発症した高齢者
- E：ICU入室
- C：一般病室
- O：30日死亡

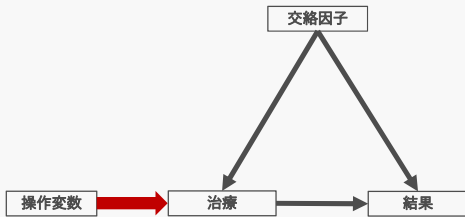
● **操作変数：differential distance**
(自宅からICU入室させる病院とそうでない病院までの距離の差)



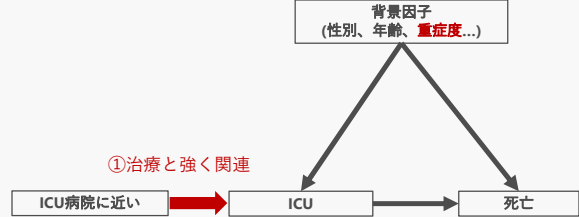
基本の仮定は3つ

- 仮定①
治療と強く関連している
- 仮定②
結果と直接関連しない
- 仮定③
結果との共通原因を持たない

①治療と強く関連している



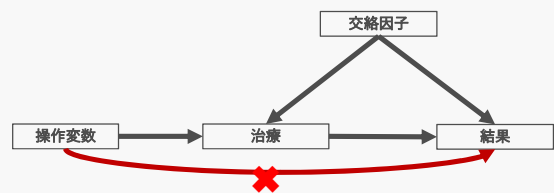
例

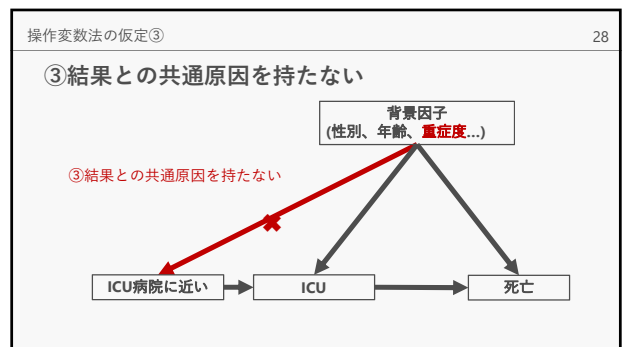
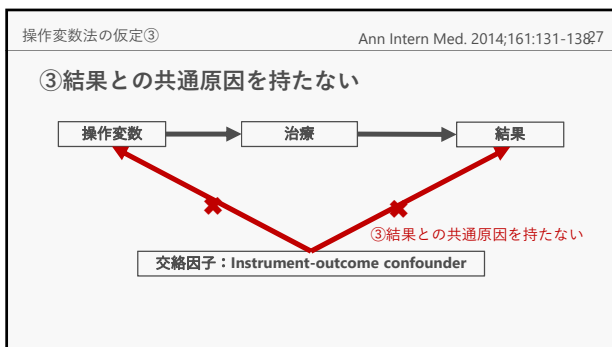
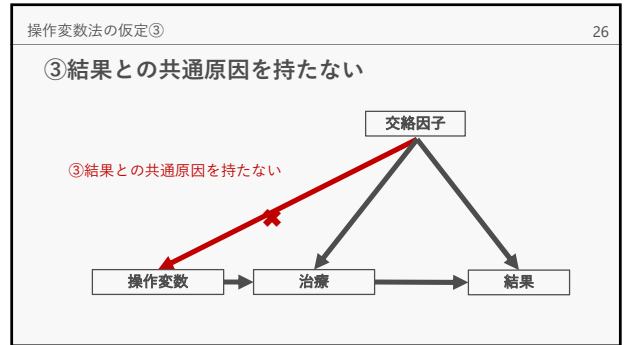
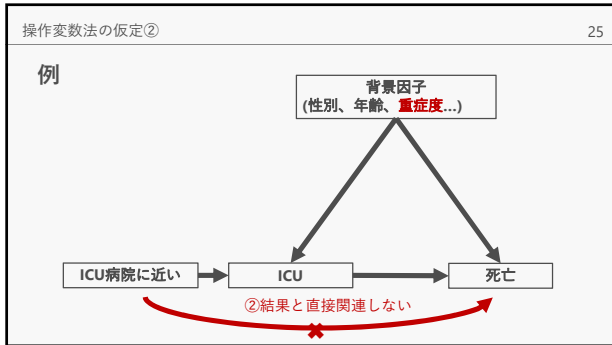


①治療と強く関連している

- くじ引きの役割「操作変数が1なら治療を受ける」(100%ではない)
- 治療を操作変数で回帰した際のF値で評価 (F > 10が目安)
- F値 ≤ 10: 弱い操作変数(weak instrument)で誤った結果を導く

②結果と直接関連しない(治療を介してのみ関連あり)





- 操作変数法の仮定②③ 29
- ②結果と直接関連しない
- ③結果との共通原因を持たない
- ②③は証明できない
 - 臨床的にメカニズムを判断
 - ③はIVのカテゴリごとに測定されている背景因子を比較し傍証とすることも

操作変数法の仮定 30

仮定を検証できるのは①のみ

仮定① 治療と強く関連している	仮定② 結果と直接関連しない
仮定③ 結果との共通原因を持たない	

操作変数の仮定 31

仮定を満たさない効果変数は推定にバイアス

効果推定量

操作変数の②③の仮定を満たさないバイアス

$$\frac{(\text{操作変数1の時のアウトカム}) - (\text{操作変数0の時のアウトカム})}{(\text{操作変数1の時の治療}) - (\text{操作変数0の時の治療})}$$

仮定①を満たさず、弱い操作変数でゼロに近づく

操作変数の仮定 32

追加の仮定は1つ

仮定① 治療と強く関連している	仮定② 結果と直接関連しない
仮定③ 結果との共通原因を持たない	追加の仮定④ 単調性

操作変数法の仮定④ 33

④単調性

ICU病院近い	ICU病院遠い	タイプ
一般病室	一般病室	Never-taker
ICU	一般病室	Complier
一般病室	ICU	Defier X
ICU	ICU	Always-taker

割り当てと逆の治療を受ける人は存在しないと仮定

単調性が必要な理由 34

Z: 操作変数 (治療群に割付けられる場合を1, それ以外を0)
 A: 受ける治療 (治療を受ける場合は1, それ以外を0)
 Y: アウトカム
 X: 交絡

単調性が必要な理由 35

$A^{Z=1} = A^1$ Z=1の場合に受ける治療
 $A^{Z=0} = A^0$ Z=0の場合に受ける治療

単調性が必要な理由 36

4つのタイプ

A^0	A^1	タイプ
0	0	Never-taker
0	1	Complier
1	0	Defier
1	1	Always-taker

単調性の仮定が必要な理由 37

Target Inference
Complier average causal effect (CACE)
 Complier average treatment effect (CATE)
 Local average treatment effect (LATE)

$$E(Y^{Z=1}|A^0=0, A^1=1) - E(Y^{Z=0}|A^0=0, A^1=1)$$

$$= E(Y^{Z=1} - Y^{Z=0} | \text{compliers})$$

$$= E(Y^{a=1} - Y^{a=0} | \text{compliers})$$

操作変数法で得られる結果はcomplierの中での効果
 操作変数が治療の割り当てに影響しないalways-takerとnever-takerでの効果は不明
 defierがいてはなぜいけないのか？

単調性の仮定が必要な理由 38

実際に観察されるデータからはタイプを決定できない

Z	A	A ⁰	A ¹	タイプ
0	0	0	?	Complier Never-taker
0	1	1	?	Always-taker Defier
1	0	?	0	Never-taker Defier
1	1	?	1	Always-taker Complier

単調性の仮定が必要な理由 39

割り当てと逆の治療を受ける人は存在しないと仮定

A ⁰	A ¹	タイプ
0	0	Never-taker
0	1	Complier
1	0	Defier X
1	1	Always-taker

単調性の仮定が必要な理由 40

Defierがいないと仮定すると、

Z	A	A ⁰	A ¹	タイプ
0	0	0	?	Complier Never-taker
0	1	1	1	Always-taker Defier
1	0	0	0	Never-taker Defier
1	1	?	1	Always-taker Complier

単調性の仮定が必要な理由 41

Target Inference
Complier average causal effect (CACE)
 $E(Y^{a=1} - Y^{a=0} | \text{compliers})$

まずは観察データで計算できるものから、

$$E(Y^{Z=1} - Y^{Z=0}) = E(Y|Z=1) - E(Y|Z=0)$$

単調性の仮定が必要な理由 42

単調性を仮定した場合

$$E(Y^{Z=1} - Y^{Z=0}) = E(Y|Z=1) - E(Y|Z=0)$$

ここで

$$E(Y|Z=1) = E(Y|Z=1, \text{always-taker})P(\text{always-taker})$$

$$+ E(Y|Z=1, \text{never-taker})P(\text{never-taker})$$

$$+ E(Y|Z=1, \text{complier})P(\text{complier})$$

$$E(Y|Z=0) = E(Y|Z=0, \text{always-taker})P(\text{always-taker})$$

$$+ E(Y|Z=0, \text{never-taker})P(\text{never-taker})$$

$$+ E(Y|Z=0, \text{complier})P(\text{complier})$$

単調性の仮定が必要な理由 43

never-takerとalways-takerはZに影響を受けない

$$E(Y^{Z=1} - Y^{Z=0}) = E(Y|Z=1) - E(Y|Z=0)$$

ここで

$$E(Y|Z=1) = E(Y|Z=1, \text{always-taker})P(\text{always-taker}) + E(Y|Z=1, \text{never-taker})P(\text{never-taker}) + E(Y|Z=1, \text{complier})P(\text{complier})$$

$$E(Y|Z=0) = E(Y|Z=0, \text{always-taker})P(\text{always-taker}) + E(Y|Z=0, \text{never-taker})P(\text{never-taker}) + E(Y|Z=0, \text{complier})P(\text{complier})$$

単調性の仮定が必要な理由 44

$$E(Y|Z=1) - E(Y|Z=0) = E(Y|Z=1, \text{complier})P(\text{complier}) - E(Y|Z=0, \text{complier})P(\text{complier})$$

両辺をP(complier)で割って

$$\frac{E(Y|Z=1) - E(Y|Z=0)}{P(\text{complier})} = E(Y|Z=1, \text{complier}) - E(Y|Z=0, \text{complier}) = E(Y^{a=1} | \text{complier}) - E(Y^{a=0} | \text{complier}) = \text{CACE}$$

単調性の仮定が必要な理由 45

Target Inference
Complier average causal effect (CACE)

$$\text{CACE} = \frac{E(Y|Z=1) - E(Y|Z=0)}{P(\text{complier})}$$

$$P(\text{complier}) = E(A|Z=1) - E(A|Z=0)$$

always-takerまたはcomplierの割合

always-takerの割合

単調性の仮定が必要な理由 46

単調性の仮定により観察データからCACEを得ることができる

$$\text{CACE} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(A|Z=1) - E(A|Z=0)}$$

操作変数の仮定 47

仮定を検証できるのは①のみ

仮定① 治療と強く関連している	仮定② 結果と直接関連しない
仮定③ 結果との共通原因を持たない	追加の仮定④ 単調性

推定の実際 48

操作変数法ののいろいろな推定方法

- Wald推定量
- 2-stage least squares (TSLS)
- 2-stage residual inclusion (TSRI)
- 2-stage prediction substitution (TSPS)
- etc.

Wald推定量

$$\frac{E(Y|Z = 1) - E(Y|Z = 0)}{E(A|Z = 1) - E(A|Z = 0)}$$

$$\frac{(\text{操作変数1の時のアウトカム}) - (\text{操作変数0の時のアウトカム})}{(\text{操作変数1の時の治療}) - (\text{操作変数0の時の治療})}$$

2SLS

- 線形モデルを仮定
- アウトカムが連続変数の時に適用されやすい
- パッケージが存在

2SLS

- ① 治療の変数を操作変数と背景因子で線形回帰する (F値確認)

$$A = \alpha_0 + \alpha_1 Z + \alpha_2 C + \varepsilon_1$$
 - ② ①で得られた係数をもとに治療の予測値を求める

$$\hat{A} = \hat{\alpha}_0 + \hat{\alpha}_1 Z + \hat{\alpha}_2 C$$
 - ③ アウトカムを②で求めた予測値と背景因子で線形回帰する

$$Y = \beta_0 + \beta_1 \hat{A} + \beta_2 C + \varepsilon_2$$
- 分散は修正が必要(Palmer et al., AJE 2017)

2SRI

- 非線形モデルを仮定
- アウトカムが2値変数に適用されることが多い
- 2回一般化線形モデルを使う

2SRI

- ① 治療の変数を操作変数と背景因子で回帰する

$$A = \alpha_0 + \alpha_1 Z + \alpha_2 C + \varepsilon_1$$
 - ② 実際の治療の値と①から計算できる予測値の差(残差)を計算

$$\varepsilon_1 = A - \hat{A}$$
 - ③ アウトカムを②で求めた残差と治療と背景因子で回帰する

$$Y = \beta_0 + \beta_1 A + \beta_2 C + \varepsilon_2$$
- 分散は修正が必要(Palmer et al., AJE 2017)

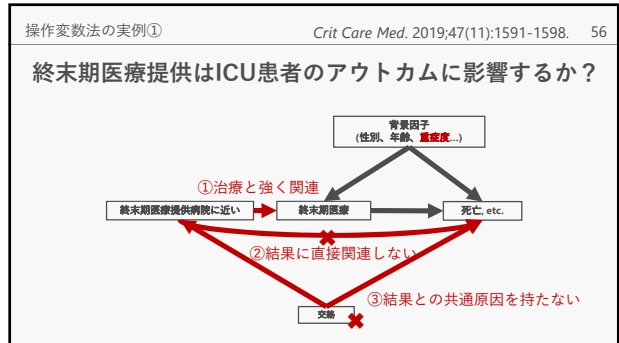
複数の操作変数を投入する場合

- 1つの治療効果に対して2つ以上操作変数を投入できる
- F値が高くなる可能性があり、弱い操作変数でなくなる
- 過剰識別制約検定が必要
「全ての操作変数が内生変数ではない」ことを検定
(1つでも操作変数が内生変数でなければよい)
- 内生変数：説明変数だが、別な説明変数で説明されてしまう変数
- 操作変数が内生変数だと条件を満たさないため不適當

操作変数法の実例① Crit Care Med. 2019;47(11):1591-1598. 55

終末期医療提供はICU患者のアウトカムに影響するか？

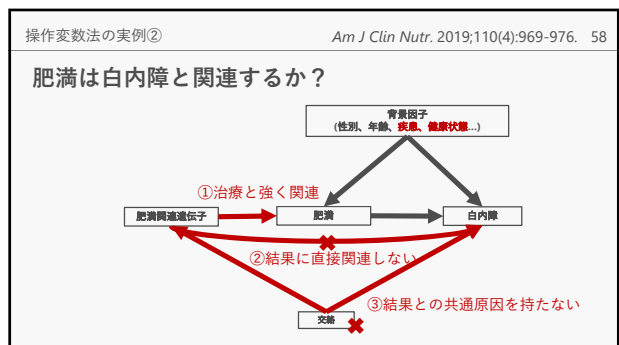
- P：ICU患者
- E：終末期医療提供病院
- C：通常医療提供病院
- O：院内死亡/ホスピス退院/
 - 増加すると予測
- 操作変数：Differential distance
 - 在院日数/治療強度
 - 減少すると予測



操作変数法の実例② Am J Clin Nutr. 2019;110(4):969-976. 57

肥満は白内障と関連するか？

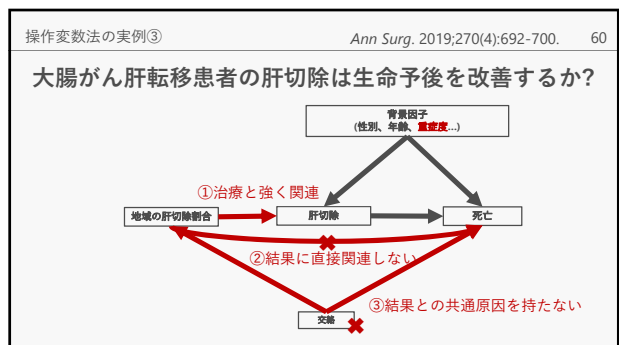
- P：49歳以上の成人コホート
- E：肥満
- C：肥満でない
- O：白内障
- 操作変数：肥満関連遺伝子 (Mendelian randomization)



操作変数法の実例③ Ann Surg. 2019;270(4):692-700. 59

大腸がん肝転移患者の肝切除は生命予後を改善するか？

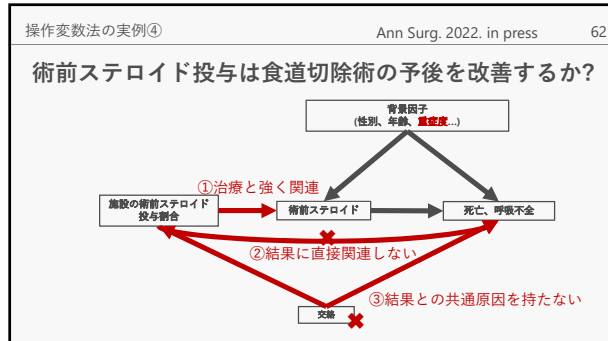
- P：大腸がん肝転移患者
- E：肝切除
- C：非切除
- O：1, 2, 3, 4, 5年生存
- 操作変数：地域の肝切除施術割合



操作変数法の実例④ Ann Surg. 2022. in press 61

術前ステロイド投与は食道切除術の予後を改善するか？

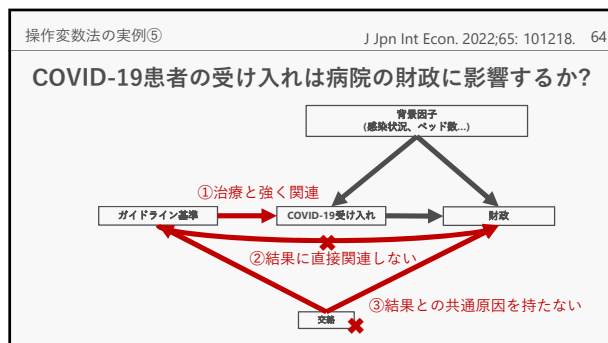
- P：食道がん患者
- E：術前ステロイド投与
- C：術前ステロイド投与なし
- O：入院死亡、呼吸不全
- 操作変数：施設の術前ステロイド投与割合



操作変数法の実例⑤ J Jpn Int Econ. 2022;65: 101218. 63

COVID-19患者の受け入れは病院の財政に影響するか？

- P：東京のCOVID-19患者受け入れ可能な病院
- E：COVID-19患者受け入れ
- C：COVID-19患者受け入れなし
- O：1ベッドあたりの利益/月
- 操作変数：COVID-19受け入れに関するガイドライン（病床あたりの呼吸器専門医数、病床あたり個室数）



操作変数checklist Epidemiology 2013;24: 363–369 65

論文を読むとき、書くときの注意点

- 推定する対象集団への効果であること（CACE）や仮定を明記
- 操作変数と治療の関連をF値を使って明記
- 測定可能な交絡因子と操作変数や治療の関係を記載
- 2値のアウトカム、治療、操作変数の組合せを明記
- 分散は修正が必要

操作変数の限界 66

実は限界が多い

- 適切な操作変数を見つけることが難しい
- ①操作変数と治療に強い関連がないと、誤った結果を導く
- ②操作変数と結果に関連がないことを確実に証明できない
- ③操作変数と結果が共通の原因を持たないことを確実に証明できない
- 仮定が多い（操作変数の条件①–③、④単調性）

バイアス増幅の危険性

CACE

$$\frac{E(Y|Z = 1) - E(Y|Z = 0)}{E(A|Z = 1) - E(A|Z = 0)}$$

操作変数の仮定②③を満たさないとバイアス

仮定①弱い操作変数でゼロに近づく

実は限界が多い

- 重要な交絡が欠落しているときに操作変数法単独使用を検討
- 操作変数法単独の研究は多くはない（4割程度）
- 傾向スコア分析などの方法と併用されることが多い
- 傾向スコア分析と操作変数法で、結果が異なるということも
- 交絡因子が全て測定されている仮定 VS. 操作変数の仮定

本日の内容

- 交絡
- 傾向スコア
- ランダム化比較試験
- 操作変数法
- 操作変数の例
- 操作変数法の仮定
- 単調性
- 操作変数法の推定の実際
- 操作変数法の実例
- 操作変数法checklist
- 操作変数法の限界

- 操作変数法は未測定の交絡因子を調整できる
- 適切な操作変数を見つけることは難しい
- 仮定が厳しい
- 仮定を満たさない、弱い操作変数でバイアスを増幅
- 単独で用いることは少ない

不連続回帰デザイン・ 差の差分析

自治医科大学データサイエンスセンター
山名隼人

Agenda

テーマ

- 不連続回帰デザイン (regression discontinuity)
- 差の差分析 (difference in differences)

内容

- 研究デザインの背景
- 研究デザインの基礎・注意点
- どのような臨床疑問の検証に適しているか
- 活用例

Agenda

テーマ

- 不連続回帰デザイン (regression discontinuity)
- 差の差分析 (difference in differences)

内容

- 研究デザインの背景
- 研究デザインの基礎・注意点
- どのような臨床疑問の検証に適しているか
- 活用例

臨床研究の種類

- 介入研究
 - 実験的な環境
 - 2群間の比較を行う場合、背景を揃えることが可能
 - 例：ランダム化比較試験
- 観察研究
 - 実際に行われた診療を“見るだけ”
 - 2群間の比較を行う場合、“結果的に”治療を受けた2群を比較する
 - 例：保健医療データベースを用いた研究

臨床研究の種類

例：「降圧薬の使用有無による脳卒中予防効果」

- ランダム化比較試験の場合
 - 基礎疾患、血圧などが揃った2群で比較が可能
- 観察研究の場合
 - 患者背景は2群で異なるのが当然
 - ベースラインの血圧は降圧薬あり群で高い

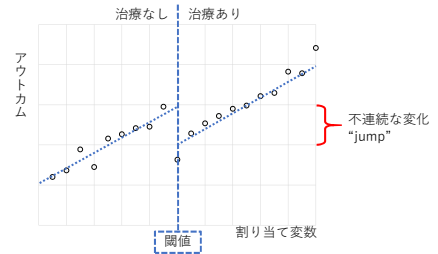
連続値と閾値

- 臨床的判断は不連続
 - 特定の閾値を境に治療方針が変わることがある
 - 例：収縮期血圧 ≥ 140 mmHgの場合に降圧薬が開始
- 血圧は連続値
- 血圧が1mmHg上がっても脳卒中のリスクはほぼ不変
- 測定時の1mmHgの誤差は偶然による

Regression discontinuity

- 治療の判断に使われる**割り当て変数**（連続）に**閾値**がある場合に適用
- 閾値の近傍で、閾値以上 vs 閾値未満でアウトカムを比較する
- ランダム割り付けに近い状況を作り出す
- 例
 収縮期血圧140mmHgで降圧薬が開始された群
 vs
 収縮期血圧139mmHgで降圧薬が開始されなかった群

Regression discontinuity

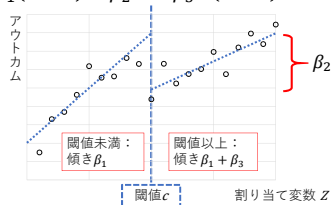


Regression discontinuity

回帰式を用いる場合

$$Outcome = \beta_0 + \beta_1(Z - c) + \beta_2D + \beta_3D(Z - c)$$

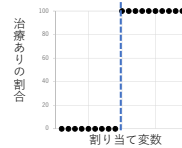
- Z: 割り当て変数
- c: 閾値
- D: $Z \geq c$ なら1
 $Z < c$ なら0



Sharp/fuzzy regression discontinuity

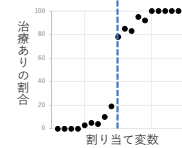
Sharp RD

閾値により割り当てが「全か無か」で決まる場合



Fuzzy RD

閾値前後で割り当てが不連続に変わるが、完全に2分されるわけではない場合



Fuzzy RDの場合

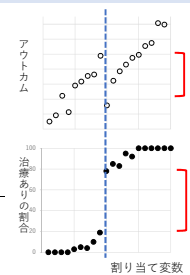
- 閾値を超えたかどうかは、
 - 治療と強く関連
 - 結果と直接は関連しない
 - 結果との交絡が存在しない
 操作変数と同じ
- 対象者は下記に分類される：
 - 閾値を超えたかに関わらず、治療を受けたであろう人 (always-taker)
 - 閾値を超えたかに関わらず、治療を受けなかったであろう人 (never-taker)
 - 閾値をまたいだことで、治療を受けるようになった人 (complier)
 - 閾値をまたいだことで、治療を受けなくなった人 (defier)

Fuzzy RDの場合

操作変数として考える

- アウトカムのjumpはintention to treat (ITT)に相当
- 単調性の仮定のもとで、Wald推定量を用いてcomplier average causal effect (CACE)を計算できる

$$CACE = \frac{\text{閾値前後のアウトカムのjump}}{\text{閾値前後の治療割合のjump}}$$



不連続回帰デザインの活用法 (1)

検査値等と治療

- 閾値前後で治療（介入）が変化する場合に、介入の効果を検証する

例

- 介入： 降圧薬 輸血 手術/保存的
- 割り当て変数： 血圧 Hb 年齢

不連続回帰デザインの活用法 (2)

時期による治療方針の変化

- 制度変更やガイドラインの更新などにより治療方針（介入）が急に変化した場合

例

- 介入： 治療方針・ワクチン接種
- 割り当て変数： 時期・年度

不連続回帰デザインの活用法 (3)

医療政策の評価

- 政策の対象者が限定され、一定の基準がある場合に、政策の効果を検証する

例

- 介入： 保険・助成
- 割り当て変数： 収入・年齢・地域

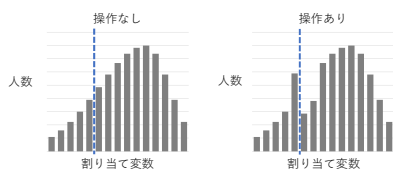
不連続回帰デザインの前提 (1)

1. 割り当てのルールと閾値が明確である
 - 検査値等が一定の値を超える（or下回る）と介入
 - 個人によって閾値が異なる場合は×
2. 介入以外の要因は閾値付近で連続である
 - 他の要因も同時にjumpをしてはいけない
 - 血圧の治療と同時に脂質・血糖値の治療も開始される場合は×

不連続回帰デザインの前提 (2)

3. 対象者が介入前に割り当て変数を操作できない (manipulation)

- メタボ健診の腹囲測定は×



不連続回帰デザインの利点と限界

利点

- ランダム化比較試験に準じて因果関係が推定できる
- リアルワールドにおける効果を検証できる

限界

- 閾値付近の結果を閾値から遠い対象者にあてはめることには限界がある (外的妥当性)
- 閾値付近の人数は少ないため、比較的大きなN数が必要

不連続回帰デザインの例 (1)

HIV患者に対する抗レトロウイルス療法開始

- 対象者 南アフリカ HIV陽性患者
- 介入 抗レトロウイルス療法開始
- 割り当て変数 初診時CD4 count
- 閾値 200 cells/ μ L
- アウトカム 死亡

Bor et al. Epidemiology 2014;25:729-37.

不連続回帰デザインの例 (2)

HPVワクチン接種と性行動

- 対象者 カナダ オンタリオ州の女性
- 介入 HPVワクチン接種の政策
- 割り当て変数 中学2年生になった年
- 閾値 2008年
- アウトカム 妊娠、STD

Smith et al. CMAJ 2015;187:E74-81.

不連続回帰デザインの例 (3)

生活保護費の支給額と医療費

- 対象者 日本 生活保護世帯
- 介入 生活保護費支給額
- 割り当て変数 第1子の年齢
- 閾値 3歳 (児童養育加算が減額)
- アウトカム 世帯の医療費

Nishioka et al. J Epidemiol Community Health 2022;76:505-11.

不連続回帰デザイン：まとめ

- 割り当てに関わる連続変数に閾値が存在する場合に、閾値近傍でのアウトカムの不連続な変化を比較する手法
- 検査値により治療が変わる場合や、時期により治療方針が変化した場合などの検証に適している
- 前提条件
 - 割り当てのルールと閾値が明確である
 - 介入以外の要因は閾値付近で連続である
 - 対象者が介入前に割り当て変数を操作できない

Agenda

テーマ

- 不連続回帰デザイン (regression discontinuity)
- 差の差分析 (difference in differences)

内容

- 研究デザインの背景
- 研究デザインの基礎・注意点
- どのような臨床疑問の検証に適しているか
- 活用例

差の差分析：背景

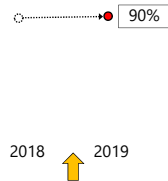
- 経時変化が関係するclinical questionの例
 - ガイドライン改訂により患者のアウトカムは改善したか？
 - 医療の質を改善するプログラムを導入した病院で、臨床評価指標 (quality indicator) は改善したか？
 - 新しい保健医療政策の導入により格差は是正されたか？
- 想定されるデザイン：前後比較
- 課題
 - 自然な経時変化の効果を見ているだけではないのか
 - 対照群がない

経時変化を評価する

例：医療の質を改善する“介入”の効果を検証する

•方法1：

導入後の指標を見る
⇒ 介入前より改善?悪化?

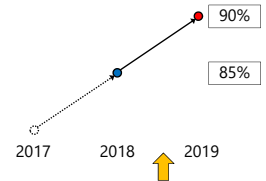


経時変化を評価する

例：医療の質を改善する“介入”の効果を検証する

•方法2：

導入前後の指標を比較
⇒ 介入がなくても自然と改善しているのでは?

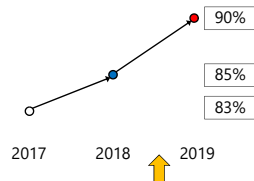


経時変化を評価する

例：医療の質を改善する“介入”の効果を検証する

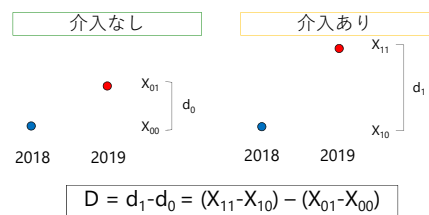
•方法3：

導入前のトレンドを考慮
⇒ 介入がなかった対照群
ではどうだったか?



差の差 (difference in differences)

介入の影響を受けない対照群を見つけ、経時変化を比較



差の差分析の実際

交互作用を導入した回帰モデルを利用

アウトカム
= $\alpha + \beta_1 \times (\text{介入}) + \beta_2 \times (\text{時点}) + \beta_3 \times (\text{介入}) \times (\text{時点})$

介入：介入群=1、対照群=0
時点：介入後=1、介入前=0
介入×時点：交互作用項

差の差分析の実際

交互作用を導入した回帰モデルを利用

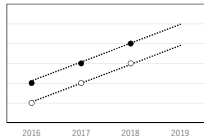
アウトカム
= $\alpha + \beta_1 \times (\text{介入}) + \beta_2 \times (\text{時点}) + \beta_3 \times (\text{介入}) \times (\text{時点})$

	介入前(時点=0)	介入後(時点=1)	後-前	差の差 β_3
介入群 (介入=1)	$\alpha + \beta_1$	$\alpha + \beta_1 + \beta_2 + \beta_3$	$\beta_2 + \beta_3$	
対照群 (介入=0)	α	$\alpha + \beta_2$	β_2	

差の差分分析の仮定 (1)

仮定 1. 平行トレンド (parallel trends)

- 介入群において介入がなかったと仮定した場合、アウトカムの経時変化は対照群と同じ
- 検証方法
 - 介入前の両群のトレンドを図示
 - 介入前で、時期×群の交互作用がないことを確認



差の差分分析の仮定 (2)

仮定 2. 共通ショック (common shocks)

- 介入と無関係にアウトカムに影響する外的要因(shock)が発生した場合、影響の程度が介入群と対照群で同じ
- 例：
 - 効果を検証したい政策と同時に施行された、全く別の制度変更
 - 天災、新興感染症
- 検証方法：なし

差の差分分析の活用法

研究テーマ

- 医療制度改革の効果
- 院内での取り組みの効果 (前後比較と同様)
- ガイドライン改訂の影響

注意点

- 適切な対照群を選択する (2つの仮定を満たすか)
- 介入前から、複数時点のデータが必要
- 各群・時点で同等のデータ (データベース研究に適している)

差の差分分析の例 (1)

医師の働き方改革と患者アウトカム

- データ： 米国Medicareデータ
- 介入群： Teaching hospitals (レジデントが多い病院)
- 対照群： Non-teaching hospitals (レジデントが少ない病院)
- 時期： 2011年のACGME duty hour reforms前後
- アウトカム： 入院患者の30日以内死亡率・再入院率

ACGME = Accreditation Council for Graduate Medical Education
Patel et al. JAMA 2014;312:2364-73.

差の差分分析の例 (2)

東日本大震災が「避けられる入院」に与えた影響

- データ： 厚生労働科学研究DPC研究班データベース
- 介入群： 岩手県・宮城県・福島県の病院
- 対照群： それ以外の病院
- 時期： 2010/7~2011/2 vs 2012/7~2013/2
- アウトカム： 各種ambulatory care sensitive conditionによる入院

Sasabuchi et al. J Epidemiol Community Health 2017;71:248-52.

差の差分分析の例 (3)

子ども医療費助成と歯科受診・口腔健康状態

- データ： 熊本県 国民健康保険
- 介入群： 自治体の子ども医療費助成が終了
- 対照群： // 継続
- 時期： 2015年3月の前後
- アウトカム： 歯科受診 (初回・全受診)、重症歯科疾患の発生

Ono et al. Community Dent Oral Epidemiol 2022 (epub).

差の差分分析：まとめ

- 介入群における経時変化と対照群における経時変化の差をとることで、介入の影響を検証する手法
- 分析には、介入×時点の交互作用を加えた回帰モデルを用いる
- 適切な対照群を選択する必要がある
 - 平行トレンドの仮定： 介入群において介入がなかった場合アウトカムの経時変化は対照群と同じ
 - 共通ショックの仮定： 外的要因が発生した場合、影響の程度は介入群と対照群で同じ

時間依存性交絡と周辺構造モデル

東京大学大学院 医学系研究科 臨床疫学・経済学
大邊寛幸

目次

1. 時間依存性交絡について
2. 時間依存性交絡の対処法と周辺構造モデル
3. 周辺構造モデルの逆確率重み付け法に必要な仮定
4. 周辺構造モデルの逆確率重み付け法の手順

目次

1. 時間依存性交絡について
2. 時間依存性交絡の対処法と周辺構造モデル
3. 周辺構造モデルの逆確率重み付け法に必要な仮定
4. 周辺構造モデルの逆確率重み付け法の手順

周辺構造モデルの逆確率重み付け法を用いた最初の論文

HIV陽性の男性に対するZidovudineは死亡を減らすか？

Design 観察研究
Patient HIV陽性男性で、AIDSを発症しておらず、治療薬も投与されていない
Exposure Zidovudineの投与
Control Zidovudineの非投与
Outcome 死亡

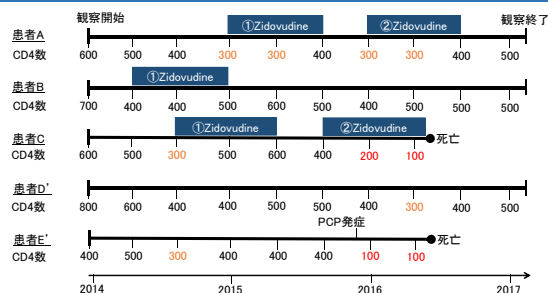
Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

HIV陽性の男性に対するZidovudineは死亡を減らすか？

1. AIDS発症していない場合
 - (1)CD4陽性Tリンパ球数が**350/uL**より多い
→治療を行わず経過観察する
 - (2)CD4陽性Tリンパ球数が**350/uL**以下
→治療を開始する
2. AIDS発症している場合
→治療を開始するが、カリニ肺炎など重篤な場合はその治療を優先する場合がある

抗HIV治療ガイドラインより

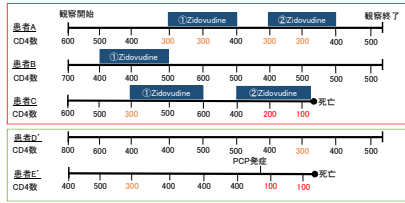
HIV陽性の男性に対するZidovudineは死亡を減らすか？



粗解析

Crude

アウトカム: 死亡
 治療群: 患者A,B,C
 コントロール群: 患者D',E'
 共変量: なし



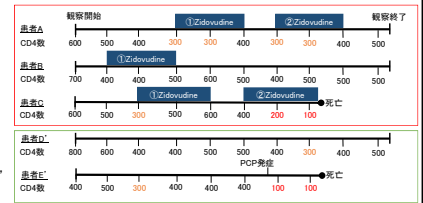
Zidovudineは死亡率をハザード比3.6倍上昇させる

Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

標準的な統計解析方法

Cox比例ハザードモデル

アウトカム: 死亡
 治療群: 患者A,B,C
 コントロール群: 患者D',E'
 共変量:
 観察開始時点での年齢,性別,
 既往症,CD4数など



Zidovudineは死亡率をハザード比2.3倍上昇させる

Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

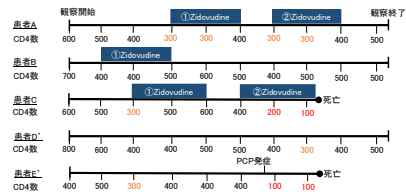
標準的な統計解析ではバイアスを生じる

時間依存性交絡(=CD4数)が存在するため

Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

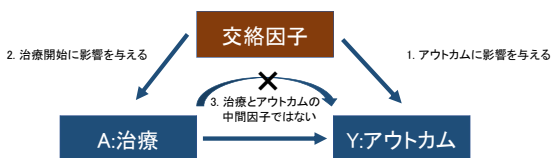
時間依存性変数(Time varying variables)

観察期間中に何度も測定し、時間に依存して変化していく変数



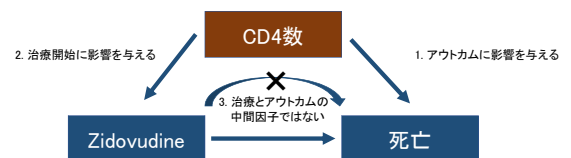
交絡因子(Confounder)

興味あるアウトカム(例えば死亡)のリスク因子(予測因子)かつ、
 治療開始に影響する(予測する)変数で中間因子でない



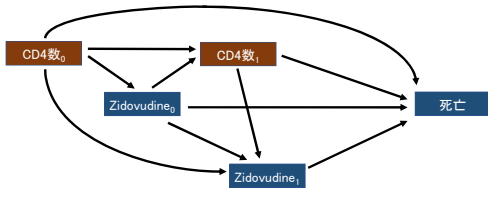
交絡因子(Confounder)

CD4数はZidovudineの開始を予測し、
 かつ死亡のリスク因子であるので交絡因子である



時間依存性交絡(Time dependent confounder)

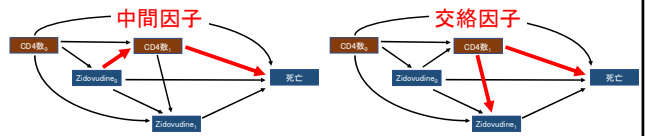
時間依存性交絡 = 時間依存性変数 + 交絡因子



時間依存性交絡の問題①過調整バイアス

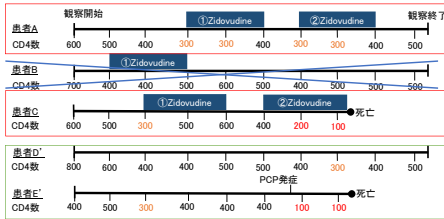
治療によって変化する中間結果 (CD4数₁) は

- 中間因子なので
→ 調整すると治療効果の推定にバイアス (過調整バイアス)
- 交絡因子なので
→ 調整しないと治療効果の推定にバイアス



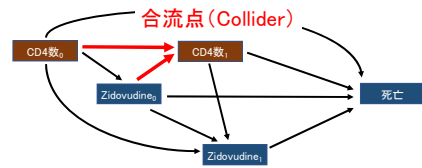
時間依存性交絡の問題②選択バイアス

CD4数が300以下の患者のみ含める



患者Bのように初回 Zidovudine 効果を見逃す

時間依存性交絡の問題②選択バイアス



*合流点層別バイアス (Collider stratification bias) とも呼ぶ

時間依存性交絡の問題まとめ

①過調整バイアスや②選択バイアスのため
通常の層別解析・回帰モデル・傾向スコア法では
バイアス補正ができない

BMJ. 2017 Oct 16;359:4587.
Epidemiology. 2009;20(4):488-95.
Epidemiology. 2003;14(3):300-6.

時間依存性交絡の例1

例1: 高齢腎不全患者に対する慢性維持透析と死亡との関連

高齢の腎不全患者に対して慢性維持透析を行う場合と行わない場合のどちらが、生命予後が良いか?

慢性維持透析はeGFRの値を基準に導入を決定する。

eGFRは死亡と関連し、透析という治療によりその値が変化。

つまり、eGFRは時間依存性交絡因子、維持透析の開始は時間依存性治療、となる。

Kidney Int. 2018;94(3):582-8.

時間依存性交絡の例2

例2:炎症性腸疾患に対する免疫抑制剤と重症感染症との関連
炎症性腸疾患に対する免疫抑制剤 (thiopurine単独, anti-TNF単独, 併用)の種類により、重症感染症の発症頻度は変わるか？
免疫抑制剤の実施は、炎症性腸疾患の活動性により決定される。
炎症性腸疾患の活動性は重症感染症と関連し、免疫抑制剤という治療により変化する。
つまり、**炎症性腸疾患の活動性は時間依存性交絡因子**、免疫抑制剤治療は時間依存性治療、となる。

Gastroenterology, 2018;155(2):337-346.e10.

目次

1. 時間依存性交絡について
2. 時間依存性交絡の対処法と周辺構造モデル
3. 周辺構造モデルの逆確率重み付け法に必要な仮定
4. 周辺構造モデルの逆確率重み付け法の実際の手順

時間依存性交絡の対処法

1. 周辺構造モデルの逆確率重み付け法 (MSM-IPW)
2. パラメトリックG-formula
3. 構造ネストモデルのG推定法

Int J Epidemiol. 2017 01;46(2):756-62.

時間依存性交絡の対処法の比較

	利点	欠点
MSM-IPW	標準的な統計手法と似ていて理解しやすい ほとんどの統計ソフトで利用可能 暴露割付けの理由がわかっている場合に有用	極端な重みの存在下では不安定 曝露と時間変化する交絡因子の間の相互作用を 研究するためにはあまり有用ではない
G-formula	複数のリスク因子に対する介入 (共同介入) およ び動的介入を検討する研究に最適 リスク比やリスク差など、関心のある因果関係の 測定値を計算する	計算量が多く、余分なプログラミングを必要とし、 過適合問題を引き起こす可能性がある 結果と同様に交絡因子のモデルが必要 "G null paradox"
G推定法	曝露と時間的に変化する交絡因子の間の相互 作用を研究するのに有用	計算量が多く、余分なプログラミングを必要とし、 適合性の問題を引き起こす可能性がある

時間依存性交絡の対処法

1. 周辺構造モデルの逆確率重み付け法 (MSM-IPW)
2. パラメトリックG-formula
3. 構造ネストモデルのG推定法

時間依存性交絡の対処としてMSM-IPWが最も広く使用され
臨床医にも理解しやすい

重要用語の説明

- 潜在アウトカム
- 周辺構造モデル
- 因果推論
- 疑似集団

潜在アウトカム

	実際に治療を受けたかどうか	治療を受けた場合のアウトカム	治療を受けなかった場合のアウトカム
Aさん	YES	1	?
Bさん	NO	?	1
Cさん	NO	?	0

もし"?"を知りたいければタイムマシンを使って治療受けるかどうか以前に戻らないといけないが、実世界では観測できない(神のみぞ知る)。それぞれの治療パターンにおけるアウトカムを**潜在アウトカム**と呼ぶ。

潜在アウトカム

	実際に治療を受けたかどうか	治療を受けた場合のアウトカム	治療を受けなかった場合のアウトカム
Aさん	YES	1	?
Bさん	NO	?	1
Cさん	NO	?	0

"?"は実世界では観測できない(神のみぞ知る)ため個人に対して、直接因果を比較することはできない。
個人は調べられないから集団を対象として、疑似集団を作成して因果推論を行う。

周辺構造モデル(Marginal Structural Model)

- 因果推論のモデルの一つ
- 潜在アウトカムモデル
 - 時間依存性交絡の調整に応用

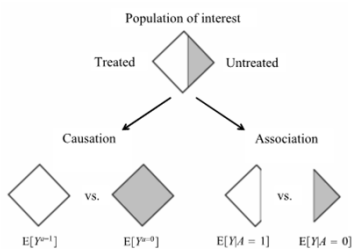
Epidemiology 2000;11:550-60.

周辺構造モデル(Marginal Structural Model)

- 周辺(マージナル) = 『潜在アウトカムの周辺分布モデル』
- 構造(ストラクチャー) = 『潜在アウトカムに対する回帰モデル』

潜在アウトカムという概念を使用した
疑似集団全体での回帰モデル

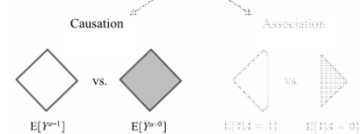
因果推論



Hernán MA, Robins JM (2020). Causal Inference.

因果推論

周辺構造モデルの逆確率重み付け法によって
疑似集団を作成し平均因果効果を推定する



Hernán MA, Robins JM (2020). Causal Inference.

目次

1. 時間依存性交絡について
2. 時間依存性交絡の対処法と周辺構造モデル
3. 周辺構造モデルの逆確率重み付け法に必要な仮定
4. 周辺構造モデルの逆確率重み付け法の実際の手順

周辺構造モデルの逆確率重み付け法に必要な仮定

1. Consistency -治療方法は単一であり、かつその効果は単一である
2. Exchangeability -未測定の変数因子が無い
3. Positivity -全ての患者層で、治療を受けた人と受けなかった人が存在する必要がある
4. No misspecification of the model -特に重み作成のモデルが正しい

Pharmacoepidemiol Drug Saf. 2014 June ; 23(6): 560-571.
Am J Epidemiol. 2008;168(6):656-664

周辺構造モデルの逆確率重み付け法に必要な仮定

1. Consistency -治療方法は単一であり、かつその効果は単一である
2. Exchangeability -未測定の変数因子が無い
3. Positivity -全ての患者層で、治療を受けた人と受けなかった人が存在する必要がある
4. No misspecification of the model -特に重み作成のモデルが正しい

Pharmacoepidemiol Drug Saf. 2014 June ; 23(6): 560-571.
Am J Epidemiol. 2008;168(6):656-664

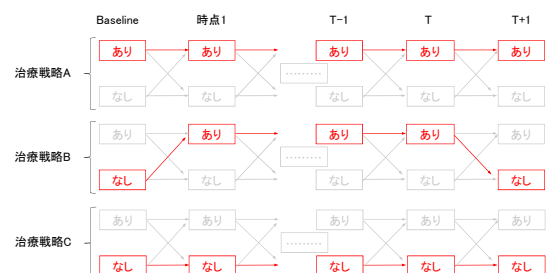
時間依存性治療(Time varying treatment)

- 対義語: Time fixed treatment 0 時点で治療状況を決定する
- 2つの時間依存性治療戦略とRandom treatment strategy
 1. Dynamic treatment strategy (動的な治療戦略)
時間依存性共変量に依存して治療が決定される (Treatment-by-indication)
“CD4数が300以下になった時に治療を開始する”など
 2. Static treatment strategy (静的な治療戦略)
時間依存性共変量に依存しない
“1ヶ月置きに治療を行う”、“最初の月以外は治療を行う”など

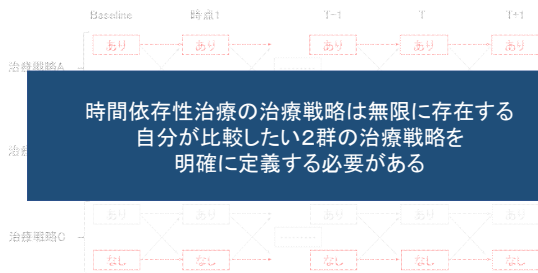
時間依存性治療(Time varying treatment)

- Random treatment strategy
 - より一般化して、特定の値で治療に割りつけるのではなく、治療確率を考慮する戦略(静的にも動的にもなり得る)
 - 例: CD4が300なので、確率0.7で治療を行い0.3で治療を行わない

時間依存性治療(Time varying treatment)



時間依存性治療(Time varying treatment)



周辺構造モデルの逆確率重み付け法に必要な仮定

1. Consistency - 治療方法は単一であり、かつその効果は単一である
2. Exchangeability - 未測定の変絡因子が無い
3. Positivity - 全ての患者層で、治療を受けた人と受けなかった人が存在する必要がある
4. No misspecification of the model - 特に重み作成のモデルが正しい

Pharmacoepidemiol Drug Saf. 2014 June ; 23(6): 560-571.
Am J Epidemiol. 2008;168(6):656-664

Exchangeability

- Exchangeability = No unmeasured confounding
- この仮定は観測されたデータからは検証できない
- Unmeasured confoundingに対する感度解析が推奨されるが 临床上重要な交絡をきちんと測定し、解析の中に組み込むことが大事

Stat Med. 2004 Mar 15;23(5):749-67

周辺構造モデルの逆確率重み付け法に必要な仮定

1. Consistency - 治療方法は単一であり、かつその効果は単一である
2. Exchangeability - 未測定の変絡因子が無い
3. Positivity - 全ての患者層で、治療を受けた人と受けなかった人が存在する必要がある
4. No misspecification of the model - 特に重み作成のモデルが正しい

Pharmacoepidemiol Drug Saf. 2014 June ; 23(6): 560-571.
Am J Epidemiol. 2008;168(6):656-664

Positivity

- 全ての交絡因子の患者層において、治療を受けた人と受けなかった人が存在している必要がある
- Positivityの存在下では重みが極端な値をとってバイアスを生じる
- 多くの交絡因子をいれることでPositivityの可能性が上昇
→ Bias-variance tradeoff
- Positivityのチェックを行う
- 重みをTruncationした感度解析(99%位、95%位など)を行う

Am J Epidemiol. 2008;168(6):656-664

周辺構造モデルの逆確率重み付け法に必要な仮定

1. Consistency - 治療方法は単一であり、かつその効果は単一である
2. Exchangeability - 未測定の変絡因子が無い
3. Positivity - 全ての患者層で、治療を受けた人と受けなかった人が存在する必要がある
4. No misspecification of the model - 特に重み作成のモデルが正しい

Pharmacoepidemiol Drug Saf. 2014 June ; 23(6): 560-571.
Am J Epidemiol. 2008;168(6):656-664

No misspecification of the model

- 重みにより、測定された交絡(measured confounders)と独立な pseudo-population (擬似集団)を作り出す
- 重みの分布を描く事がとても大事
- MSM-IPWの信頼性は、重みを正確にモデルできているかに依存(治療の重み、打ち切りの重み)

Am J Epidemiol. 2008;168(6):656-664

目次

1. 時間依存性交絡について
2. 時間依存性交絡の対処法と周辺構造モデル
3. 周辺構造モデルの逆確率重み付け法に必要な仮定
4. 周辺構造モデルの逆確率重み付け法の実際の手順

MSM-IPWの実際の手順

1. 治療に対する逆確率の重み(IPTW)を作成する
 - 1.1. 安定化された重み (Stabilized treatment weight) の作成
 - 1.2. 打ち切りの重み (censoring weight) の作成
 - 1.3. 最終的な重み (overall weight) の作成
2. IPTWで重み付けしたアウトカム-治療回帰モデルを適用する

The Stata Journal 2004;4:402-420.

MSM-IPWを用いた最初の論文(再掲)

HIV陽性の男性に対するZidovudineは死亡を減らすか？

Design	観察研究
Patient	HIV陽性男性で、AIDSを発症しておらず、治療薬も投与されていない
Exposure	Zidovudineの投与
Control	Zidovudineの非投与
Outcome	死亡

Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

1. 治療に対する逆確率の重み(IPTW)を作成する

- 1.1. 治療の重み (treatment weight) の作成

時間非依存性変数

年齢
西暦
観察開始時の検査値
(CD4数, CD8数, WBC, RBC, Plt)
観察開始時の症状の有無
(発熱, 口腔内カンジダ, 下痢,
体重減少, 口唇ヘルペス)

時間依存性変数

検査値
(CD4数, CD8数, WBC, RBC, Plt)
症状の有無
(発熱, 口腔内カンジダ, 下痢,
体重減少, 口唇ヘルペス)
AIDS疾患の有無

1. 治療に対する逆確率の重み(IPTW)を作成する

- 1.1. 治療の重み (treatment weight) の作成

治療を受ける時点ごとに傾向スコアを算出し、最終的に掛け算

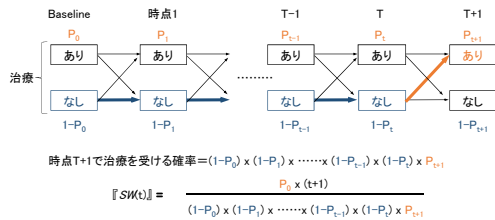
$$SW(t) = \prod_{k=0}^t \frac{f\{A(k)|\bar{A}(k-1), V\}}{f\{A(k)|\bar{A}(k-1), \bar{L}(k)\}}$$

掛け算
 時間非依存性交絡
 重み
 時点kで治療を受ける確率
 治療歴
 時間依存性交絡の歴史

1. 治療に対する逆確率の重み(IPTW)を作成する

1.1. 治療の重み (treatment weight) の作成

治療を受ける時点ごとに傾向スコアを算出し、最終的に掛け算



1. 治療に対する逆確率の重み(IPTW)を作成する

1.2. 打ち切りの重み (censoring weight) の作成

$$SW^\dagger(t) = \prod_{k=0}^t \frac{\Pr\{C(k) = 0 | \bar{C}(k-1) = 0, \bar{A}(k-1), V, T > k\}}{\Pr\{C(k) = 0 | \bar{C}(k-1) = 0, \bar{A}(k-1), \bar{L}(k-1), T > k\}}$$

重み

時点kでCensoring しない確率

時点k-1でCensoring していない歴史

治療歴

時間依存性交絡の歴史

時間非依存性交絡

Failure timeより時点kは小さいという条件

1. 治療に対する逆確率の重み(IPTW)を作成する

1.3. 最終的な重み (overall weight) の作成

$$SW(t) \times SW^\dagger(t)$$

治療の重み (treatment weight)

打ち切りの重み (censoring weight)

1. 治療に対する逆確率の重み(IPTW)を作成する

1.3. 最終的な重み (overall weight) の作成

重みの分布の確認とTruncationを考慮

治療の重み (treatment weight)

打ち切りの重み (censoring weight)

2. IPTWで重み付けしたアウトカム-治療回帰モデルを適用する

- 『最終的な重み(overall weight)』により、擬似集団を作り出す
- その擬似集団にどのような周辺構造モデルを当てはめるかで、求まる治療効果は異なる
- 周辺構造ロジスティックモデル
 - Marginal Structural Logistic Model
- 周辺構造Cox比例ハザードモデル
 - Marginal Structural Cox proportional hazards Model

2. IPTWで重み付けしたアウトカム-治療回帰モデルを適用する

- 同じ人が何回も出てくるので、クラスターを考慮
- 治療の重みを作成する際に分子に投入した変数は、全てアウトカムモデルに投入

HIV陽性の男性に対するZidovudineは死亡を減らすか

	ハザード比(95%信頼区間)
Crudeハザード比	3.55 (2.95-4.27)
Cox比例ハザードモデル	2.32 (1.92-2.81)
周辺構造Cox比例ハザードモデル	0.74 (0.57-0.96)

Hernán MA, Brumback B, Robins JM. Epidemiology. 2000;11(5):561-70.

まとめ

- 疫学的な曝露(治療)の多くは時間依存性
- 時間依存性交絡の存在下では、通常の解析方法では結果にバイアス
- 時間依存性交絡・時間依存性治療の対処法は周辺構造モデルの逆確率重み付け法(MSM-IPW)、パラメトリックG-formula、構造ネストモデルのG推定法
- 周辺構造化モデルとは、潜在アウトカムという概念を使用した疑似集団全体での回帰モデルのこと

まとめ

- 周辺構造モデルの逆確率重み付け法は"Consistency", "Exchangeability", "Positivity", "No misspecification of the model"を仮定している
- 疑似集団の作成には逆確率重み付け法を用いる
- 重みの作成の後には、重み付き回帰モデルを適用
- 重みの分布の確認を

生存時間分析における 競合リスクモデル

東京大学大学院医学系研究科ヘルスサービスマニagment学講座
道端伸明

本日の内容

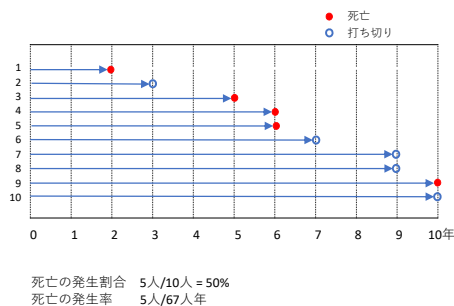
- 生存時間分析とは
- 従来の生存時間分析
 - カプランマイヤー法・ログランク検定
 - Cox回帰分析
 - 比例ハザード性の確認
- 競合リスクを考慮した生存時間分析
 - 原因別ハザード比の推定
 - 部分分布ハザード比の推定 (Gray検定、Fine and Grayモデル)

なぜ生存時間分析？

- アウトカムを在院死亡としたロジスティック回帰分析
入院3日で死亡、入院3か月で死亡。同じアウトカム!?
 - 生存時間解析 アウトカムが発生するまでの時間を考慮
- “生存時間”分析 → 生存時間 (survival time) 生死だけでない
例：がんの再発、術後退院までの時間など

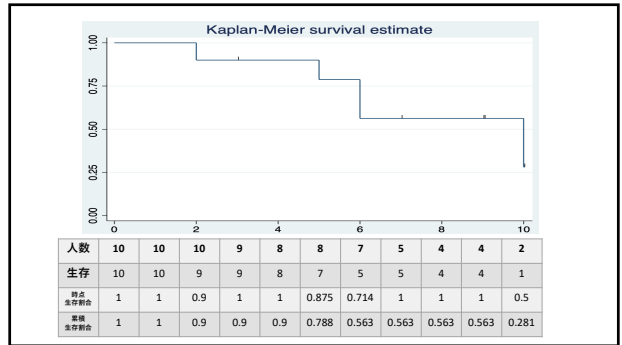
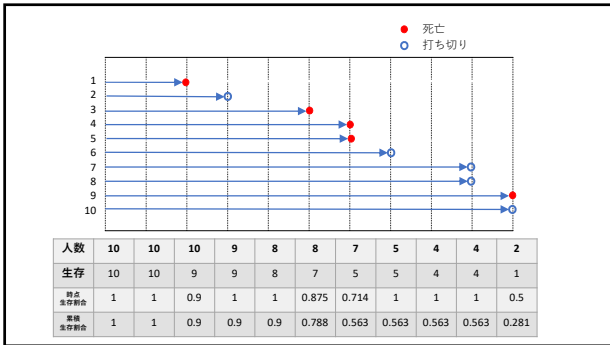
用語

- 生存時間 survival time
観察開始からアウトカム (イベント・エンドポイント) が発生するまでの時間
- ハザード hazard
観察期間のある点までアウトカムが発生していなかった人々が次の瞬間アウトカムを発生する確率
- 打ち切り censoring
研究終了期間前にアウトカムが発生せずに脱落 (転居など) したり、研究終了時までアウトカムが発生しない場合を打ち切りとする



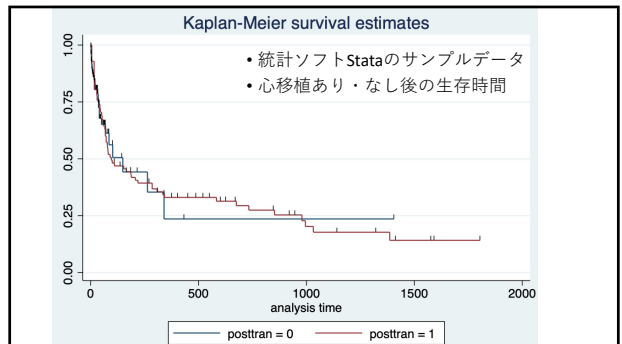
カプランマイヤー(Kaplan-Meier)法

- ある時点までに生存している者の割合 (生存割合) を表す生存関数をグラフ化する方法。生存曲線。
- アウトカムが発生する度に、その時点での生存割合を算出し、前時点までの生存割合に、時点での生存割合を掛けた累積生存割合を算出する。
- ノンパラメトリック



Kaplan-Meier (Kaplan-Meier)法

- 2群間の比較



ログランク(log-rank)検定

```

. sts test posttran
      failure_d: died
      analysis_time_t1: t1
      id: id

Log-rank test for equality of survivor functions
+-----+-----+-----+
| posttran | Events | Events |
|           | observed | expected |
+-----+-----+-----+
| 0         | 38      | 31.28    |
| 1         | 45      | 43.88    |
+-----+-----+-----+
| Total    | 75      | 75.00    |
+-----+-----+-----+

      ch12(1) = 0.13
      p1>ch12 = 0.7225
  
```

- 2群間に有意差は無い。
- 心移植の有無と生存時間に有意な関連があるとは言えない

Cox 回帰分析

- 生存時間分析における多変量回帰モデル
- 比例ハザード性を仮定している
 - ハザードは時間と共に変化して良い
 - 比較する2つのハザードの比が一定
- セミパラメトリック

Cox 回帰分析

Cox regression -- Breslow method for ties

No. of subjects = 103 Number of obs = 172
 No. of failures = 75
 Time at risk = 31938.1 LR chi2(3) = 14.38
 Log likelihood = -291.12672 Prob > chi2 = 0.0024

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
posttran	1.001496	.3058315	0.00	0.996	.5504453 1.822152
age	1.033265	.0144652	2.34	0.019	1.005299 1.062009
surgery	.3305094	.1420734	-2.58	0.010	.1423262 .7675082

比例ハザード性の確認

- シェーンフィールド残渣
- 二重対数プロット

シェーンフィールド残渣 (Schoenfeld residual)

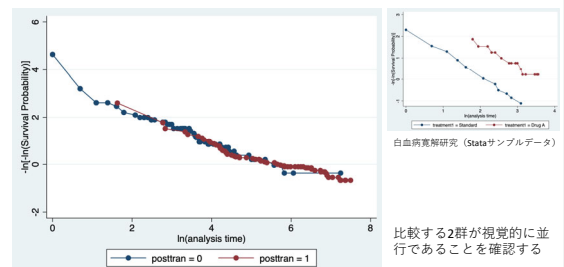
Test of proportional-hazards assumption

Time: Time

	rho	chi2	df	Prob>chi2
posttran	-0.00521	0.00	1	0.9638
age	0.12258	1.43	1	0.2314
surgery	0.30630	6.18	1	0.0129
global test		7.78	3	0.0508

※ 帰無仮説：比例ハザード性が一定 → 棄却してはいけない

二重対数プロット (log-log plot)



競合リスクを考慮した生存時間分析

競合リスク(competing risk)とは

- 関心のあるアウトカム (イベント・エンドポイント) 以外のアウトカムが生じることにより関心のあるアウトカムが観察できなくなる場合、**競合リスク**があるという。

例1：がんによる死亡がアウトカム
 心血管障害で死亡 → 競合リスク

例2：維持透析患者の死亡
 腎移植により維持透析から離脱 → 競合リスク

心血管障害による死亡 → → → 心血管障害で死亡しなければ、がんで死亡していたはず

従来の生存時間解析との違い

- カプランマイヤー法、Cox回帰はランダムな打ち切りを仮定
- 競合リスクが有る場合はランダムな打ち切りと仮定できない
- 心血管障害で死亡した患者と、観察を継続できた患者とは
関心のあるアウトカムの発生割合が異なる可能性がある

競合リスクが有る場合のバイアス

- 競合リスクによる打ち切りを受けた参加者が
(打ち切りを受けずに観察を継続できたとした場合)
関心のあるアウトカムの発生割合が
その後、他の観察を継続できている参加者の発生割合と
1. 同じ場合
→ バイアスはなし
 2. 比較して高い場合
→ 生存割合を過大評価してしまう
 3. 比較して低い場合
→ 生存割合を過小評価してしまう

競合リスクの対処法

- 複合エンドポイント (composite endpoint)
→ 心血管イベントまたはがんによる死亡
- 統計解析による対処法
 - 原因別ハザード比の推定
 - 部分分布ハザード比の推定

生存時間分析における統計手法

解析目的	従来の生存時間解析	競合リスクを考慮した生存時間解析	
		原因別ハザード比の推定	部分分布ハザード比の推定
リスクの推定 (グラフの描出)	Kaplan-Meier法 による生存割合、 累積発生割合	累積発生割合	累積発生関数 (cumulative incidence function, CIF)
2群 (以上) の単純比較	ログランク検定	ログランク検定	Gray検定
多変量回帰モデル	Cox回帰モデル	原因別Cox回帰モデル	Fine and Gray モデル

康永秀生、森田光治良ら、超絶解説医学論文の難解な統計手法が
手に取るようにわかる本、東京：金原出版、2019. より改変

原因別ハザード比の推定

- 競合リスク(アウトカム・イベント)の発生をランダムな
打ち切りとして扱う方法
- 次のことを仮定
 - 関心があるアウトカム以外は発生しない
 - 競合リスクは互いに独立 (関連が無い)
- 実際独立でなく、互いに正の相関があるときには、
関心があるアウトカムの発生割合を過大評価する
- 原因別ハザードは、ある時点まで、どのアウトカム
(関心あるアウトカム・競合リスクを含む) も発生していない時、
次の瞬間に関心のあるアウトカムが発生する確率
- 予後比較よりも、アウトカムの**リスク要因**に適した解析手法

部分分布ハザード比の推定

- 競合リスク(アウトカム・イベント)が発生した場合、
打ち切りとせず解析上は観察集団に残り続けるよう扱う方法
- 2群 (以上) の単純 (単変量) 比較には、各群の累積発生関数
(cumulative incidence function, CIF)を推定しGray検定を行う
- 多変量解析 (交絡因子を調整した回帰分析) には
部分分布ハザード比を推定するFine and Grayモデルを用いる
- 部分分布ハザードは、ある時点まで、関心あるアウトカムが
発生していない時、次の瞬間に関心のあるアウトカムが発生する確率
- アウトカム・イベント毎の**発生確率の推定・予測モデル**作成
- 2群間の**予後と比較**したい場合に適した解析手法

統計ソフトの対応について

- 原因別ハザード比の推定
 - Stata, R, SASで通常の生存時間分析のコマンドで可能
- 部分分布ハザード比の推定
 - R, SASは対応している
 - StataはGray検定以外には対応

原因別ハザード比と部分分布ハザード比の比較

ヨーロッパ腎臓透析移植学会のデータベースに登録された患者 (N=73382) を5年間追跡したデータ

関心のあるアウトカム：維持透析患者の死亡
競合リスク：腎移植

- Noordzij M, et al. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant.* 2013;28:2670-7.

原因別ハザード比と部分分布ハザード比の比較

5年間時点の累積発生割合	原因別ハザード比	部分分布ハザード比
死亡	60%	51%
腎移植	33%	24%
その他	25%	25%
合計	118%	100%

Noordzij M, et al. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant.* 2013;28:2670-7.

原因別ハザード比と部分分布ハザード比の比較

維持透析患者死亡をアウトカムにした多変量解析	原因別ハザード比	部分分布ハザード比
男性	1.04 (1.02-1.07)	1.03 (0.87-1.23)
高齢者	2.57 (2.52-2.63)	3.47 (3.39-3.55)

性別は競合リスクである腎移植と関連が低いためどちらの方法でもほぼ同じ結果。一方、高齢者は腎移植を受けにくくなるため、原因別ハザード比では過小評価される。

Noordzij M, et al. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant.* 2013;28:2670-7.

論文紹介

P: 介護老人保健施設入所者 (N=342758)

O: 在宅復帰

競合リスク：死亡、病院へ入院、他介護施設への移動

在宅復帰に関連するリスク要因の探索が目的、原因別ハザード比の推定を行った。

結果：追跡期間中央値は137日で、19%の入所者が在宅復帰した。

原因別Cox回帰分析の結果、高齢、高い要介護度、複数の合併症、介護施設の運営形態が私立、施設規模が大きいなどが有意に在宅復帰を妨げる要因として関連していた。

- Morita K, et al. Factors Affecting Discharge to Home of Geriatric Intermediate Care Facility Residents in Japan. *J Am Geriatr Soc.* 2018;66:728-34.

論文紹介

P: 初回冠動脈バイパス術/弁置換・形成術を受けた65歳以上心房細動患者 (N=10524)

E: 左心耳閉鎖術併用あり (n=3892)

C: 左心耳閉鎖術併用なし (n=6632)

O: 血栓塞栓症の発症

競合リスク：死亡

結果：追跡期間平均2.6年、全体の5.4%に血栓塞栓症発症し、全体の21.5%が死亡。

Fine and Grayモデルを用いた調整済み部分分布ハザード比は、0.67 (95%信頼区間 0.56-0.81)と、左心耳閉鎖術併用は、血栓塞栓症の発症を減少させる可能性を示唆した。

- Friedman DJ, et al. Association Between Left Atrial Appendage Occlusion and Readmission for Thromboembolism Among Patients With Atrial Fibrillation Undergoing Concomitant Cardiac Surgery. *JAMA.* 2018;319:368-74.

参考文献

1. 康永秀生、森田光治良ら. 超絶解説医学論文の難解な統計手法が手に取るようにわかる本. 東京: 金原出版; 2019.
2. Noordzij M, et al. When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant*. 2013;28:2670-7.
3. Morita K, et al. Factors Affecting Discharge to Home of Geriatric Intermediate Care Facility Residents in Japan. *J Am Geriatr Soc*. 2018;66:728-34.
4. Friedman DJ, et al. Association Between Left Atrial Appendage Occlusion and Readmission for Thromboembolism Among Patients With Atrial Fibrillation Undergoing Concomitant Cardiac Surgery. *JAMA*. 2018;319:365-74.

欠損値の取扱と多重代入法

東京大学大学院医学系研究科 生物統計情報学講座

麻生将太郎

Contents

- ◆ 欠損値
- ◆ 欠損値の種類
- ◆ 単一代入法
- ◆ 多重代入法
- ◆ 使用例

欠損値

例えばカルテからデータ収集する…

既往歴から高血圧の記載は、
「高血圧あり」、「高血圧なし」、「」（聞いていない）のいずれか

入院前のADLの記載は、
「自立」、「一部介助」、「全介助」、「」（聞いていない）のいずれか

受診時の呼吸数は、
「1~∞の数値」、「」（測定していない）

ID	高血圧	ADL	呼吸数
1	「あり」	「自立」	「15」
2	「なし」	「」	「」
3	「」	「一部介助」	「」
4	「あり」	「」	「18」
5	「あり」	「全介助」	「」
...

「」が欠損値

欠測が生じるには理由がある

理由で欠損値の種類が決まる

欠損値のメカニズムによる分類

- ◆ MCAR
- ◆ MAR
- ◆ MNAR

MCAR

- ◆ Missing completely at random (完全にランダムな欠測)
- ◆ データの欠測が完全にランダムに発生している状況

MCARの例

- ◆ 体重計が壊れた
→ 体重が測れない
→ 体重の値が欠測

ID	体重	年齢	性別
1	40	30	男
2	(35)	60	女
3	60	65	女
4	(65)	35	男

- ◆ 体重の値、他の変数(年齢や性別)で欠測したわけではない(偶然)

MAR

- ◆ Missing at random (ランダムな欠測)
- ◆ データの欠測が完全にランダムではない
- ◆ 欠測した変数自体で欠測した理由は説明できない
- ◆ しかし、他の変数で欠測を説明できる

MARの例

- ◆ 若年者で体重の欠測が多く発生
→ 若年者は体重測定しない
- ◆ 体重の値自体で欠測したわけではない
- ◆ しかし、年齢によって欠測が発生

ID	体重	年齢	性別
1	(35)	30	女
2	40	60	男
3	60	65	女
4	(65)	35	男

MNAR

- ◆ Missing not at random (ランダムではない欠測)
- ◆ データの欠測する理由が、欠測した変数自体で説明できる

MNARの例

- ◆ 肥満の人の体重の欠測が多く発生
→ 肥満の人は測定しない傾向
- ◆ 体重の値自体で欠測が生じる

ID	体重	年齢	性別
1	40	30	男
2	(80)	60	女
3	45	65	女
4	(90)	35	男

欠損値の種類を見分ける

- ◆ MCAR、MAR、MNARを客観的に判断するのは難しい
- ◆ 欠損値を予測する変数を増やすことでMARの仮定を満たす確率上昇
- ◆ Little's test → MCARの判断 → 帰無仮説がMCAR
→ $P \geq 0.05$ でMCAR ($P < 0.05$ でMAR or MNAR)

対処法

- ◆ 完全ケース分析 (complete case analysis)
- ◆ 欠損値を表す指標を用いる方法
- ◆ 単一代入法
- ◆ 多重代入法 (multiple imputation)

完全ケース分析 (complete case analysis)

- ◆ 欠測のない対象者のみで解析
- ◆ MCARでは結果は変わらない

欠点


- ◆ 母集団を代表していない
- ◆ MNARの欠損値を除外するとバイアス
- ◆ 症例数が少 → 検出力が低下
→ 信頼区間が広がる

ID	体重	年齢	性別
1	40	30	男
2	(80)	60	女
3	45	65	女
4	(90)	35	男

欠損値を表す指標を用いる方法

- ◆ 欠損値に適切な値を代入してカテゴリとするような方法

ID	高血圧	ADL	呼吸数
1	1	1	15
2	0		
3		2	
4	1		18
5	1	3	
...



ID	高血圧	ADL	呼吸数
1	1	1	15
2	0	9	99
3	9	2	99
4	1	9	18
5	1	3	99
...

変数	オッズ比	標準誤差	P値
高血圧			
なし	Reference		
あり	2.42	3.07	0.488



変数	オッズ比	標準誤差	P値
高血圧			
なし	Reference		
あり	2.42	3.07	0.488
欠損	9.15	6.32	0.001

- ◆ 高血圧あり、なしの2群 → あり、なし、欠損の3群にする
- ◆ 欠損群は高血圧なし群に比べて、オッズ比9倍?
- ◆ 解釈困難な結果

推奨しない

代入法

欠測値を何らかの値に置き換えてデータを疑似完全状態にする

- ◆ 単一代入法
- ◆ 多重代入法

単一代入法

- ◆ 平均値代入法
- ◆ 比率代入法
- ◆ Hot-deck法
- ◆ Cold-deck法
- ◆ 回帰代入法
- ◆ LOCF法

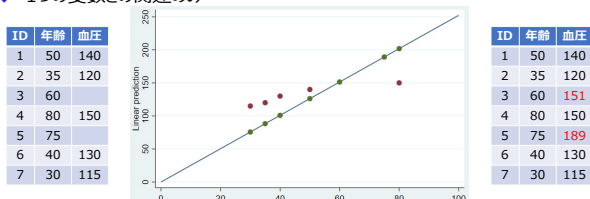
平均値代入法

- ◆ データのある集団から平均値を求め、欠損値を平均値に置き換える
- ◆ 百害あって一利なし!!!

ID	高血圧	ADL	呼吸数
1	1	1	15
2	0	2	16.5
3	0.75?	2	16.5
4	1	2	18
5	1	3	16.5

比率代入法

- ◆ 欠損値を変数と関連する変数を切片なしで単回帰して予測した値を代入
- ◆ 1つの変数との関連のみ



Hot-deck法

- ◆ 欠損値のある属性に近いデータを同じデータセットから見つけ、その値を代入
- ◆ 例えば、血圧が欠損の65歳女性は、同じデータセットの中から65歳女性を探し出して、その血圧の値を欠損値を代入
- ◆ 同じ属性が複数の場合は、その中から無作為に1件を選択

Cold-deck法

- ◆ 欠損値のある属性に近いデータを異なるデータセットから見つけ、その値を代入
- ◆ 例えば、血圧が欠損の65歳女性は、異なるデータセットの中から65歳女性を探し出して、その血圧の値を欠損値を代入
- ◆ 同じ属性が複数の場合は、その中から無作為に1件を選択

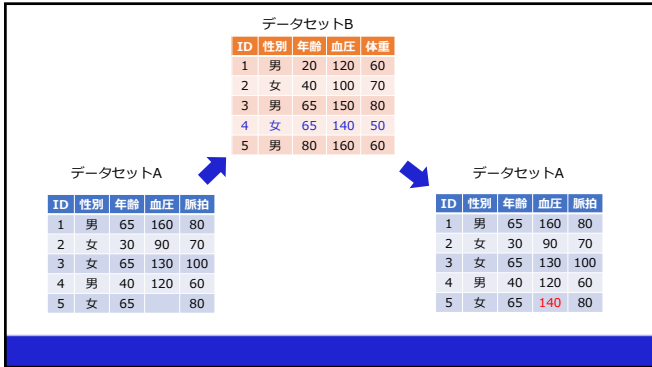
ID	性別	年齢	血圧
1	男	65	160
2	女	30	90
3	女	65	130
4	男	40	120
5	女	65	

➡

ID	性別	年齢	血圧
1	男	65	160
2	女	30	90
3	女	65	130
4	男	40	120
5	女	65	

➡

ID	性別	年齢	血圧
1	男	65	160
2	女	30	90
3	女	65	130
4	男	40	120
5	女	65	130



回帰代入法

- ◆ 欠損値のある変数を他の変数で回帰分析してモデルを作成
- ◆ モデルから推定値を出して欠損値に代入する
- ◆ モデルに誤差をランダムに加える場合もある（確率的代入法）

呼吸数 = a × 性別 + b × 年齢 + c × 血圧 + d × 脈拍 + e × 体温 + f

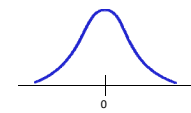
ID	性別	年齢	血圧	脈拍	呼吸数	体温
1	男	30	120	60	15	36.5
2	女	45	120	80	16.9	37.0
3	男	50	140	70	10	38.0
4	女	40	110	65	20	37.5
5	男	60	150	90	7.5	36.0

呼吸数 = a × 0 + b × 45 + c × 120 + d × 80 + e × 37.0 + f

呼吸数 = a × 1 + b × 60 + c × 150 + d × 90 + e × 36.0 + f

確率的回帰代入法

- ◆ 回帰代入法で得られた予測値に誤差項を加えた値を代入
- ◆ 誤差項は平均値0、分散が回帰モデルの残差分散の正規分布から無作為に抽出



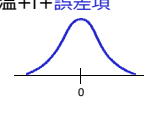
※ 残差分散：実測値回帰の予測値の差の平方和を(データ数-変数の数-1)で割った値

呼吸数 = a × 性別 + b × 年齢 + c × 血圧 + d × 脈拍 + e × 体温 + f + 誤差項

ID	性別	年齢	血圧	脈拍	呼吸数	体温
1	男	30	120	60	15	36.5
2	女	45	120	80	14.9	37.0
3	男	50	140	70	10	38.0
4	女	40	110	65	20	37.5
5	男	60	150	90	9	36.0


呼吸数 = a × 0 + b × 45 + c × 120 + d × 80 + e × 37.0 + f - 2

呼吸数 = a × 1 + b × 60 + c × 150 + d × 90 + e × 36.0 + f + 1.5



LOCF法

- ◆ Last observation carried forward法
- ◆ 経時データで欠測が起きた場合、最後に観察された値で欠損値を埋める
- ◆ 「脱落後の変数の推移は最後に観察された値のまま変化しない」強い仮定が必要
- ◆ 推定精度を過小評価
- ◆ 最悪値で補完する方法もあるが、保守的で強い仮定が必要



ID	時点1	時点2	時点3	時点4	時点5
1	200	1000	5000	3500	2000
2	5000	8000	10000	15000	15000
3	1000	3000	3000	4000	1000
4	10000	8000	8000	8000	6000
5	100	100	100	100	100

ID	時点1	時点2	時点3	時点4	時点5
1	200	1000	5000	3500	2000
2	5000	8000	10000	15000	
3	1000	3000		4000	1000
4	10000	8000			6000
5	100				

単一代入法

- ◆ 個々の欠測データを1つの値で代入する方法
- ◆ 欠測値5%未満だと良好な結果を得ることもある
- ◆ 1つのモデルで推定した推定値を代入するため、不確実性が高い
(代入値が測定された値に近いか不明)
- ◆ 推定値の分散を過小評価 (信頼区間狭くなる) → 誤った結果

推奨しない

多重代入法

なぜ多重代入法か

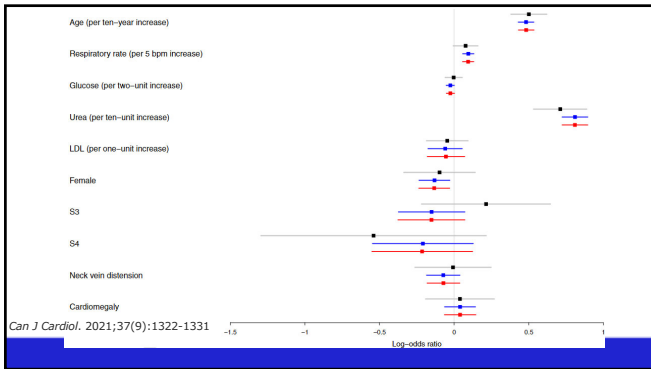
- ◆ 欠損値によるバイアスは代入法では是正可能
- ◆ 代入法では完全に欠損値を復元できるわけではない(単一でも多重でも)
- ◆ 代入法の目的はバイアスを減少し、真の値を推定
- ◆ 代入法の誤差を適切に評価する必要

なぜ多重代入法か

- ◆ 単一代入法だと代入モデルによって代入した値は変わる
- ◆ 単一代入法で得られた代入値の背後には様々な可能性が隠されている
- ◆ 一つの値を代入して解析することは不確実性が高い
- ◆ 複数の値を代入して不確実性を減少することが必要

多重代入法のメリット

- ◆ 不確実性下げる
- ◆ バイアスを減らし精度を高める
- ◆ Complete case analysisより検出率向上、信頼区間も広くならない

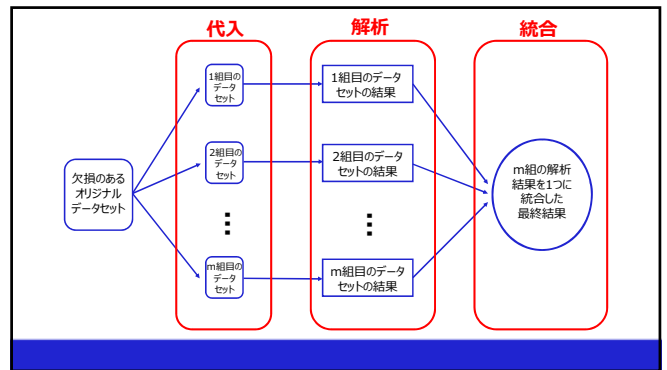


多重代入法の手順

- ① 代入
- ② 解析
- ③ 統合

多重代入法の手順

- ① 代入
- ② 解析
- ③ 統合



① 代入

- ◆ オリジナルデータから欠損を埋めた完全データをm個作成
- ◆ 欠損値はMARを仮定

アルゴリズム

- ◆ DA (Data Argumentation)法
- ◆ EMB (Expectation-Maximization with Bootstrapping)法
- ◆ FCS (Fully Conditional Specification)法

アルゴリズム

- ◆ ベイズ統計学の枠組みを使用
- ◆ どの方法もMARを仮定
- ◆ DA法とEMB法は正規分布を仮定
- ◆ DA法とFCS法はマルコフ連鎖モンテカルロ法に基づく
- ◆ 少しだけFCS法の精度が良い

FCS法

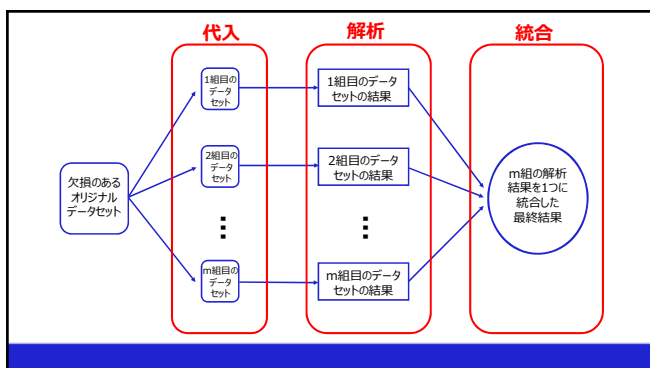
- ◆ MICE (multivariate imputation by chained equations)
- ◆ 柔軟な方法で、多変量の正規分布の仮定が不要 (条件付き分布)
- ◆ 変数ごとに回帰モデルを作成
- ◆ 2値変数、順序変数をアウトカムとしたlogistic回帰などのモデルも使用可

FCS法

- ◆ 「各変数の欠測に対する補完モデルとして、欠測を補完したい変数以外の他のすべての変数によってお互いに説明可能」という強い仮定
- ◆ この仮定を満たせない場合、推定値の分散が小さい (信頼区間が狭くなる)

多重代入法の手順

- ① 代入
- ② 解析
- ③ 統合

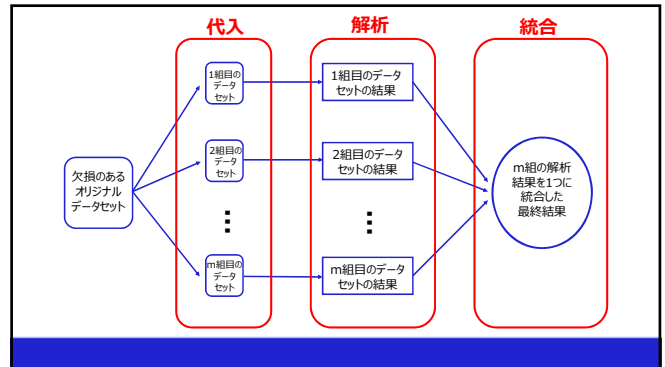


②解析

- ◆ 作成したm個のデータセットを各々別々に解析
- ◆ m個のデータセットをまとめて1つのデータセットとして解析してはいけない

多重代入法の手順

- ① 代入
- ② 解析
- ③ 統合



③統合

- ◆ 各データセットで得られた結果を統合する
- ◆ 統合の方法はRubin's ruleに従う

Rubin's rule

推定値はm個のデータセットの推定値の平均 $\bar{Q} = \frac{1}{m} \sum_{i=1}^m \hat{Q}_i$

分散はm個のデータセットの分散の平均 $\bar{U} = \frac{1}{m} \sum_{i=1}^m U_i$

データセット間の分散 $B = \frac{1}{m-1} \sum_{i=1}^m (\hat{Q}_i - \bar{Q})^2$

全体の分散 $T = \bar{U} + \left(1 + \frac{1}{m}\right) B$

95%信頼区間 $\bar{Q} \pm 1.96 * \sqrt{T}$

m データセットの数
 \bar{Q} 推定値の平均
 \hat{Q}_i 各データセットの推定値
 T 全体の分散
 \bar{U} 分散の平均
 U_i 各データセットの分散
 B データセット間の分散

データセット	推定値	分散
1	10	5
2	15	3
3	7	2
4	6	10
5	12	7

$\bar{Q} = \frac{1}{m} \sum_{i=1}^m \hat{Q}_i$ 推定値 = (10+15+7+6+12)/5 = 10

$\bar{U} = \frac{1}{m} \sum_{i=1}^m U_i$ 分散 = (5+3+2+10+7)/5 = 5.4

$B = \frac{1}{m-1} \sum_{i=1}^m (\hat{Q}_i - \bar{Q})^2$ データセット間の分散
 = ((10-10)² + (15-10)² + (7-10)² + (6-10)² + (12-10)²) / (5-1) = 13.5

$T = \bar{U} + \left(1 + \frac{1}{m}\right) B$ 全体の分散 = 5.4 + (1+1/5)*13.5 = 21.6

Rubin's rule

全体の分散 $T = \bar{U} + \left(1 + \frac{1}{m}\right) B$

$\bar{U} < B$ データセット数を増やすと精度が上がる

$\bar{U} > B$ データセット数を増やしても精度は上がらない

データセット	推定値	分散
1	10	5
2	15	3
3	7	2
4	6	10
5	12	7

推定値 = 10

分散 = 5.4 データセット間の分散 = 13.5

$\bar{u} < B$

全体の分散 = $5.4 + (1 + 1/5) * 13.5 = 21.6$

全体の分散 $T = \bar{u} + (1 + \frac{1}{B})B$

データセット数を増やすと全体の分散が小さくなる

データセットの数は？

- ◆ 開発された当初は5セットで十分とされていた
- ◆ しかし、最近では50セット、100セット以上必要という主張も
- ◆ ガイドラインでは少なくとも20セット作成が推奨

代入モデルで用いる変数は？

- ◆ MARで欠測が起きる状況を説明できる変数（補助変数）を使用
- ◆ できるだけ多くの補助変数と欠測を説明するのに重要な変数は必須
- ◆ 解析で使用する変数だけでなく、
解析に使用しなくても欠損に関連する変数も投入

代入モデルで用いる変数は？

- ◆ MNARでも、補助変数を多数投入 → MARを満たす確率が上昇
- ◆ 説明変数が欠損値でも、代入モデルにアウトカム変数の投入が推奨
- ◆ 解析のステップでは代入モデルに含まれていない変数の使用×

アウトカム変数の欠損に使えるか？

- ◆ 使えます
- ◆ RCTでアウトカムの追跡不能問題から開発
- ◆ MARの仮定が重要
- ◆ 従属変数と説明変数を数学的に区別しない

THE NEW ENGLAND JOURNAL OF MEDICINE

ORIGINAL ARTICLE

Trial of Satralizumab in Neuromyelitis Optica Spectrum Disorder

T. Yamamoto, I. Kleiter, K. Fujihira, J. Pilsatz, B. Greenberg, B. Zakrzewska-Pomaska, F. Patti, C.-P. Tsai, A. Sato, H. Yamazaki, Y. Kawata, P. Wright, and J. De Seze

N Engl J Med 2019;381:2114-24.

使用例

Antenatal blood pressure for prediction of pre-eclampsia, preterm birth, and small for gestational age babies: development and validation in two general population cohorts
BMJ 2015;101:h15948

妊婦の平均血圧が子癇前症、早産、胎内発育遅延を
予測するモデルの開発と妥当性を研究

記載例

For both the ALSPAC and SWS datasets there were missing data on maternal characteristics and also for the derived values of initial mean arterial pressure and mean arterial pressure at 20, 25, 28, 31, 34, and 36 weeks for some women.

To increase power and minimise selection bias in the second stage prediction models, we next used multivariable multiple imputation of missing data.

For each imputation model, we included all early pregnancy characteristics and derived mean arterial pressure values as well as the outcomes.

データセットには母体の背景因子と平均血圧に欠損値が含まれる

検出力を高め、選択バイアスを最小化するために多重代入法を用いた

モデルには患者背景、平均血圧、アウトカムも投入

記載例

For each imputation we generated 20 imputed datasets and combined coefficient estimates across these using Rubin's rules.

Imputation of missing data in this way assumes that data are missing at random—that is, that any reasons for missingness are explained by the observed variables included in the imputation model.

The assumption seems reasonable here as most women will not be aware of their current blood pressure (only their previous values) so this is unlikely to influence their decision to have their blood pressure measured at any given gestational age.

20の完全なデータセットを作成
Rubin's ruleで結果を統合

欠損値はMARを仮定して代入
MARは他の変数で欠損が説明
できる

MARの仮定は妥当である説明

最近は、当然の如く、あっさり説明

The NEW ENGLAND
JOURNAL of MEDICINE

Childhood Cardiovascular Risk Factors and Adult
Cardiovascular Events

All primary analyses were performed after multiple imputation of missing values by means of chained equations with fully conditional specification (10 replications) in PC-SAS software (version 9.4, SAS Institute); data were assumed to be missing at random.¹¹

多重代入法を行い解析している(10セット複製)。MARを仮定している。

N Engl J Med. 2022;386(20):1877-1888.

まとめ

- ◆ 欠測値の種類 (MCAR, MAR, MNAR)
- ◆ 単一代入法
- ◆ 多重代入法

マルチレベル分析

東京大学大学院医学系研究科リアルワールドエビデンス講座 特任准教授
笹淵 裕介

Agenda

- マルチレベル分析とは
- マルチレベル分析の統計モデル
- マルチレベルモデルを利用した研究紹介

Agenda

- マルチレベル分析とは
- マルチレベル分析の統計モデル
- マルチレベルモデルを利用した研究紹介

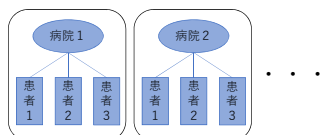
マルチレベル分析とは

階層構造をもつデータを分析する方法

階層構造

複数のレベルをもつデータ

- ✓病院 — 患者
- ✓地域 — 住民
- ✓患者 — 時点
- ✓病院 — 患者 — 時点



なぜマルチレベル分析が必要なのか

階層構造を考慮しない回帰分析

- 各個人のデータは独立であることを仮定

階層構造をもつデータ

- 同じ集団に属する個人は独立であると言えない
- 同じ患者から得られたデータは似通っている
グループ内のデータには相関がある（級内相関）

なぜマルチレベル分析が必要なのか

✓病院 - 患者

施設によって

- 専門性
- スタッフの習熟度
- 最新技術が利用可能
- 管理体制
などが異なる



背景が一致する患者を
同じ病院から選ぶと

異なる病院から選ぶ
ときと比較して

アウトカムは似通う

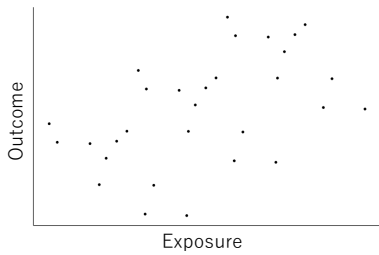
なぜマルチレベル分析が必要なのか

✓患者 - 時点

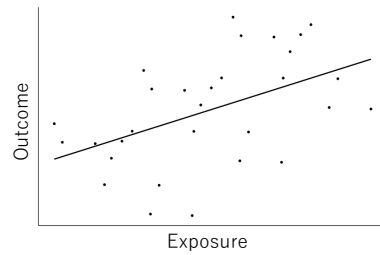
患者の重症度を複数時点にわたって測定

- 最初に測定した重症度はその後の重症度に影響がある
- どこかの時点で差があれば他の時点でも差がつきやすい

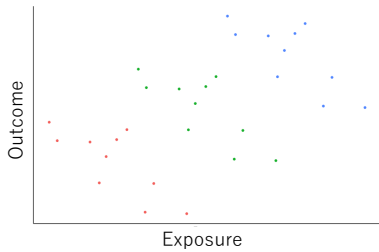
階層を考慮しない分析



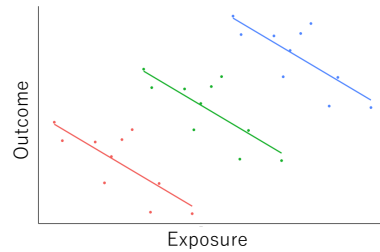
階層を考慮しない分析



階層を考慮した分析



階層を考慮した分析



階層構造を考慮しないと

- 誤った結果を導く
- 第一種の過誤の可能性が増大する
 - サンプルをコピーして検出力を上げているかのよう

マルチレベル分析を行う目的

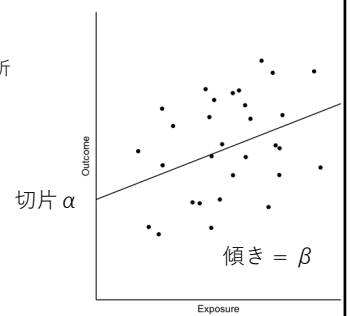
- ▶ 階層構造を考慮する
 - 多施設研究やアウトカムの繰り返し測定
- ▶ グループの特性が個人に与える影響を調べる
 - 施設の特性が個人に与える影響を知りたい
- ▶ グループ間のばらつきを調べる
 - グループ間でのばらつき、あるいはその要因を調べる

Agenda

- マルチレベル分析とは
- **マルチレベル分析の統計モデル**
- マルチレベルモデルを利用した研究紹介

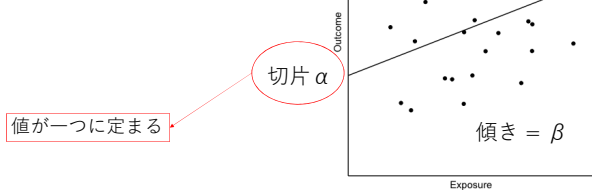
統計モデル

階層構造を考慮しない回帰分析
 $Y = \alpha + \beta \times x + \varepsilon$



統計モデル

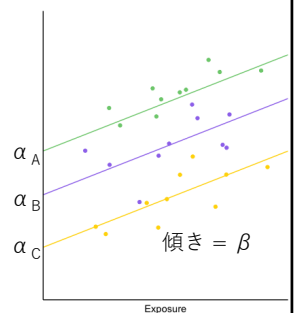
階層構造を考慮しない回帰分析
 $Y = \alpha + \beta \times x + \varepsilon$



統計モデル

マルチレベルモデル

- A病院 $Y_A = \alpha_A + \beta \times x + \varepsilon$
- B病院 $Y_B = \alpha_B + \beta \times x + \varepsilon$
- C病院 $Y_C = \alpha_C + \beta \times x + \varepsilon$



統計モデル

マルチレベルモデル

A病院

$$Y_A = \alpha_A + \beta \times x + \varepsilon$$

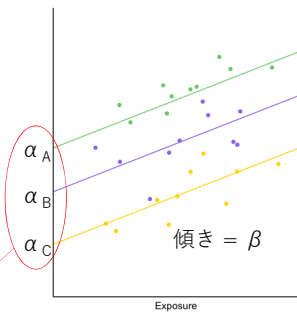
B病院

$$Y_B = \alpha_B + \beta \times x + \varepsilon$$

C病院

$$Y_C = \alpha_C + \beta \times x + \varepsilon$$

切片の全体平均 + 施設ごとの誤差
全体平均を中心にランダムに分布



統計モデル

マルチレベルモデル

$$Y = \alpha + \beta \times x + \varepsilon$$

$$\alpha = \gamma + u$$

γ : 切片の全体平均

u : 切片のグループごとの誤差

ランダム切片モデル

結果の解釈

$$Y = \alpha + \beta \times x + \varepsilon$$

$$\alpha = \gamma + u$$

γ : 切片の全体平均

u : 切片のグループごとの誤差

係数 β : x が1単位増加したときのアウトカムの変化

切片 α : 各グループのアウトカムの大きさ

誤差 u の分散: 切片のばらつき大きさ

分散が0でないならばグループごとに切片が異なる

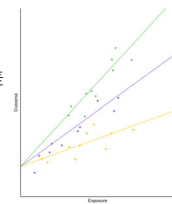
様々なマルチレベルモデル

$$Y = \alpha + \beta \times x + \varepsilon$$

$$\beta = \gamma + u$$

γ : 傾きの全体平均

u : 傾きのグループごとの誤差



様々なマルチレベルモデル

$$Y = \alpha + \beta \times x + \varepsilon$$

$$\alpha = \gamma_0 + u_0$$

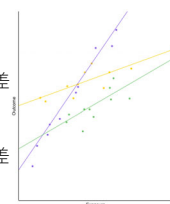
γ_0 : 切片の全体平均

u_0 : 切片のグループごとの誤差

$$\beta = \gamma_1 + u_1$$

γ_1 : 傾きの全体平均

u_1 : 傾きのグループごとの誤差



Agenda

- マルチレベル分析とは
- マルチレベル分析の統計モデル
- マルチレベルモデルを利用した研究紹介

マルチレベル分析の実例

1. 腰椎すべり症に対する硬膜外ステロイド投与の効果
2. 小児の下気道感染による入院後の再入院割合

Epidural Steroid Injections for Management of Degenerative Spondylolisthesis

J Bone Joint Surg Am. 2020;102(15):1297-1304.

背景

- 腰椎変性すべり症は神経性跛行や腰痛の一般的な要因
- 理学療法や硬膜外腔へのステロイド投与が一般的
- ステロイド投与
 - 椎間板ヘルニア→疼痛緩和・手術回避
 - 脊柱管狭窄症→QOL低下と関連
 - 腰椎変性すべり症→効果はわかっていない

研究目的

- 腰椎変性すべり症に対するRCTと前向きコホート研究のデータ

- P 腰椎変性すべり症
E 硬膜外ステロイド投与
I ステロイド投与なし
O 疼痛・身体機能（最大4年まで反復測定）

統計モデル

- 個人-時点
- ランダム切片モデル
$$\text{疼痛} \cdot \text{身体機能} = \alpha + \beta \times \text{交絡要因} + \varepsilon$$
$$\alpha = \gamma + u$$

γ : 切片の全体平均
 u : 切片の個人毎の誤差

結果

- 非手術患者
疼痛・身体機能とも硬膜外ステロイド投与により悪化
- 手術患者
疼痛・身体機能とも硬膜外投与の有無で有意差なし

Pediatric Readmissions After Hospitalizations for Lower Respiratory Infections

Pediatrics. 2017 Aug;140(2):e20160938.

背景

- 下気道感染
小児科入院の最も一般的な理由の一つ
多くの再入院につながっている
- 再入院率
20-30%は避けられる可能性

研究目的

- 診療報酬請求 (Medicaid) データ
- 2007年12月1日~2009年11月30日

下気道感染で入院した18歳以下の患者

30日以内の予定外の入院に影響する要因
再入院が病院によって異なるかどうか

統計モデル

- 病院-患者
- ランダム切片モデル
$$\text{再入院} = \alpha + \beta \times \text{患者背景} + \varepsilon$$
$$\alpha = \gamma + u$$

γ : 切片の全体平均
 u : 病院ごとの誤差

統計モデル

患者背景の係数
再入院に与える影響の大きさ

各病院の切片
各病院の再入院率の大きさ

切片の分散
再入院率のばらつき

結果

再入院のリスク因子
年齢・性別・慢性疾患

各病院の調整後再入院確率
5.2%(3.8 - 8.8%)
1SD上の病院→6.5%
1SD下の病院→4.2%

まとめ

- 階層構造を考慮しない分析を行うと誤った結果を導く
- マルチレベル分析は3つの目的に利用する
 - 階層構造を考慮する
 - グループの特性が個人に与える影響を調べる
 - グループ間のばらつきを調べる
- 切片・傾きがグループごとに異なるモデル

自己対照ケースシリーズ (Self-Controlled Case Series: SCCS)

東京大学大学院医学系研究科臨床疫学・経済学
橋本 洋平

1

本日の内容

1. イントロダクション
2. SCCS研究とは
3. 古典的な研究デザインとの比較
4. SCCSの必要条件
5. よく受ける質問

2

コホート研究/ケースコントロール研究

- ランダム化比較試験 (RCT) は倫理的・費用的制約が大きい
- そこで・コホート研究やケースコントロール研究が行われる
- しかし・未測定交絡因子の問題が残る
- たとえ交絡因子が全て測定できたとしても・因果パス (causal pathway) を完璧に描くのは不可能である
- 結果・残差交絡が生じる

3

自己対照研究デザイン (SCCS: self-controlled study design)

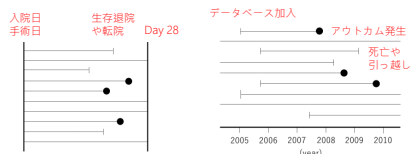
- Event (アウトカム) が起きた人のみを対象とする
- 同一人の過去および未来の期間と比較する
- 同一人内の比較のため・時間で変化しない因子 (性別・遺伝子・生活習慣・社会経済的背景) は自然に相殺される
- 因果の追求のため (だけ) に用いる

	Exposure-indexed	Outcome-indexed
個人内比較	SCCS	Case-crossover
個人間比較	コホート	ケースコントロール

4

入院データ vs. 外来 (一般住民) データ

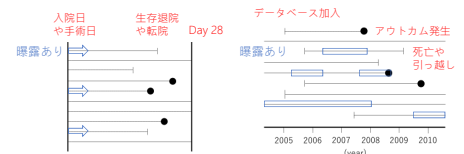
	入院データベース (例: DPC, MDV)	外来 (一般住民) データベース (例: JMDC, NDB)
(i) コホートの特性	Closed	Open
(ii) 曝露因子の特性	(多くが) time-invariant	(多くが) time-variant
(iii) アウトカムの特性	(多くが) 死亡	様々なアウトカム



5

入院データ vs. 外来 (一般住民) データ

	入院データベース (例: DPC, MDV)	外来 (一般住民) データベース (例: JMDC, NDB)
(i) コホートの特性	Closed	Open
(ii) 曝露因子の特性	(多くが) time-invariant	(多くが) time-variant
(iii) アウトカムの特性	(多くが) 死亡	様々なアウトカム



6

入院データ vs. 外来 (一般住民) データ

	入院データベース (例: DPC, MDV)	外来 (一般住民) データベース (例: JMDC, NDB)
(i) コホートの特性	Closed	Open
(ii) 曝露因子の特性	(多くが) time-invariant	(多くが) time-variant
(iii) アウトカムの特性	(多くが) 死亡	様々なアウトカム



7

記述的 vs. 分析的

記述的な研究

実態を把握したい
- 疾病の負担
- 地域格差
- 時代による変化

分析的な研究

因果を追求したい



8

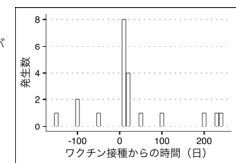
本日の内容

1. イントロダクション
2. SCCS研究とは
3. 古典的な研究デザインとの比較
4. SCCSの必要条件
5. よく受ける質問

9

なぜSCCSが生まれたか？

- 1988年・イギリスでは、浦部株を含むMMR (麻疹・おたふくかぜ・風疹) ワクチンが導入された。
- もともとこのワクチンは無菌性髄膜炎を起こす可能性が指摘されていた。
- 無菌性髄膜炎になった児を調べると、明らかにワクチン接種「後」に発症が集積していた。

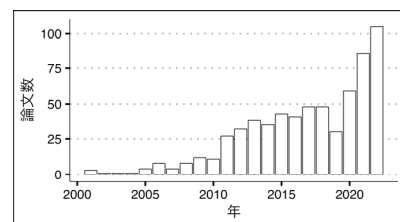


1) Farrington P et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 10

- 時間的關係から因果關係はありそう。
- コホート研究は、イベント発症頻度が低すぎるため実行困難。
- ケースコントロール研究は、コントロールを選ぶ源集団の設定が困難。
- なんとかして、ケース (無菌性髄膜炎患者) のみから、リスク比やハザード比のような統計量を出すことができないか？
- 開発された手法がSCCS。

11

最近の論文数の推移



self-controlled case series [Title/Abstract]で検索 (2022/11/17, PubMed)

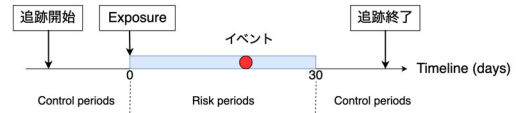
12

SCCSに必要な情報

- ケースのみのデータが良い
 1. 追跡開始日
 2. 追跡終了日
 3. Exposureを受けた日
 4. Event発生日
- Timelineはageで考えるのが一般的（日単位）。
- 観察開始日の年齢が20歳なら、 $365.25 \times 20 = 7305$ （日）
- イベント発生日が20歳2ヶ月なら、 $7305 + 30.5 \times 2 = 7366$ （日）

13

SCCSで求められるもの



- Control periodsにおける発生率を基準とした・risk periodsにおける発生率の比が求められる（発生率比・incidence rate ratio [IRR]）。
- 条件付きポアソン回帰が用いられる。

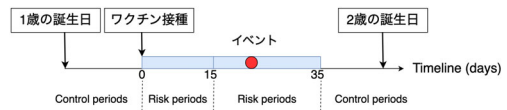
14

SCCSのメリット・デメリット

- 同一人内の比較のため・時間で変化しない因子（性別・遺伝子・生活習慣・社会経済的背景）は自然に相殺される
- ケースのみを用いた手法なので・コントロールを選んでくる必要がない
→ コントロール選択に伴う選択バイアスが発生しない
- 絶対的指標（発生率）を求めることはできない
→ コホート研究を同時に行い・発生率を提示すると説得力が増す。

15

MMRワクチンと熱性けいれん・無菌性髄膜炎



Control periodと比べたrisk periodの発生率比 (IRR)は・
6-11日後 3.04 (2.27-4.07)
15-35日後 1.51 (1.21-1.90)

Farrington P, et al. Lancet. 1995.

16

本日の内容

1. イントロダクション
2. SCCS研究とは
3. 古典的な研究デザインとの比較
4. SCCSの必要条件
5. よく受ける質問

17

SCCSとコホート研究の比較 ①

—MMRワクチンと熱性けいれん¹⁾

	サンプルサイズ	Risk period	Relative incidence
コホート研究 ²⁾	679942人 (487人event発生)	8-14 days	2.83 (1.44-5.55)
SCCS ³⁾	952人 (全員event発生)	6-11 days	3.04 (2.27-4.07)

- 点推定値は両者ではほぼ同じ。
- SCCSの方がイベント数が多いため信頼区間が狭い。
- SCCSの方が全体のサンプル数は小さく・研究遂行が容易。

1) Farrington P, et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018.
2) Barlow WE, et al. N Engl J Med. 2001.
3) Farrington P, et al. Lancet. 1995.

18

SCCSとコホート研究の比較 ②

-MMRワクチンと自閉症¹⁾

	サンプルサイズ	Risk period	Relative incidence
コホート ²⁾	537303人 (316件event発生)	Indefinite	0.92 (0.68-1.24)
SCCS ³⁾	357人 (全員event発生)	Indefinite	0.88 (0.40-1.95)

- 点推定値は両者でほぼ同じだが・SCCSの方が信頼区間が広い・
- Risk periodがindefiniteの時は・SCCSは検出力低くなる・
- SCCSの方が全体のサンプルサイズが小さく研究遂行が容易・

1) Farrington P et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018.
2) Madsen KM, et al. N Engl J Med. 2002.
3) Farrington P, et al. Journal of the Royal Statistical Society: Series C. 2006.

SCCSとコホート研究の比較 ③

-コロナウイルス内服中患者における・プロトンポンプ阻害薬 (PPI) と心筋梗塞¹⁾

	サンプルサイズ	Risk period	Relative incidence
コホート	29025人 (734件event発生)	PPI内服期間	1.30 (1.12-1.50)
SCCS	774件event発生	PPI内服期間	0.75 (0.55-1.01)

- 点推定値は両者で異なる方向・
- コホート研究では残差交絡が起きていた可能性がある・

1) Douglas LJ, et al. BMJ. 2012. 20

SCCSとコホート研究の比較 ④

-コロナワクチン接種と眼有害事象 (ぶどう膜炎・視神経炎など)¹⁾

	サンプルサイズ	Risk period	Relative incidence
コホート	99718人 (107件event発生)	接種後3ヶ月	1.8 (1.2-2.9)
SCCS	115件event発生	接種後3ヶ月	0.9 (0.7-1.1)

- 複数回のexposure (ワクチン1回目と2回目) を区別した解析
- 点推定値は両者で異なる方向・
- コホート研究では残差交絡が起きていた可能性がある・
(肥満・血圧値の情報が不明)

1) Hashimoto Y, et al. Ophthalmology. 2022. 21

本日の内容

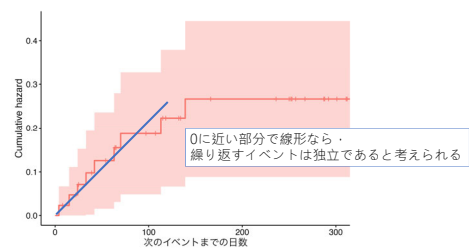
1. イントロダクション
2. SCCS研究とは
3. 古典的な研究デザインとの比較
4. SCCSの必要条件
5. よく受ける質問

条件①： 反復するeventの場合は・互いに独立している
反復しないeventの場合は・その確率がレアである

- 心筋梗塞は1回発生すると・2回目以降も起きやすいはず・
- 初回的心筋梗塞のみを解析に含めた解析を行う¹⁾・
- 2回目以降の心筋梗塞も含めた解析とほぼ同じ結果なら・robustと言える・
- 心筋梗塞の発生頻度はレアであるという仮定をおいている・

1) Minassian C, et al. Annals of Internal Medicine. 2010. 23

Cumulative hazardでイベントが独立であることを確認



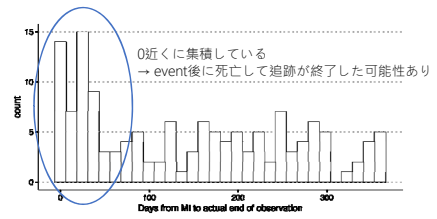
Farrington P, et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 24

条件②：Event発生は後の追跡期間に影響を及ぼさない

- 短期死亡率が高いeventで問題となる・
- eventが心筋梗塞の場合・その後早期に死亡してしまう確率が高く・後の追跡期間に影響を及ぼす（極端な例は・event = 死亡）・

25

Eventから追跡終了までの日数のヒストグラム



Farrington P. et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 25

Terminal risk periodの導入

- I. baseとなるSCCSモデル
- II. 追跡終了～30日前をterminal risk periodとして入れたモデル
- モデル I・IIを尤度比検定し有意となれば・eventが追跡期間に影響を及ぼしている可能性が高い・
- Eventが後の追跡期間に影響を及ぼしていると考えられる際は・RのSCCS package内の eventdependobs()関数で対応可能・

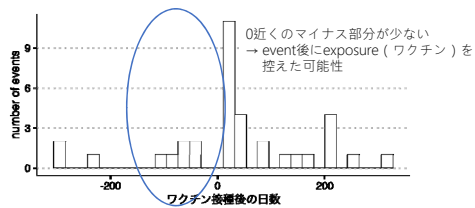
Farrington P. et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 27

条件③：Event発生は後のexposureに影響を及ぼさない

- 薬剤の副作用を検討する場合が代表的・
- ワクチンとけいれんの関連を見たい場合・けいれんを起こした患者へのワクチン投与は控えられため・後のexposureに影響を及ぼす・

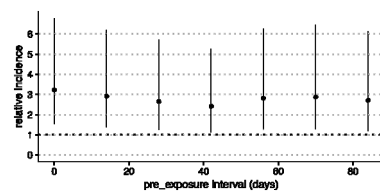
28

Exposureからeventまでの日数のヒストグラム



Farrington P. et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 29

Pre-exposure periodをモデルへ組みこむ



Pre-exposure periodを動かしても点推定値は大きく変化しない→ robust
Robustではない場合は・RのSCCS package内のeventdependexp()関数で対応可能・

Farrington P. et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 30

条件④：Exposureは後のevent捕捉に影響を及ぼさない

- 薬剤有害事象の自己申告型調査が代表例・
- 薬剤A内服後の下痢をeventとする場合・はじめから薬剤Aと下痢の関連が疑われたケース（内服後早期の下痢）のみが含まれる・SCCSの結果・有意に関連がでるのは当然である・
- ケースコントロール研究と同様に・源集団を明確に意識する・（薬剤Aとの時間的關係によらず下痢患者を集める）

31

- 4つの条件全てを満たすのは困難である・
- コホート研究では未測定交絡因子がないことを前提にしているが・それも現実的には無理な話である・
- コホート研究・ケースコントロール研究といった古典的な手法に加え・SCCS結果も提示するのが好ましい・
- 両者は異なるassumptionの上に成立している
→ 結果が一致すれば・robustと言える
→ 結果が一致しなければ・より深く結果を考察することが可能となる

32

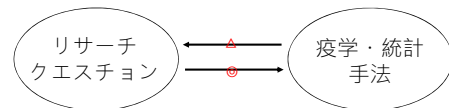
本日の内容

1. イントロダクション
2. SCCS研究とは
3. 古典的な研究デザインとの比較
4. SCCSの必要条件
5. よく受ける質問

33

Q1: SCCSを使ってみたい・どんなリサーチクエスチョンが良いか？

- 「SCCSを使う」ことを目的にするのは・おすすめできません・
- まずはリサーチクエスチョンを設定し・曝露とアウトカムの特性がSCCSを使用するための条件を満たしている場合・是非トライしてみてください・

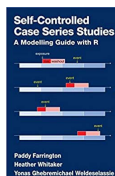


34

Q2: 自己対照研究デザインの統計コマンドを教えてください・

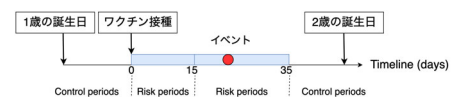
- 下記ウェブサイトにR・STATA・SASの練習用データセットとコマンド例（データセット作成から条件付きポアソン回帰分析まで）が載っています・
<https://sccs-studies.info/about-us.html>

- Rユーザーは右の本がおすすめです・
(Farrington P, et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018.)



35

Q3: サンプルサイズはどれくらい必要か？



```

samplesize( eexpo=2.5, # 発生率比
            risk=21,   # Risk period (ワクチン接種後15-35日)
            astart=366, # 観察開始日
            aend=730,  # 観察終了日
            p=1)       # 集団の中でexposureを受けた割合
    
```

→ 110例のイベント数が必要

Farrington P, et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018. 36

Q4: 解析に用いる症例は・eventを生じ・かつexposureを受けた人の情報のみでよいか (exposureを受けていない人は除くか) ?

- 時間依存性交絡因子を調整しない場合には・exposureを受けた人のみに絞っても良い・
- 時間依存性交絡因子を調整する場合は・exposureを受けた人も受けなかった人も全員解析対象とすべき (効率が上がる)¹⁾・
- 計画的にexposureを受けた人のみに絞った例もある²⁾

1) Farrington P, et al. Self-Controlled Case Series Studies: A Modelling Guide with R. 2018.
2) Kwong JC, et al. *New England Journal of Medicine*. 2016.

サンプルサイズ計算 (Exposureの割合を50%にしてみる)

```
samplesize( eexpo=2.5, # 発生率比
            risk=21,   # Risk period (ワクチン接種後15-35日)
            astart=366, # 観察開始日
            aend=730,  # 観察終了日
            p= 0.5)    # 集団の中でexposureを受けた割合
```

→ 210例のイベント数が必要

- $p = 1.0$ → 必要なサンプルサイズは110例だった・
- $P = 0.5$ → 必要なサンプルサイズは210例 (約2倍) に増加・
- Exposureなしの人の情報量は乏しい (時間依存性交絡因子を考慮しない場合)

本日のまとめ

- SCCSは・eventが起きた人のみを対象に・その人の過去および未来の期間と比較したりする手法である・
- コホート研究よりも効率が良い・
- 時間非依存性交絡因子は自動的にキャンセルアウトされる・
- 発生率比を求めることで・因果推論に絶大な効果を発揮する・
- 絶対リスクを求めることはできない・
- 4つのassumptionの上に成立していることを意識する・
- コホート/ケースコントロール研究と併行して行うことで説得力が増す・

39

機械学習

東京大学大学院医学系研究科糖尿病・生活習慣予防講座
岡田 啓

本日の内容の対象者

- 機械学習を学びたいと思っているが
ハードルが高く、導入の本すら読めずにいる人
- 機械学習に興味があるけれども
どう論文作成に活かせばよいかわからない人

本日の目標

- ① 機械学習で用いられる方法・用語を説明出来る
Lasso回帰、決定木、ランダムフォレスト、XGBoost
用語：交差検証法、正則化、過学習
- ② 予測と推論で、機械学習を使い分け、論文を読み書き出来るようになる

本日の内容

1. 機械学習の導入 (Lasso回帰を中心に)
2. 機械学習を用いた論文の紹介と使われ方
3. ほかの機械学習 (決定木 → ランダムフォレスト/XGBoost)

はじめに

基本的な用語説明

機械学習とは

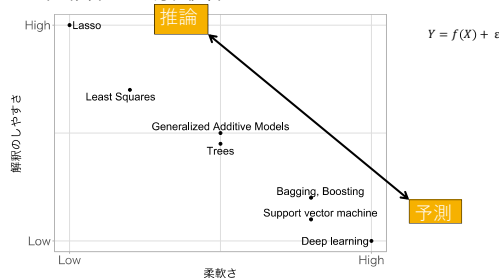
- 大量のデータから法則性を見つけ出す一連の流れのこと
- 説明変数を X ・ 応答変数を Y ・ ランダム誤差を ε とする
- X と Y の間に何らかの関係があると仮定する
$$Y = f(X) + \varepsilon$$
- f を推定するのが機械学習

予測 (prediction) と推論 (inference)

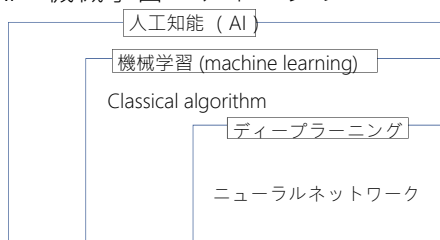
$$Y = f(X) + \epsilon$$

- 予測
 - X から Y を予測することが最重要・ f の形はブラックボックスで良い・
- 推論
 - X と Y の関係を知ることが最重要・ X が変化すると Y はどのように変化するのか知りたい・ f の形はブラックボックスというわけにはいかない・

柔軟性と解釈性のトレードオフ



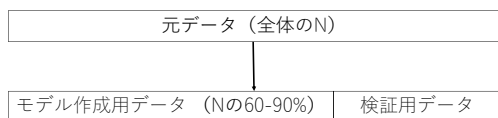
用語の説明 AI・機械学習・ディープラーニング



Introduction

多変量回帰からみた、Lasso回帰について

用語の確認



一般的に予測モデルでは、モデル作成用データと検証用データが違うデータセットの方が良いが、難しい場合、元データをモデル作成用と検証用に分けるのが一般的である。

身近な機械学習① 多変量回帰 ~ 一般線形モデル ~

①最小二乗法

$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ というデータがあるとして

$$\text{Predicted } y_i = F(x) = ax_i + b$$

とすると
実測値とのずれの和

$$J = \sum_{i=1}^n (y_i - f(x_i))^2$$

が最小となる a と b を求める

身近な機械学習② 多変量回帰 ～ 一般化線形モデル ～

②link関数を用いたモデルの作成例：ロジスティック回帰

リンク関数はlogitなので、

$$\text{logit}(y) = \frac{\ln(y)}{1 - \ln(y)} = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \varepsilon$$

実測値

$$\left\{ \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \end{pmatrix}, \begin{pmatrix} x_{21} \\ \vdots \\ x_{2n} \end{pmatrix}, \dots, \begin{pmatrix} x_{N1} \\ \vdots \\ x_{Nn} \end{pmatrix} \right\}$$

①②を組み合わせると

$$J = \sum_{i=1}^N \left(y_i - \sum_{j=1}^n x_{ij} \beta_j \right)^2$$

が最小となるように計算する

予測(\hat{y})を考える

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

これは計算出来そうだし、予測出来そう

では1000個のパラメータを使う場合は？

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{100} x_{100} + \beta_{101} x_{101} + \dots + \beta_{1000} x_{1000}$$

つまり、 y を1000個のパラメータで予測する

臨床疫学研究としての問題は？

①予測能は高いが、推論が難しい

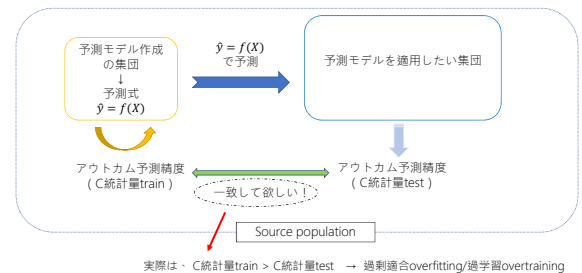
②統計学的な問題も・・・

多変量回帰での主な問題

- 過剰適合overfitting/過学習overtraining/optimism
提供データ（訓練データ）に対して学習されているが
未知データ（テストデータ）に対しては適合できていない状態
← 過多の回帰変数、過多の自由度など

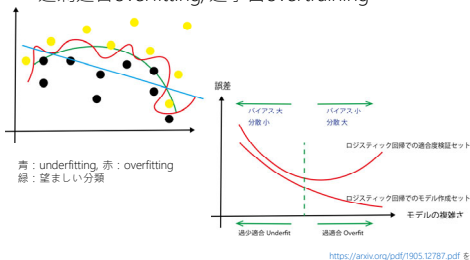
(・ 複数の多重共線性multiple collinearity)

予測モデルのoptimism



過適合の概念

- 過剰適合overfitting/過学習overtraining



過適合への対応

対策：

①交差検証Cross-validation

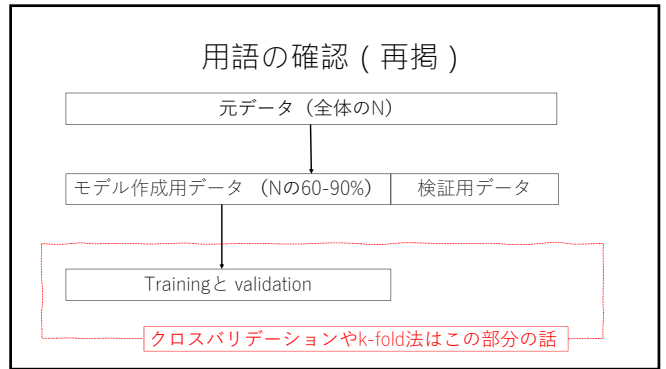
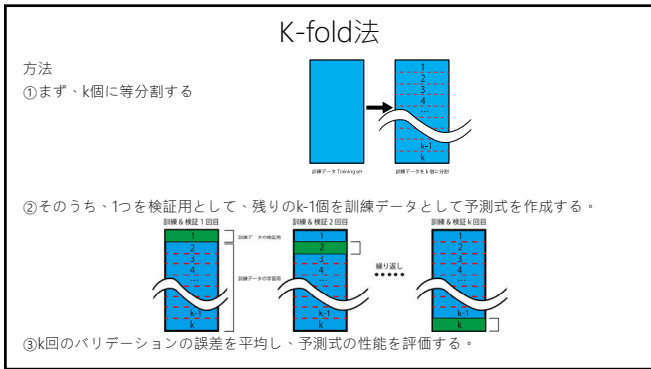
訓練データ内でもサンプリングとテストセットを作り、それを何度も繰り返して適合度を確認する。例：k-fold法など（次スライド）

→ 外的妥当性generalizabilityを高める

②正則化Regularization (→ 縮小推定shrinkage)

誤差関数に正則化項を追加して、モデルの複雑度・自由度を抑制し、変数を減らす。

→ わかりやすさを重視する (より「推論」に寄せる)



罰則項の追加による正則化

回帰で最小にしたい値

$$J = \sum_{i=1}^N \left(y_i - \sum_{j=1}^n x_{ij} \beta_j \right)^2$$

に罰則項

$$\lambda \sum_{j=1}^n |\beta_j|$$

を加えた

$$\sum_{i=1}^N \left(y_i - \sum_{j=1}^n x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^n |\beta_j|$$

を最小にするようにβを推定する。λは値を振ってみて適切な値を決める。

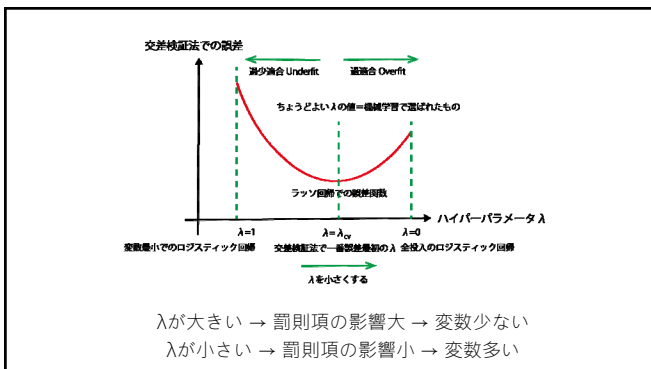
絶対値のみ：Lasso回帰
二乗のみ：Ridge回帰
その両方：Elastic net

Lasso回帰で行われていること

$$\sum_{i=1}^N \left(y_i - \sum_{j=1}^n x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^n |\beta_j|$$

- ①ハイパーパラメータλの値を100個くらい振る
- ②k-fold法での誤差関数を値を計算
- ③誤差関数の最も小さいところのλを採用 (交差検証法で得られたλ)

Cf. Hyperparameterとは、自動計算で出されたparameterと違い、自分で設定するパラメータなのでhyper- + parameter



小括

1. 推論/予測
2. Lasso回帰のキーワード：過学習、K-fold法、罰則項

Lasso回帰論文例：シンプルなモデル

Diabetes Care

1



A Machine Learning-Based Predictive Model to Identify Patients Who Failed to Attend a Follow-up Visit for Diabetes Care After Recommendations From a National Screening Program

<https://doi.org/10.2337/1841>

Akira Okada,¹ Yohei Hashimoto,^{2,3}
Tadahiro Goto,^{2,4} Satoko Yamaguchi,²
Sachiko Ono,⁵ Kayo Ikeda Kurakawa,¹
Masayumi Nangaku,⁶
Toshinasa Yamaguchi,⁷
Hideo Yasunaga,⁸ and
Takashi Kadowaki^{1,7,8}



背景

- ①糖尿病スクリーニング後の受診率は低い
- ②受診中断の因子は数々の研究があるが、受診勧奨後未受診の因子などを検討した論文は殆ど無い

既存研究では、重要と思われる13因子で回帰
→ BMI、HbA1c値、蛋白尿、降圧剤処方、脂質異常症薬処方、抗うつ剤処方などが、未受診に関連していた。

- ③機械学習などの新たな手法の出現の変化もあり、より「良く」「効率的に」糖尿病受診勧奨後未受診を予測すれば政策立案に役立つ可能性がある

- ④目的：機械学習を用いて、受診勧奨後の未受診予測モデルの構築を試みた

方法

「既存ロジスティックモデル（Diabetes Res Clin Pract 2014;105:176-184）」

VS

「既知の糖尿病関連マーカーなどを含めたものからlasso回帰にて予測因子を選択したモデル」

モデル作成・評価方法：

- ①trainデータ（80%）とtestデータ（20%）に分割
- ②交差検証法により、trainデータで予測モデル作成（lasso回帰の1SE ruleを適用）
- ③testデータでを用いてc統計量（ROC曲線のグラフ下面積：AUC）で予測能の評価（Delongテストで検定）を行う

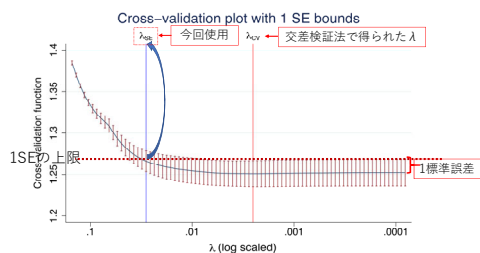
PECOによる概要

P: 過去12ヶ月に糖尿病関連レセプトがなく、健診でHbA1c $\geq 6.5\%$ and 空腹時血糖値 ≥ 126 mg/dLを満たした患者

E/C: 受診勧奨後受診に関わりそうな変数からlasso回帰で選択された変数の曝露有無

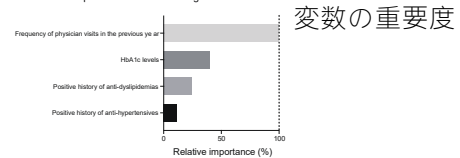
O: 健診後6ヶ月以内の糖尿病関連レセプトの発生

ハイパーパラメータλの設定



(A)

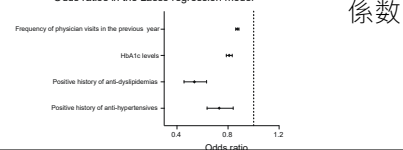
Variable importance in the Lasso regression model



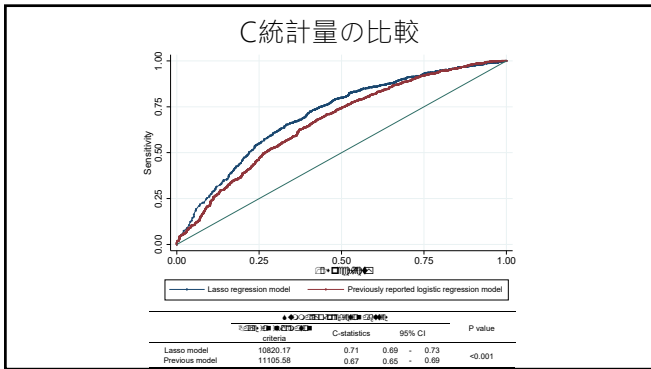
変数の重要度

(B)

Odds ratios in the Lasso regression model



係数



結論

既存モデル(13因子)のモデルよりも、機械学習で選ばれた4因子のモデルの方が受診勧奨後未受診をより正確に予測

機械学習を用いたメリット

よりわかりやすい(少ない変数で)モデルを作成するために、Lasso回帰を用いた。特に、変数を減らすために、ISE ruleを用いた。推論に特に有用である。

論文例② : スコア作成+ほかの機械学習

Surgery 171 (2022) 1036–1042

Surgery
journal homepage: www.elsevier.com/locate/surg

New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using a Japanese nationwide inpatient database

Takaaki Konishi, MD^{a,b,c}, Tadahiro Goto, MD, MPH, PhD^{b,c}, Michimasa Fujigoi, MD^{b,d}, Nobuaki Michihata, MD, MPH, PhD^e, Ryosuke Kumazawa, MPharm^b, Hiroki Matsui, MPH^b, Kiyohide Fushimi, MD, PhD^f, Masahiko Tanabe, MD, PhD^f, Yasuyuki Seto, MD, PhD^{a,g}, Hideo Yasunaga, MD, PhD^h

背景と目的

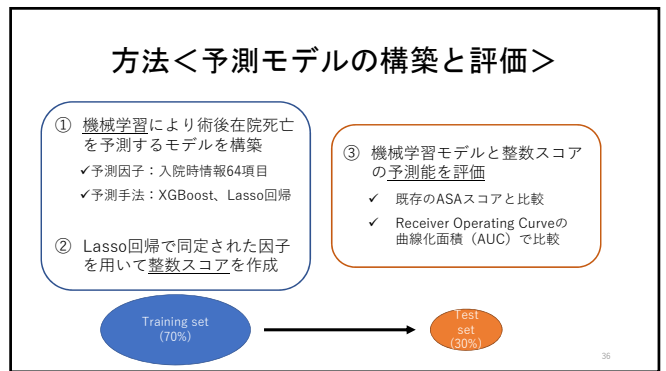
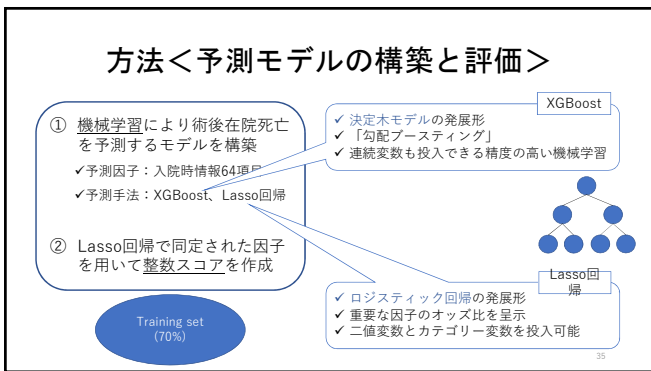
- 胃十二指腸潰瘍穿孔は一般的な外科的救急疾患である
- 術後死亡予測としてASA (American Society of Anesthesiology) スコア等が知られる
- しかし、それらの予測能は不十分である

Anbalakan K et al. Int J Surg 2015; 14: 38–44.

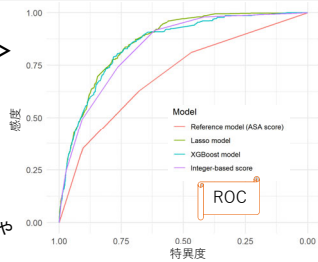
↓

本研究では、DPCデータベースを用いて下記を行うことを目的とする

- ① 機械学習による予測モデルの構築
- ② 臨床現場で使いやすい整数スコアの作成
- ③ ASAスコアと比較した①②の予測能の評価



結果<予測能>



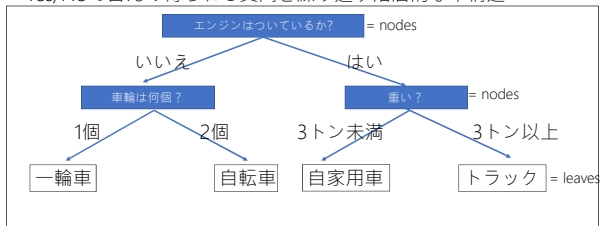
ASAスコアより機械学習モデルや整数スコアの方が予測能が高い

		C統計量 (95%信頼区間)	正確度 (95%信頼区間)	感度	特異度	陽性的中率	陰性的中率	P値
既存	ASAスコア	0.70 (0.67-0.73)	0.68 (0.67-0.69)	0.63	0.68	0.09	0.97	基準
今回作成	Lasso回帰	0.86 (0.85-0.88)	0.76 (0.75-0.77)	0.80	0.76	0.14	0.99	<0.001
	XGBoost	0.85 (0.83-0.87)	0.76 (0.75-0.76)	0.81	0.75	0.13	0.99	<0.001
	整数スコア	0.94 (0.82-0.86)	0.63 (0.62-0.64)	0.92	0.62	0.10	0.99	<0.001

決定木とその発展版 XGBoost、ランダムフォレスト

決定木 (decision tree)

- Yes/Noで答えの得られる質問を繰り返す階層的な木構造



決定木の特徴

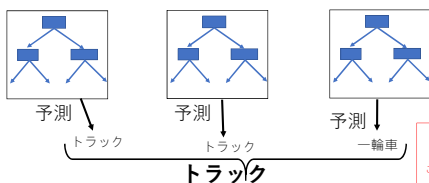
- メリット
- 分類・回帰とも利用できる
 - 説明変数はカテゴリカル変数・連続変数とも対応可
 - 得られた結果の意味を解釈しやすい

- デメリット
- 外れ値やノイズに弱い (= overfittingしやすい)

ランダムフォレスト

- 決定木を複数作り・多数決を取ることでoverfittingを防ぐ

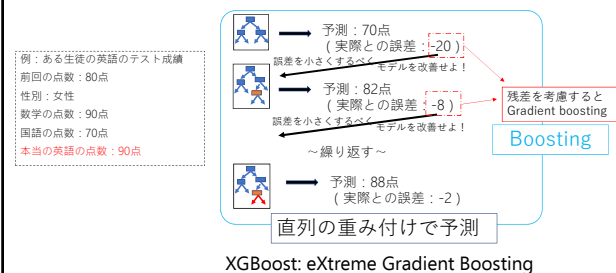
- サンプル数nのデータから、n個のサンプルを復元抽出する
- このデータセットを用いて決定木を作成する (いくつかの説明変数を選択する = 捨てる説明変数が発生する)
- k回繰り返す
- k個の決定木から、多数決によりクラスを予測する



多数決により、決定木の弱点のoverfittingを防ぐ
この多数決のことを「バギング」と呼ぶ

XGBoostなどのBoosting

- 最初の決定木から誤差を求め、より誤差の少ない (精度のよい) 決定木に、徐々に洗練させていき、重みつけて予測する方法



本日のまとめ (目標の再掲)

①機械学習で用いられる方法・用語を説明出来る

Lasso回帰、決定木、ランダムフォレスト、XGBoost

用語：交差検証法、正則化、過学習

②予測と推論で、機械学習を使い分け、論文を読み書き出来るようになる

データベース研究における バリデーション研究

自治医科大学データサイエンスセンター
山名隼人

Agenda

- バリデーション研究とは
なぜバリデーション研究が必要か
- バリデーション研究の活用例
- バリデーション研究の例と実際
- 現状と課題

大規模データベース研究の特徴

保健医療介護ビッグデータ

- NDB、DPC、介護レセプトデータ
- JMDC、MDV
など

利点

- データの量、N数
- 全国規模、集団代表性
- データ取得が(比較的)容易
- 実際の臨床現場を反映(リアルワールド)

大規模データベース研究の特徴

研究の特徴

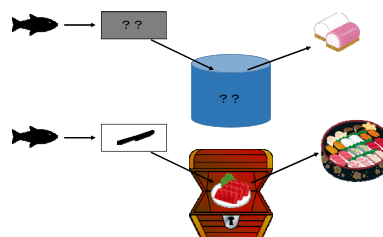
- 後ろ向き研究 ↔ 前向きにデータ取得
- 観察研究 ↔ RCTなどの介入研究
→ 研究デザイン・統計手法を駆使、
観察研究から重要なエビデンスを生み出す
- 既存の情報を二次利用
→ 情報の「質」はどうか？

情報の質

情報の妥当性(Validity)

- 査読者コメントで非常に多いもの：“Is the data validated?”
- データベースに格納されている情報と、研究者がイメージする情報の差
- 例：傷病名
 - DPCデータの主病名・併存症名、“レセプト病名”
 - 緊急入院の場合の併存症名
- 100%の正確性は現実的ではないが、定量化することでデータを活用する研究の質を向上させる

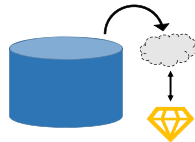
情報の質



バリデーション研究

データベースに記録された情報の妥当性を検証する

- データベースのうちの一部について、より正確な別の情報と比較する (至適基準、reference standard)
- データベース全体の妥当性を担保する



妥当性の指標

感度・特異度

至適基準 (reference standard) との比較が必要

例: 「脳梗塞」の病名

- 至適基準 = カルテ上の病名
- 検証したいもの = データベース上の病名の場合、
- 感度: カルテ上、脳梗塞がある人のうち、データベース上の病名でも脳梗塞ありとされる人の割合
- 特異度: カルテ上、脳梗塞がない人のうち、データベース上の病名でも脳梗塞なしとされる人の割合

妥当性の指標

感度・特異度・陽性的中率・陰性的中率

		カルテ上の脳梗塞		
		あり	なし	
データベース上の脳梗塞	あり	真の陽性 a	偽陰性 b	陽性的中率 =a/(a+b)
	なし	偽陰性 c	真の陰性 d	陰性的中率 =d/(c+d)
		感度 =a/(a+c)	特異度 =d/(b+d)	

- 感度90%, 特異度90%, 実際の割合5% → 陽性的中率は32%

妥当性の指標

感度・特異度

		カルテ上の脳梗塞		
		あり	なし	
データベース上の脳梗塞	あり	真の陽性	偽陰性	
	なし	偽陰性	真の陰性	

- 感度90%, 特異度90%, 実際の割合5% → 陽性的中率は32%

バリデーション研究の必要性

感度・特異度が研究結果に与える影響

- A薬 vs B薬でアウトカムを比較したい
- 実際:

A薬 100/2000 (5%)

B薬 200/2000 (10%)

< データベースにおけるアウトカム特定の感度50%・特異度80%の場合 >

A薬		真のアウトカム		
		あり	なし	
観測されるアウトカム	あり	50	380	430
	なし	50	1520	1570
		100	1900	2000

B薬		真のアウトカム		
		あり	なし	
観測されるアウトカム	あり	100	360	460
	なし	100	1440	1540
		200	1800	2000

バリデーション研究の必要性

感度・特異度が研究結果に与える影響

- A薬 vs B薬でアウトカムを比較したい
- 実際:

A薬 100/2000 (5%)

B薬 200/2000 (10%)

< データベースにおけるアウトカム特定の感度50%・特異度80%の場合 >

- 観測:

A薬 430/2000 (21.5%)

B薬 460/2000 (23%)

Real-World??

バリデーション研究の活用例 (1)

Angiotensin converting enzyme inhibitors and risk of lung cancer.

BMJ 2018;363:k4209

- Database UK Clinical Practice Research Datalink (CPRD)
- Patient 高血圧で治療開始
- Exposure アンジオテンシン変換酵素阻害薬
- Control アンジオテンシン受容体拮抗薬
- Outcome 肺癌の発症

バリデーション研究の活用例 (1)

Angiotensin converting enzyme inhibitors and risk of lung cancer.

BMJ 2018;363:k4209

- Method内の記載
“CPRDの記録は妥当であり質が高いことが示されている[引用]。さらに、CPRDにおける肺癌の診断名はUK National Cancer Data Repositoryの診断名とも高い一致率 (>93%) を示している[引用]。”

バリデーション研究の活用例 (2)

Use of haloperidol versus atypical antipsychotics and risk of in-hospital death in patients with acute myocardial infarction.

BMJ 2018;360:k1218

- Database 米国Premier Research Database
- Patient 急性心筋梗塞で入院 (+せん妄あり)
- Exposure ハロペリドール
- Control 非定型抗精神病薬
- Outcome 7日以内の在院死亡

バリデーション研究の活用例 (2)

Use of haloperidol versus atypical antipsychotics and risk of in-hospital death in patients with acute myocardial infarction.

BMJ 2018;360:k1218

- Method内の記載
“せん妄の病名が記録されていることは選択基準に含めなかった。せん妄があっても病名が記録されることが少なく、ICD-9コードによるせん妄の定義の感度は3%と低いことが示されているためである[引用]。代わりに、事前に精神疾患の病名がない患者において抗精神病薬が開始されたことをせん妄と定義した。”

バリデーション研究の種類

目的

- Target: 疾患・処置・薬剤・アウトカム
- 求める指標: 感度／特異度
陽性的中率／陰性的中率

方法

- カルテレビュー
- より信頼性が高いデータとの結合
例: 院内がん登録・症例データベース
- 他の情報からの“間接証拠”
例: 先行研究の発生率

バリデーション研究の例

Validity of diagnoses, procedures, and laboratory data in Japanese administrative data.

J Epidemiol 2017;27:476-82.

概要

- Target: DPCデータの疾患・処置
- 求める指標: 感度／特異度、陽性的中率／陰性的中率
- 方法: カルテレビュー
小規模 (N=315)

Methods (1)

対象

- 国立病院機構の4病院
- 2014年度入退院、入院時18歳以上、DPC入院
- 各施設100入院をランダムに抽出

内容

- 病名： Charlson Comorbidity Indexの基準となる17疾患
(主or入院時併存として有無)
- 処置： 10種類の処置 (入院日の実施有無)

Methods (2)

カルテレビュー方法

- 2名(看護師+医師)が独立に判断、相違があれば協議

データベースからの抽出方法

- 病名： 様式1 主傷病, 入院の契機となった傷病, 入院時併存病名×4 (疑い除く)
 - 処置： EFファイル 入院日の実施有無
- 解析**
- Validity： 感度・特異度、陽性・陰性的中率

Results (1)

病名の妥当性

疾患	カルテ上の 病名出現頻度 n	%	DPCデータの 感度 (%)	DPCデータの 特異度 (%)	DPCデータの 陽性的中率 (%)	DPCデータの 陰性的中率 (%)
心筋梗塞	23	7.3	52.2	99.7	92.3	96.4
うっ血性心不全	32	10.2	68.8	97.5	75.9	96.5
末梢血管疾患	29	9.2	34.5	99.3	83.3	93.7
脳血管疾患	38	12.1	50.0	98.9	86.4	93.5
糖尿病(合併症なし)	46	14.6	52.2	96.7	72.7	92.2
糖尿病(合併症あり)	17	5.4	29.4	99.7	83.3	96.1
悪性腫瘍	97	30.8	83.5	97.7	94.2	93.0

Results (2)

処置の妥当性

処置	カルテ上の 処置実施頻度 n	%	DPCデータの 感度 (%)	DPCデータの 特異度 (%)	DPCデータの 陽性的中率 (%)	DPCデータの 陰性的中率 (%)
尿検査	74	23.5	98.6	98.3	94.8	99.6
細菌鏡見	35	11.1	91.4	100	100	98.9
細菌培養	35	11.1	97.1	100	100	99.5
呼吸心拍監視	30	12.8	66.7	92.2	55.6	95.0
X線撮影	161	51.1	97.5	99.4	99.4	97.5
CT	93	29.5	100	99.5	98.9	100
尿路カテーテル挿入	29	9.2	65.5	97.2	70.4	96.5

バリデーション研究の実際 (1)

事前準備

- 症例数
10%の有病率、感度50%の場合：
200例では真の陽性は10人のみ。300~400例は必要
- カルテレビューにかかる時間・費用
10分×100例=16.7時間
- 病院への計画説明&協力依頼
- レビュー用シート作成

バリデーション研究の実際 (2)

カルテレビューの実施

- 日程
 - 連続する2日間で実施
 - 共同研究者・施設側との日程調整
- その他
 - 電子カルテの使用権限、環境の確保
 - 慣れない電子カルテの利用
 - 限られた時間 再訪問は困難
 - レビュー結果データのやりとり

バリデーション研究の実際 (3)

トラブル

- 予定していた症例全ては終わらなかった
Reviewerに“why?”と指摘され、論文に“due to time constraints”と追加
- ある病院では、鼻腔培養MRSAがほとんどの患者で実施されていた
- 別の病院では、過去の温度板が参照できなかった
電子カルテシステム更新の影響

バリデーション研究の課題 (1)

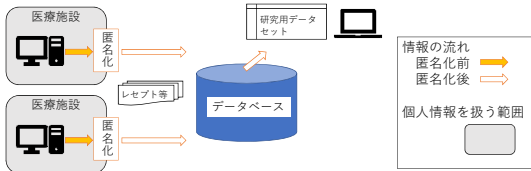
バリデーション研究は実施が難しい

- van Walraven et al. Administrative database research infrequently used validated diagnostic or procedural codes. *J Clin Epidemiol* 2011より：
- Measuring code accuracy can be time consuming and expensive.
 - Securing grant support for such a study is difficult because a code-validation study may hold little appeal for many granting agencies.
 - Finally, journals may decide to remove data regarding code validation in an effort to save space.

バリデーション研究の課題 (2)

データ収集と匿名化の壁

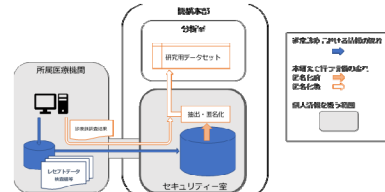
- 各施設から匿名化データを収集して活用する場合、カルテに戻ってのバリデーションは難しい



バリデーション研究の課題 (2)

データ収集と匿名化の壁

- 国立病院機構におけるバリデーション研究の場合



バリデーション研究の現状

日本のバリデーション研究は増加傾向

- 報告数
2017年11月～2022年3月：20編超の英語論文が発表
2018: 1 → 2019: 2 → 2020: 1 → 2021: 13
- 領域 : 循環器、がん、糖尿病が多い
- 規模 : 単施設・小規模が多い
- 至適基準 : カルテ調査が多いが、レジストリ等も活用

バリデーション研究の現状

例

筆頭著者、掲載誌、年	データ	対象	施設数	至適基準
Imai et al, J Med Virol 2019	医科 レセプト	B型肝炎ウイルスの既感染	4	カルテ調査
Shigemi et al, Cancer Epidemiol 2021	DPC データ	がん (15種類)	31	がん登録
Ono et al, BMC Health Serv Res 2021	歯科 レセプト	各種診断名、 歯科処置等	1	カルテ調査

バリデーション研究の現状

例

筆頭著者、掲載雑誌、年	データ	対象	施設数	至適基準
Yamana et al, Pharmacoepidemiol Drug Saf 2022	DPC データ	術後感染症	4	カルテ調査
Konishi et al, Surg Today 2022	DPC データ	手術情報 (術式等)	1	カルテ調査
Fujita et al, J Glaucoma (in press)	医科 レセプト	緑内障	1	カルテ調査

データベース研究におけるバリデーション研究： まとめ

- 保健医療介護ビッグデータの懸念：研究に必要な情報が正確に記録されていないかもしれない
- バリデーション研究：データベースのうちの一部について、より正確な別の情報（至適基準）と比較して情報の妥当性を検証する研究
- データベースを活用する研究の質を向上させることができる
- 日本においても増加傾向

分担研究名 NDB データ研究の実践と人材育成プログラム開発・検証

研究分担者

中山 健夫 京都大学医学研究科 健康情報学分野教授

研究協力者

加藤 源太 同医学部附属病院診療報酬センター
・病床運営管理部 部長／病院教授
植嶋 大晃 同国際高等教育院附属データ科学イノベーション教育
研究センター 特定講師
齋藤 良行 京都大学医学研究科 健康情報学分野 研究員
黒田 知宏 同医学部附属病院 医療情報企画部教授

要旨

京都大学では、2019 年度より文部科学省が管轄する「関西広域・医療データ人材育成拠点形成事業」を開始しており、主コースである修士課程追加履修コースのほかに、民間向けに展開する 2 つのインテンシブコースを併設している。この 2 つのコースは、講師やコンテンツに一部の重複はあるものの、それぞれに運用形式や目指す方向性には相違があり、参加者が習得を目指すスキルも異なっていて、1 期生、2 期生と回を重ねるごとに独自の展開を見せるようになってきている。本課題ではアカデミアにおける研究者養成とは異なる形式での保健医療介護ビッグデータ研究人材育成の可能性について報告する。

A. 目的

大学院課程における研究者養成とは異なる、企業の勤務者等社会人を対象とする保健医療介護ビッグデータ研究人材育成プログラムの可能性と課題を検討する。

B. 研究方法

京都大学は 2019 年度より文部科学省の助成を受け、関西広域の計 13 大学とコンソーシアムを構築して「関西広域・医療データ人材育成拠点形成事業 (Kansai Union / Kyoto University Education Program for Digital Health

Innovation: KUEP-DHI. 責任者: 黒田知宏・京都大学医学部附属病院医療情報企画部教授)」に取り組んでいる。本事業は医療現場から大規模に収集される多様なデータの利活用を推進し、質の高い医療を実現するため、医療データの活用基盤を構築・運営する人材、医療データを利活用できる人材を育成することを目指すものである。KUEP-DHI の根幹は、大学院生(主に修士学生)を対象とした「医療データ取扱専門家育成コース」であり、京都大学大学院医学研究科と同大大学院情報学研究科に追加履修コースとして設置されている。次世代医療基盤法や保健医療データプラットフォームに蓄積されるデータを活用し、医療データ

サイエンスの国際的牽引役を担えるよう本邦の医療データサイエンスを発展させるために、医療データが生まれてから活用されるまでの情報流の始点から終点までを確実に支え、正しく統制できる人材を育成することを目指している。

本研究は同事業の一環として実施されている2つの社会人向けインテンシブコースの運営経験に基づく事例検討である。

C. 研究結果

KUEP-DHI では大学院生という立場に依らず、企業の勤務者等、社会人を想定して次の2種類のインテンシブコースを提供している。

(1) DHIEP Program (ディープ プログラム. 責任者: 奥野恭史・京都大学大学院医学研究科ビッグデータ医科学教授) Data-driven Healthcare Innovation Evangelist Promotion Program 分析だけを請け負う従来型データサイエンティストではなく、ライフサイエンス・ヘルスケア領域に深い造詣を有し、社会変革を実現しうる人材育成を目指すプログラム。対象は医療データの利活用を志す個人であり、個人参加、登録等も個人で行う。出願資格は大卒資格を有する者で、ライフサイエンス・IT 関連企業に所属する者及び今後当該領域に参入を検討している者、医療従事者、地方自治体職員、規制当局関係者、大学院生など定員は20名。カリキュラムの修業期間は半年間であり、大きく経営・社会変革、医学・生命科学、情報・データサイエンスの3領域で基礎知識を e-learning で学び、必修科目 45 時間(プレゼンテーション・コミュニケーション、環境・市場理解、ビジネスモデリング、デザインシンキング、チェンジマ

ネジメント、医療情報倫理学、バイオインフォマティクス、可視化、AI・機械学習など)、選択科目 21 時間(ゲノム医療、IT 創薬、ライフサイエンスデータベース、時系列分析、テキストマイニング、ベイジアンネットワーク等)、実習(必修)20 時間(医療データから患者を層別化し、診断アプローチを提案、既存データベースを活用して分子標的探索を行い現状プロセスの改善点を提案、電子カルテの副作用情報を用いたドラッグリポジショニングの探索、ライフログと医療データを用いた生活の中からの医療シーズの発見等)で構成されている。修了要件は、合計 60 時間以上の履修と各科目評価での合格(修了要件を満たせば、履修証明書が授与される)。受講料は社会人が 50 万円(税抜)、大学官公庁等の職員と大学院生は無料である。

(2) 医療データ人材育成拠点形成事業 ビジネス特化型インテンシブコース Kansai Union / Kyoto University Education Program for Digital Health Innovation directly on themes of business (KUEP-DHI dot.b. 運営委員長: 中山健夫・京都大学大学院医学研究科社会健康医学系専攻健康情報学教授): 民間(企業)のニーズを踏まえつつ、医療データの活用に造詣のある人材を育成していくことを目的とするプログラム。対象は医療データ利活用に造詣のあるスタッフを確保したい企業であり、個人参加ではあるが、契約は企業-大学間で締結する。

育成する人材像は参加者が所定のプログラムを終えられた後には、各企業に戻り、医療データを企業内で有効かつ安全に利用する方針を立案したり、使用できる

データを適切に活用したりすることで有用な商品やビジネス等の開発に繋げ、社会に貢献できる人材を想定。到達目標は次の通りである。

- ・医療データを活用するための知識を系統立てて理解し、応用できる。
- ・医療データの活用に必要な現場感覚をもとに、医療データを分析して結果を解釈できる。

プログラムの内容は、dot.b 独自の講義、大学院の既存科目については科目等履修制度の活用、文献抄読、分析実習、臨床講義・実習、グループ演習で構成されている。分析実習では NDB オープンデータ等を用いた R、SQL 等の利用、グループ演習では、NDB 申出の練習や実際の NDB サンプルデータセットを用いた分析を実施している。また臨床講義・実習では、電子カルテ入力体験により、電子カルテの自習環境の操作、電子カルテのインターフェイスの理解、予め提示された模擬患者に対し、医師になったつもりでのカルテ記入等を通して、入力されるデータの背景とリアリティを実感できる機会を提供している。グループ演習では、EBM の PECO (patient, exposure, comparison, outcome) 形式を学び、構造化された研究計画を立ててデータを抽出し分析するという、データベース研究の一連のプロセスを提供している。dot.b プログラムは、参加者個人と契約を行うのではなく、まず本プログラムの趣旨に賛同する企業を募り、大学と有償の共同研究契約を締結する。そして各企業内で参加の該当者を選定し、プログラムに派遣する。参加者は京都大学と各企業の共同研究のプロジェクト研究員として、プログラム参加期間中は、京都大学での立場を持つことができ、京都大学の各種リソースを一

定の条件で利用できる。

コースは 2020 年度から開始され、2020 年度には 6 社・9 名、2021 年度は 3 社・4 名、2022 年度は 7 社・8 名が参加した。また、各年度の参加者が一同に会して NDB を分析する演習を行い、コース修了者との関係を維持し、参加者間の繋がりを深化させる機会を提供している。

D. 考察 & E. 結論

「関西広域・医療データ人材育成拠点形成事業」における 2 種類の社会人向けインテンシブコースの経験から、データサイエンスの人材育成に関するいくつかの知見を得た。下記に抜粋して記述する。

- ・医療者でない参加者の場合、医療データ以前の「医療そのもの」への距離感が大きいと、プログラム提供に当たっては、そのギャップを十分意識することが必要。
- ・NDB 等の実際のデータに触れることは学習者のモチベーションを大きく高め得る。
- ・「公益性のある研究」として NDB 申出書を書くことの意義を体感することは大いに重要。
- ・医師の診断プロセスや、電子カルテ入力体験を通じた、「データが絶対とは限らない」「バリデーションが重要」という感覚の共有はきわめて重要。
- ・レセプトでの傷病名同定における投薬や処置の情報を理解することの重要性。
- ・医療の場で入力されている情報が必ずしも正確かつ網羅的とは言えないことの実感。
- ・模擬的なデータ分析体験を通じた、臨床研究への意識の高まり。
- ・「企業が求める人材育成」と「大学教育が想定する人材育成」の違いの認識。

・受講者ニーズと提供するコンテンツとのマッチングの重要性。

本プログラムは 2022 年度も実施しており、引き続き、より良いデータサイエンス人材育成の在り方、方向性を探っていきたい。

F. 健康危機情報
なし

G. 研究発表(R4 年度分更新済)

中山健夫. 京都大学における医療データ利活用人材育成プログラム. 日本臨床疫学会第 5 回年次学術大会康永秀生理事企画シンポジウム「保健医療介護ビッグデータ研究
人材育成の取り組み」 早稲田大学
2022.11.13

1: Tsuji N, Takahashi Y, Sakai M, Ohtera S, Kaneyama J, Cho K, Kato G, Ohtsuru S, Nakayama T. Trend of anticoagulant therapy in elderly patients with atrial fibrillation considering risks of cerebral infarction and bleeding. *Sci Rep.* 2023 Jan 5;13(1):192.

2: Ishizaki M, Toyama M, Imura H, Takahashi Y, Nakayama T. Tracheostomy decannulation rates in Japan: a retrospective cohort study using a claims database. *Sci Rep.* 2022 Nov 17;12(1):19801.

3: Sakai F, Hirata K, Igarashi H, Takeshima T, Nakayama T, Sano H, Kondo H, Shibasaki Y, Koga N. A study to investigate the prevalence of

headache disorders and migraine among people registered in a health insurance association in Japan. *J Headache Pain.* 2022 Jun 23;23(1):70. doi: 10.1186/s10194-022-01439-3. Erratum in: *J Headache Pain.* 2022 Dec 28;23(1):164.

4: Hoshino N, Hida K, Fukui Y, Takahashi Y, Nakayama T, Obama K. Relationship between diverting stoma and adjuvant chemotherapy in patients with rectal cancer: a nationwide study using the National Database of Health Insurance Claims and Specific Health Checkups of Japan. *Int J Clin Oncol.* 2022 Mar;27(3):545-552.

5: Sengoku T, Ishizaki T, Goto Y, Iwao T, Ohtera S, Sakai M, Kato G, Nakayama T, Takahashi Y. Prevalence of type 2 diabetes by age, sex and geographical area among two million public assistance recipients in Japan: a cross-sectional study using a nationally representative claims database. *J Epidemiol Community Health.* 2022 Apr;76(4):391-397.

H. 知的所有権の取得状況
なし

令和4年度厚生労働行政推進調査事業費補助金
(政策科学総合研究事業(政策科学推進研究事業))
分担研究報告書

医療・介護データ活用による研究・人材育成

研究分担者 田宮菜奈子 筑波大学 医学医療系 ヘルスサービスリサーチ分野 教授
研究協力者 岩上将夫 筑波大学 医学医療系 ヘルスサービスリサーチ分野 准教授

研究要旨

本研究分担班では、昨年度に引き続き、本年度も医療・介護保険レセプトデータや Diagnosis Procedure Combination (DPC)データなどの医療ビッグデータを用いた研究を推し進め、特に博士課程・研修医・若手教員を対象にした人材育成に一層力を入れた。

近年、地域医療や地域包括ケアシステムの教育の必要性が明記されるなど、医学教育においても病院内での急性期医療のみでなく、退院後の社会的課題を含めた長期的視点が重要とされている。一方、2020年から新型コロナウイルスのパンデミックにより、リアルワールドデータを素早く解析して社会・世間の疑問に対して答えることの重要性が増している。このような中、本研究班により、その取り組みが加速でき、国際誌掲載や学会発表などの成果をあげることができた。

論文発表の成果としては、当教室(筑波大学医学医療系ヘルスサービスリサーチ分野)の博士課程生がある自治体の医療・介護保険レセプト連結データを解析し、在宅療養支援診療所・病院による緊急往診、入院、看取り医療に対する影響の評価を行った(Sun Y, BMC Health Serv Res. 2023;23(1):115.)。また当大学病院の研修医がMedical Data Visions社から提供を受けたDPCデータを解析し、新型コロナウイルス入院患者におけるデキサメタゾンとステロイドパルス療法の効果と有害事象の比較を行った(Watanabe A, et al. Crit Care Explor. 2023;5(4):e0886.)。また同データを用いて当教室の若手教員(宇田助教)が2022年度第5回臨床疫学会にて高齢COVID-19入院患者における入院早期リハビリテーションと在院死亡との関連についての解析結果を発表し、優秀賞を受賞した(宇田, et al. 宇田 和晃

また、同学会にて、本研究分担者である田宮菜奈子がシンポジウム(リアルワールドデータ研究の人材育成)で講演を行い、また本研究協力者である岩上将夫がシンポジウム(臨床疫学研究に使われるデータベース:電子カルテデータベースとその研究利用)で臨床疫学の初中級者を対象に講義を行った。

このようにリアルワールドデータを用いた臨床医学・社会医学研究に関心を持つ医療者や若手教員を対象に医学教育を行い、研究者の裾野を広げ、本事業の目標である幅広い医療ビッグデータ研究にかかわる人材育成に貢献することができた。

はじめに

昨今、WHOでWorld Health Assembly in May 2016においてFramework on integrated people-centred health services (IPCHS) が議論され、また医学教育モデル・コア・カリキュラム（平成 28 年度改訂版）においても、「地域医療や地域包括ケアシステムの教育」として章を設け、「超高齢社会を迎え地域における福祉介護等の関係機関との連携により、包括的かつ継続的な地域完結・循環型医療』の提供を行うことが必要とされ、卒前教育にも、多職種連携・多職種協働やチーム医療を具体的にイメージできるカリキュラムが求められている。『医師として求められる基本的な資質・能力』に地域医療やチーム医療、コミュニケーション能力を列挙するのみならず、A-4-1)コミュニケーション、A-4-2)患者と医師の関係、A-5-1)患者中心のチーム医療、A-7-1)地域医療への貢献、B-1-7)地域医療・地域保健（A-7-1)と学修目標を共有させた）、F-2-15)在宅医療と介護、G-4-3)地域医療実習の各項目で触れている。なお、単に高齢者に対する医療や介護だけではなく、全年齢を見据えた予防も含めた地域保健や関連する地域福祉の理解と実践が求められる。」としている。我々は、こうした視点を教育するにあたり、個別の地域医療教育に加えて、地域のデータに基づく集団の視点の分析を学ぶことの意義を考え、当初から医学生や若手に地域データの分析機会を設けてきた。

また、2020年初頭からの新型コロナウイルスのパンデミックに伴い、リアルワールドデータの有効活用に基づく evidence based policy making (EBPM)の重要性が広く一般世間・社会に認識されるようになってきた。この時に重要なことは、時間と研究意欲に溢れる若手研究者が、経験・実績の豊富な指導者の下で、素早くデータを解

析し、報告書や論文を執筆して世に発信することである。そのような教育体制・環境を日頃から構築することが重要である。

本研究班により、これらの取り組みが加速でき、昨年度に引き続き、今年度は研修医・博士課程生・若手教員が実際に国際誌掲載に成果をあげることができた。また本研究班の取り組みから得た人材育成の知見に基づき、本研究班の分担研究者・協力研究者である教員がシンポジウム講演や講義をする機会も得ることができた。以下、この1年の成果を報告する。

I 医療・介護保険レセプトデータ活用による研究・人材育成（1）論文発表

本年度は医療・介護ビッグデータ活用による論文を2本、および学会発表を1回行った。解析及び論文執筆は、本研究分担者（田宮）および本研究協力者（岩上）の指導のもと、当教室（筑波大学医学医療系ヘルスサービスリサーチ分野）に所属する博士課程学生および研修医が行った。下記に論文の要旨をまとめる。

1. 論文タイトル：Evaluation of enhanced home care support clinics regarding emergency home visits, hospitalization, and end-of-life care: a retrospective cohort study in a city of Japan (BMC Health Serv Res. 2023 Feb 3;23(1):115.)

【発表者】孫瑜（博士課程4年）

【目的】高齢化に伴い増加する在宅医療の需要に対応するため、2006年に在宅療養支援診療所・病院（HCSC）が、2012年に強化型HCSCがそれぞれ導入された。本研究は、強化型HCSCが在宅医療において期待される役割を果たしたかどうかを評価することを目的とした。

【方法】日本のある自治体の医療・介護請求データを用いて、後ろ向きコホート研究を実施した。対象者は65歳以上で、2014

年7月から2018年3月の間に新たに定期訪問診療を開始し、従来のHCSCまたは強化型HCSCのいずれかを使用した。患者は、定期訪問診療を開始してから1年間、または1年以内に終了した場合は定期訪問診療終了の翌月まで、追跡された。アウトカム指標は、(i)全時間帯および夜間・休日の緊急の往診をそれぞれ1回以上、(ii)入院を1回以上、(iii)看取り医療として死亡場所と家庭内死亡時の医師の立ち合いの有無、により評価した。多変量ロジスティック回帰分析は、緊急の往診と入院のアウトカムについて実施した。

【結果】 分析対象は802人で、そのうち強化型HCSCの405人と従来型HCSCの397人であった。強化型HCSCでは、従来型HCSCよりも全時間帯での緊急の往診が多く(65.7%対49.1%;調整オッズ比1.70、95%CI [1.26-2.28])、夜間・休日の緊急の往診が多く(33.6%対16.7%;2.20 [1.55-3.13])、入院が少なかった(21.5%対32.2%;0.55 [0.39-0.76])。追跡期間中、229人の患者(強化型HCSCでは152人、HCSCでは77人)が死亡した。在宅での死亡は、従来のHCSCよりも強化型HCSCで有意に多く(80.9%対64.9%; $p < .001$)、在宅で死亡した患者のうち医師の立ち合いによる死亡も強化型HCSCで有意に多かった(99.2%対78.0%; $p < .001$)。

【考察】 本研究により、在宅医療における重要な医療機能である緊急往診や在宅での看取りは、強化型HCSCの方が対応できる可能性が高いことが確認された。強化型HCSCのさらなる普及は有効であると考えられる。

2. 論文タイトル : Pulse Methylprednisolone versus Dexamethasone in COVID-19: A

Multicenter Cohort Study (Crit Care Explor. 2023;5(4):e0886.)

【発表者】 渡邊淳之 (筑波大学病院初期研修医)

【背景】 パルスメチルプレドニゾロン療法は、仮説的には免疫系のフレアを効果的にコントロールできるが、COVID-19におけるデキサメタゾンと比較したパルスメチルプレドニゾロンの臨床的有用性は、まだ結論が出ていない。

【目的】 COVID-19の治療薬としてパルスメチルプレドニゾロンとデキサメタゾンを比較すること。

【方法】 日本のMedical Data Visionデータベースを用いて、2020年1月~2021年12月にCOVID-19で入院し退院した成人患者のうち、入院0日目または1日目にパルスメチルプレドニゾロン(250、500、1000mg/日)またはデキサメタゾン(≥ 6 mg/日)の静脈内投与で治療を受けた患者を特定した。主要アウトカムは院内死亡率、副次的アウトカムは、30日死亡率、新規ICU入室、インスリン開始、真菌感染症、再入院と設定した。多変量ロジスティック回帰を行い、パルスメチルプレドニゾロンの用量(250、500、1,000mg/日)を区別した。さらに、侵襲的機械換気(IMV)の必要性などの特徴によるサブグループ分析も実施した。

【結果】 合計7,519人、197人、399人、1,046人の患者に、それぞれデキサメタゾン、250、500、1,000mg/日のメチルプレドニゾロンが開始された。院内粗死亡率は、それぞれ9.3% (702/7,519)、8.6% (17/197)、17.0% (68/399)、16.2% (169/1,046)であった。調整オッズ比(95%信頼区間)は、デキサメタゾンを開始した患者と比較して、メチルプレドニゾロン250、500、1,000mg/日をそれぞれ開始した患者で1.26 (0.69-2.29)、1.48 (1.07-2.04)、1.75 (1.40-2.19)であった。サブ

グループ解析では、IMV 患者では、院内死亡の調整オッズ比は 0.78 (0.25-2.47)、1.12 (0.55-2.27) および 1.04 (0.68-1.57) であったのに対し、非 IMV 患者では 1.54 (0.77-3.08)、1.62 (1.13-2.34)、2.14 (1.64-2.80) であった。

【解釈】パルスメチルプレドニゾロンの高用量 (500 または 1,000mg/日) は、デキサメタゾンと比較して、特に IMV を使用していない患者において、COVID-19 の悪い転帰と関連する可能性が示唆された。

3. 講演タイトル：高齢 COVID-19 入院患者における入院早期リハビリテーションと在院死亡との関連 (2022 年度第 5 回日本臨床疫学会 YIA 候補者発表会)

【発表者】宇田和晃 (助教)

【目的】高齢 Coronavirus disease 2019 (COVID-19) 患者に対して、入院早期リハビリテーションは複数の国際機関のステートメントで推奨されている。しかし、入院早期リハビリテーションと在院死亡との関連は不明である。本研究は入院患者データベースを用いて、高齢 COVID-19 入院患者に対する入院早期リハビリテーションと在院死亡との関連を明らかにすることを目的とした

【方法】Medical Data Vision 社から提供を受けた入院患者データベースを利用し、2020 年 1 月～2021 年 12 月の退院患者の中から、COVID-19 の診断で入院した 65 歳以上を対象とした。他院からの転院や入院 2 日以内に死亡・退院した者は除外した。入院 2 日以内にリハビリテーションを実施した者を早期リハ群、それ以外を対照群と分類した。入院時期、年齢、性別、BMI、喫煙歴、併存疾患、要介護度、入院時 Barthel Index、入院 2 日以内の診療行為、入院料区分、および地域加算区分から作成した傾向スコアを用いて 1:1 の比率で nearest neighbor matching (caliper 0.2 SD) を行っ

た。マッチされた集団で、早期リハ群 vs. 対照群の在院死亡のリスク比およびリスク差を推定した。欠損値 (BMI、喫煙歴、Barthel Index 等) は多重代入法で補完し、推定結果は Rubin's rule で統合した。

【結果】対象者は 30,054 人、早期リハ群は 2,059 人 (6.9%) であった。早期リハ群は対照群と比べ第 4 波と第 5 波の時期が多い、より高齢で要介護認定者や認知症が多い、入院時 Barthel Index が低い、併存疾患指数が高い、人工呼吸器装着が多い、などの特徴があった。マッチされた集団は早期リハ群 2,037 人 vs. 対照群 2,037 人となり、在院死亡は 10.1% vs. 13.1%、リスク比は 0.77 (95%CI: 0.65～0.92)、リスク差は -3.0% (95%CI: -1.0%～-5.0%) と推定された。

【結論】高齢 COVID-19 入院患者において入院早期リハビリテーションと在院死亡リスク減少との関連が示唆された。ただし、未測定交絡因子の影響は否定できないため、さらなる研究が必要である。

なお、本研究は、同学会にて優秀演題賞として受賞された。

以上のように、医療・介護ビッグデータを有効に活用し新たな医学的知見を得ることができた。また、研究計画立案から、データクリーニング、統計解析、論文執筆、論文投稿、レビューとのやり取り、までの一連の流れについて、次世代を担う若手に経験してもらおうと共に、明確な成果物として世に発信することができた。今回の論文執筆をきっかけに、今後の活躍が期待できる。

II 医療・介護データ活用による研究・人材育成 (2) 日本臨床疫学会年次学術総会での教育講演

2022 年 11 月 12 日～13 日に開催された日本臨床疫学会第 5 回年次学術大会にて、本研

究分担者である田宮菜奈子がシンポジウム（リアルワールドデータ研究の人材育成）で講演を行い、また本研究協力者である岩上将夫がシンポジウム（臨床疫学研究に使われるデータベース：電子カルテデータベースとその研究利用）で臨床疫学の初中級者を対象に講演を行った。

このように、本研究班の重要な目的の1つである「後進の育成」に関する活動の一環として、筑波大学の若手に限らず、日本全国を対象にした啓蒙教育活動に関わることができた。

以上のように、本研究分担を通じて、医療・介護ビッグデータ研究の後進の育成実績およびそのノウハウが順調に蓄積されている。これは、今後の医療人材のあるべき方向にも沿ったものであり、かつ学生や若手自身が自ら意欲的に取り組める内容でもあり、適切な指導により高い成果につながることが示唆された。

今後、以上の経験を生かし、それぞれの分野やレベルの違いに応じた医療・介護ビッグデータ人材教育の最適な方法論についてまとめていくことが必要であると考え

F. 研究発表

1. 論文発表 2本

1. Yu Sun, Masao Iwagami, Nobuo Sakata, Tomoko Ito, Ryota Inokuchi, Jun Komiyama, Naoaki Kuroda, Nanako Tamiya. Evaluation of enhanced home care support clinics regarding emergency home visits, hospitalization, and end-of-life care: a retrospective cohort study in a

city of Japan. BMC Health Serv Res . 2023 Feb 3;23(1):115.

2. Atsuyuki Watanabe, Ryota Inokuchi, Toshiki Kuno, Kazuaki Uda, Jun Komiyama, Motohiko Adomi, Yoshiko Ishisaka, Toshikazu Abe, Nanako Tamiya, Masao Iwagami. Pulse Methylprednisolone versus Dexamethasone in COVID-19: A Multicenter Cohort Study. Crit Care Explore. 2023;5(4):e0886..

2. 学会発表 3回

1. 宇田和晃、井口竜太、小宮山潤、田宮菜奈子、岩上将夫. 高齢 COVID-19 入院患者における入院早期リハビリテーションと在院死亡との関連、日本臨床疫学会第 5 回年次学術大会、2022 年
2. 田宮菜奈子、シンポジウム：リアルワールドデータ研究の人材育成、日本臨床疫学会第 5 回年次学術大会、2022 年
3. 岩上将夫、シンポジウム：臨床疫学研究に使われるデータベース：電子カルテデータベースとその研究利用、日本臨床疫学会第 5 回年次学術大会、2022 年

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

厚生労働行政推進調査事業補助金
政策科学総合研究事業（政策科学推進研究事業）
（分担）研究報告書

診療現場の実態に即した医療ビッグデータ（NDB 等）を利活用できる人材育成促進に資
するための研究

研究分担者 笹渕裕介 自治医科大学データサイエンスセンター

研究要旨：令和2年度までに開発した医療ビッグデータハンドリング教育プログラム、e-learning 用プログラムは一方向型であり、より学習効率の高いプログラムの必要性がある。そこで、本研究の目的を双方向型オンライン教育プログラムの開発を行うこととした。令和3年度は双方向型オンライン教育プログラム開発にあたり当センターで研究を行っている研究者・大学院生と、双方向型プログラム開発のための意見交換を行った。その結果を元に、令和4年度は研究計画立案、統計解析用ソフトウェア R について双方向型オンラインプログラムの開発を行い、研究者へ提供した。本研究事業で開発した教育プログラムは医療ビッグデータを利活用できる人材育成促進につながったと考えられる。

A. 研究目的

近年医療ビッグデータを利用した研究が爆発的に増加しており、臨床判断や医療政策策定に必要なエビデンスの創出に大きな役割を占めていることはすでに周知の事実である。医療ビッグデータを利用した研究を行うためには通常の臨床データを用いた臨床研究に必要な臨床の知識・臨床研究の知識に加えて、データベースハンドリング・統計学等の知識や技術を要する。我々は平成 29-30 年度および平成 31-令和 2 年度厚生労働科学研究において、医療ビッグデータを利用した研究を行う際に必要なデータハンドリング技術である SQL 言語、統計解析に必要な R、SPSS 等の統計ソフトの習得を目指す教育プログラム、e-learning 用教育プログラムの開発を行い多くの対象者へ提

供してきた。一方、これらのプログラムは一方向性であり学習効率が必ずしも高くなく、より学習効率の高いプログラムが求められている。そこで、本研究の目的はこれまでに開発したプログラムをベースに、双方向型教育プログラムの開発を行うこととした。

B. 研究方法

①これまで自治医科大学データサイエンスセンターにおいて医療ビッグデータを利用した研究を行っている研究者、大学院生に対して必要なプログラムについて意見交換を行う。

②研究者・大学院生の意見を元にこれまでに作成した医療ビッグデータハンドリング教育プログラムから双方向型オンラインプ

プログラムを開発するコンテンツの選別を行う。

③選別したコンテンツに関する双方向型のオンラインプログラムを開発する。

④実際にプログラムを提供し医療ビッグデータ研究実施へとつなげていく。

C. 研究結果

①②特に統計解析のプログラムは双方向型であることが習得の近道であるとの意見が多かった。また、ほぼすべての研究者、大学院生から研究計画立案について知識が足りないことについて教育を受ける機会が不十分であるとの意見があった。そこで研究者にとって必要性の高い研究計画および統計ソフトウェア R についての双方向型オンラインプログラムの開発を行うこととした。

双方向型教育プログラムのコンテンツ

(1) 研究計画立案

研究計画の立案について教育を受ける機会が少ない、書籍等での学習と自身での計画立案にはギャップがあり、自力でそのギャップを埋めることが難しい等の意見を得た。そこで、研究計画立案に際して必須の以下の項目について双方向型オンラインプログラムの開発を行った。具体的には講義に加えて、実践的な作業を行うことができるような研究計画立案に関する課題を設定し、それを解くことで能動的に講義の内容を理解できるようなコンテンツとした。テーマは以下のとおりである。

・クリニカルクエスチョン (CQ) から

リサーチクエスチョン (RQ) への作り変え

- ・ FINER による RQ の評価
- ・ 文献検索
- ・ 研究デザイン
- ・ 抽出変数の定義

(2) R による統計解析

R に関して令和 2 年度までに開発した e-learning プログラムを改良し、以下の項目について双方向型オンライン教育プログラムの開発を行った。具体的には講義中に実際に自分でスクリプトを書くことで R の基本的な使用方法を理解することに加えて、講義の内容を反映した課題を設定し、それを解くことで能動的にスクリプトを書くことができるようなコンテンツとした。テーマは以下のとおりである。

- ・ データの整形
- ・ 各変数の集計及び可視化
- ・ 検定
- ・ 回帰分析

③研究計画立案、R による統計解析に関して動画撮影が終了し、オンラインプログラムとして使用可能なプログラムの開発が完了した。

④プログラムを自治医科大学データサイエンスセンターにおいて大学院生、客員研究員、共同研究者に提供した。提供を受けた複数の研究者により研究計画の立案から論文発表まで行われた。

(倫理面への配慮)

倫理的な問題はない。

D. 考察

医療ビッグデータを用いた研究のための双方向型教育プログラムの作成にあたり、研究者・大学院生と意見交換を行い、研究者はデータハンドリングや統計解析に加えて研究デザインの教育も必要であることが明らかとなった。

当センターで行う研究についての研究計画の立案は分担者が個別に各研究者や大学院生に対して直接指導を行っている。令和2年度までにこれらの指導内容を体系化し、直接指導を行わなくとも保健医療ビッグデータ研究を行うためのデータハンドリングから解析までを習得可能な教育プログラムを開発し、研究者や大学院生に提供してきた。これらの教育プログラムによってビッグデータを用いた研究の実践を通して、学会発表・論文執筆につながった。しかしながらこれらの教育プログラムは一方向性であり、一部十分な習得に至らなかった部分に関して改良の余地が残されていた。本研究ではこれらの習得が不十分な点について、プログラム受講者から意見を集めることで特に研究者にとってデータハンドリングや統計解析に加えて研究計画の立案についても教育プログラムが必要であることが明らかとなった。

ずさんな研究計画のまま研究を開始すると、研究途中で様々な問題が生じ、研究自体が頓挫してしまうこともしばしば起こる。そ

のため、綿密な研究計画を立案する必要がある。本研究で開発した研究計画立案のための双方向型オンライン教育プログラムはこれまで分担者が行ってきた指導をオンライン用に構築したものである。また、令和2年度までに開発したR習得 e-learning プログラムを、動画視聴の合間に複数の課題を設定することにより能動的にRを習得することのできる双方向型のオンライン教育プログラムとして改良した。プログラム受講者からの評価は良好であったが、更に発展的な内容のプログラムを望む声も寄せられており、今後の課題としたい。

本研究事業により開発した教育プログラムは医療ビッグデータを利活用できる人材育成につながったと考えられる。実際にプログラム受講者が研究デザインやデータハンドリングを行うことができるようになっており、研究成果として論文発表につながっている (G. 研究発表参照) ことから有用なプログラムであったことが示された。今後このプログラムをより多くの研究者へ提供していく予定である。

E. 結論

医療ビッグデータ研究の為に必要な知識・技術を養成するための双方向型オンライン教育プログラム開発について研究者と意見交換を行い、研究計画立案、統計解析について特に効率的な教育プログラムが必要であることが明らかとなった。令和2年度までに開発した教育プログラムを能動的に学

習することのできる双方向型オンライン教育プログラムとして再構築した。この双方向型オンライン教育プログラムの提供により、これまでよりもスムーズに研究アイデアを論文執筆まで繋げられるようになった。

F. 健康危険情報

なし

G. 研究発表

1.論文発表

1-1. Ono S, Sasabuchi Y, Ishimaru M, Ono Y, Matsui H, Yasunaga H. Short-term effects of reduced cost sharing on childhood dental care utilization and dental caries prevention in Japan. *Community Dent Oral Epidemiol.* 2023;51:228-235.

1-2. Yoshida T, Uchino S, Sasabuchi Y. Epidemiology of constipation in critically ill patients and its impact on in-hospital mortality: a retrospective observational study. *J Anesth.* 2022;36:349-358.

1-3. Suzuki R, Uchino S, Sasabuchi Y, Kawarai Lefor A, Sanui M. Dopamine use and its consequences in the intensive care unit: a cohort study utilizing the Japanese Intensive care PATient Database. *Crit Care.* 2022;26:90.

1-4. Chinen T, Sasabuchi Y, Matsui H, Yamaguchi H, Yasunaga H. Oxaliplatin-versus cisplatin-based regimens for elderly individuals with advanced gastric cancer: a

retrospective cohort study. *BMC Cancer.* 2022;22:460.

1-5. Ookawara S, Ito K, Sasabuchi Y, Miyahara M, Miyashita T, Takemi N, Nagamine C, Nakahara S, Horiuchi Y, Inose N, Shiina M, Murakoshi M, Sanayama H, Hirai K, Morishita Y. Cerebral oxygenation and body mass index association with cognitive function in chronic kidney disease patients without dialysis: a longitudinal study. *Sci Rep.* 2022;12:10809.

1-6. Iwasaki Y, Sasabuchi Y, Horikita S, Furukawa T, Shiotsuka J, Lefor AK, Sanui M. The effect of preoperative sodium-glucose cotransporter 2 inhibitors on the incidence of perioperative metabolic acidosis: A retrospective cohort study. *BMC Endocr Disord.* 2022;22:209.

1-7. Wada Y, Takahashi H, Matsui H, Yasunaga H, Fujiwara H, Sasabuchi Y. Adhesion barriers and intraperitoneal or uterine infections after cesarean section: A retrospective cohort study. *Surgery.* 2022;172:1722-1727.

1-8. Hayashi K, Sasabuchi Y, Matsui H, Nakajima M, Ohbe H, Fushimi K, Ono K, Yasunaga H. Additive effectiveness of acrylonitrile-co-methallyl sulfonate surface-treated membranes in the treatment of pneumonia: A propensity score-matched retrospective cohort study. *Artif Organs.* 2023;47:408-416.

- 1-9. Shiotsuka J, Uchino S, Sasabuchi Y, Masuyama T, Lefor AK, Sanui M. Reducing the effect of immortal time bias affects the analysis of prevention of delirium by suvorexant in critically ill patients: A retrospective cohort study. *PLoS One*. 2022;17:e0277916.
- 1-10. Hayashi K, Sasabuchi Y, Matsui H, Nakajima M, Otawara M, Ohbe H, Fushimi K, Ono K, Yasunaga H. Does early excision or skin grafting of severe burns improve prognosis? A retrospective cohort study. *Burns*. 2023;49:554-561.
- 1-11. Ohbe H, Ouchi K, Miyamoto Y, Ishigami Y, Matsui H, Yasunaga H, Sasabuchi Y. One-Year Functional Outcomes After Invasive Mechanical Ventilation for Older Adults With Preexisting Long-Term Care-Needs. *Crit Care Med*. 2023:e005822. Epub ahead of print.
- 1-12. Murooka Y, Sasabuchi Y, Takazawa T, Matsui H, Yasunaga H, Saito S. Long-Term Prognosis Following Early Rehabilitation in the ICU: A Retrospective Cohort Study. *Crit Care Med*. 2023. Epub ahead of print.
- 1-13. Wada Y, Takahashi H, Sasabuchi Y, Usui R, Ogoyama M, Suzuki H, Ohkuchi A, Fujiwara H. Maternal outcomes of placental abruption with intrauterine fetal death and delivery routes: A nationwide observational study. *Acta Obstet Gynecol Scand*. 2023. Epub ahead of print.
- 2.学会発表
- 2-1. 笹渕裕介. 自治医科大学における人材育成の取り組み. 第5回日本臨床疫学会年次学術大会. 2022/11/13 (東京).
- H. 知的財産権の出願・登録状況
なし

研究成果の刊行に関する一覧表

書籍
なし

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Abe H, Sumitani M, Matsui H, Inoue R, Fushimi K, Uchida K, Yasunaga H	Use of naldemedine is associated with reduced incidence of hyperactive delirium in cancer patients with opioid-induced constipation: a nationwide retrospective cohort study in Japan	Pharmacotherapy	42(3)	241-249	2022
Abe H, Sumitani M, Matsui H, Inoue R, Konishi M, Fushimi K, Uchida K, Yasunaga H	Gabapentinoid use is associated with reduced occurrence of hyperactive delirium in older cancer patients undergoing chemotherapy: a nationwide retrospective cohort study in Japan	Anesthesia & Analgesia	135(2)	362-369	2022
Aso S, Matsui H, Fushimi K, Yasunaga H	Effect of Dobutamine on Patients with Septic Shock: A Retrospective Nationwide Study	Annals of Clinical Epidemiology	4(2)	41-48	2022
Awano N, Jo T, Izumo T, Inomata M, Morita K, Matsui H, Fushimi K, Urushiyama H, Nagase T, Yasunaga H	Recombinant human soluble thrombomodulin for acute exacerbation of idiopathic pulmonary fibrosis: A nationwide observational study	Journal of Intensive Care	10(1)	14	2022
Chinen T, Sasabuchi Y, Matsui H, Yamaguchi H, Yasunaga H	Oxaliplatin- versus cisplatin-based regimens for elderly individuals with advanced gastric cancer: A retrospective cohort study	BMC Cancer	22(1)	460	2022
Fujita A, Hashimoto Y, Matsui H, Yasunaga H, Aihara M	Recent trends in glaucoma surgery: a nationwide database study in Japan, 2011–2019	Japanese Journal of Ophthalmology	66(2)	183-192	2022
Fujita A, Hashimoto Y, Okada A, Obata R, Aihara M, Matsui H, Yasunaga H	Association Between Proteinuria and Retinal Vein Occlusion in Individuals with Preserved Renal Function: A Retrospective Cohort Study	Acta Ophthalmologica	100(7)	e1510-e1517	2022
Fukui R, Kaneko H, Okada A, Yano Y, Itoh H, Matsuoka S, Morita K, Kiriyaama H, Kamon T, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Nishiyama A, Node K, Yokoo T, Nangaku M, Yasunaga H, Komuro I	Semiquantitative assessed proteinuria and risk of heart failure: Analysis of a nationwide epidemiological database	Nephrology Dialysis Transplantation	37(9)	1691-1699	2022
Furuya H, Nakajima M, Ikeda K, Nakamura K, Ohbe H, Aso S, Kumazawa R, Iwamoto T, Iwata A, Furuta S, Matsui H, Fushimi K, Yasunaga H, Nakajima H	Prognosis and treatment of myositis-associated severe interstitial lung disease	Arthritis Care & Research	74(3)	478-483	2022
Fuseda Y, Michihata N, Kumazawa R, Fushimi K, Yasunaga H	Annual Trend in Lumbar Puncture for Infants Younger Than 3 Months Hospitalized with Suspected Serious Bacterial Infection: A Nationwide Inpatient Database Study	Pediatric Infectious Disease Journal	41(8)	631-635	2022
Hamada S, Nakajima M, Kaszynski RH, Otaka S, Goto H, Matui H, Fushimi K, Yamaguchi Y, Yasunaga H	Association between adjunct clindamycin and in-hospital mortality in patients with necrotizing soft tissue infection due to group A Streptococcus: A nationwide cohort study	European Journal of Clinical Microbiology & Infectious Diseases	41(2)	263-270	2022
Hashimoto Y, Kaneko H, Okada A, Matsui H, Yasunaga H, Aihara M, Obata R	Association between retinal vein occlusion and Life's Simple 7 cardiovascular health metrics: A large claims database study	Ophthalmology Retina	6(8)	684-692	2022
Hashimoto Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, Kaburaki T	Incidence of sympathetic ophthalmia after inciting events: a national database study in Japan	Ophthalmology	129(3)	344-352	2022
Hashimoto Y, Michihata N, Yamana H, Shigemi D, Morita K, Matsui H, Yasunaga H, Aihara M	Safety of topical ophthalmic antibiotics in pregnant women with hordeola, chalazia, blepharitis, or bacterial conjunctivitis: propensity score analyses	EYE	36(5)	1066-1073	2022
Hashimoto Y, Yamana H, Michihata N, Shigemi D, Ishimaru M, Matsui H, Yasunaga H, Aihara M	Eye drops for dry eye disease during pregnancy and adverse neonatal outcomes: high-dimensional propensity score analyses	Ophthalmic Epidemiology	29(4)	384-393	2022

Hirano T, Nakajima M, Ohbe H, Kaszynski RH, Iwasaki Y, Arakawa Y, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H	Corticosteroid use with extracorporeal cardiopulmonary resuscitation for out-of-hospital cardiac arrest: A nationwide observational study	Resuscitation Plus	12	100308	2022
Hirano Y, Kaneko H, Konishi T, Itoh H, Matsuda S, Kawakubo H, Uda K, Matsui H, Fushimi K, Daiko H, Itano O, Yasunaga H, Kitagawa Y	Short-Term Outcomes of Epidural Analgesia in Minimally Invasive Esophagectomy for Esophageal Cancer: Nationwide Inpatient Data Study in Japan	Annals of Surgical Oncology	29(13)	8225-8234	2022
Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H	Association between nurse aide staffing and patient mortality after major cancer surgeries in acute care settings: A retrospective cohort study using a national inpatient database	Nursing and Health Sciences	24(1)	283-292	2022
Hirose N, Morita K, Matsui H, Fushimi K, Yasunaga H	Dose-response association between nurse staffing and patient outcomes following major cancer surgeries using a nationwide inpatient database in Japan	Journal of Clinical Nursing	31(17-18)	2562-2573	2022
Honda A, Michihata N, Iizuka Y, Uda K, Morita K, Mieda T, Takasawa E, Ishiwata S, Tajika T, Matsui H, Fushimi K, Yasunaga H, Chikuda H	Risk factors for severe lower extremity ischemia following venoarterial extracorporeal membrane oxygenation: An analysis using a nationwide inpatient database	Trauma Surgery & Acute Care Open	7	e000776	2022
Hosoi T, Yamana H, Tamiya H, Matsui H, Fushimi K, Akishita M, Yasunaga H, Ogawa S	Association between comprehensive geriatric assessment and polypharmacy at discharge in patients with ischaemic stroke: a nationwide retrospective cohort study	eClinicalMedicine	50	101528	2022
Inooka Y, Yamana H, Shinoda Y, Inokuchi H, Matsui H, Fushimi K, Yasunaga H, Haga N	Predictive factors for oral intake recovery after acute stroke: analysis of a Japanese nationwide inpatient database	Dysphagia	37(6)	1623-1632	2022
Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H	Prevalence, incidence rate and risk factors of medication-related osteonecrosis of the jaw (MRONJ) in patients with osteoporosis and cancer: a nationwide population-based study in Japan	Journal of Oral and Maxillofacial Surgery	80(4)	714-727	2022
Ishimaru T, Fujiogi M, Michihata N, Konishi T, Morita K, Matsui H, Uda K, Fushimi K, Kawashima H, Fujishiro J, Yasunaga H	Perioperative Outcomes of Laparoscopically Assisted Anorectoplasty versus Conventional Procedures for Anorectal Malformation: A Retrospective Nationwide Database Study	Pediatric Surgery International	38(12)	1785-1791	2022
Itoh H, Kaneko H, Okada A, Suzuki Y, Fujiu K, Matsuoka S, Michihata N, Jo T, Nakanishi K, Takeda N, Morita H, Node K, Di Tullio MR, Homma S, Yasunaga H, Komuro I	Age-Specific Relation of Cardiovascular Health Metrics With Incident Cardiovascular Disease	American Journal of Cardiology	177	34-39	2022
Iwasaki Y, Ohbe H, Shigemi D, Fushimi K, Yasunaga H	Effect of antithrombin III among patients with disseminated intravascular coagulation in obstetrics: a nationwide observational study in Japan	BJOG: An International Journal of Obstetrics and Gynaecology	129(5)	805-811	2022
Kamijo K, Nakajima M, Shigemi D, Kaszynski RH, Ohbe H, Goto T, Sasabuchi Y, Fushimi K, Matsui H, Yasunaga H	Resuscitative endovascular balloon occlusion of the aorta for life-threatening postpartum hemorrhage: A nationwide observational study in Japan	The Journal of Trauma and Acute Care Surgery	93(3)	418-423	2022
Kamon T, Kaneko H, Itoh H, Okada A, Matsuoka S, Kiriyama H, Fujiu K, Morita K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I	Sex difference in the association between lipid profile and incident cardiovascular disease among young adults	Journal of Atherosclerosis and Thrombosis	29(10)	1475-1486	2022
Kaneko H, Suzuki Y, Ueno K, Okada A, Fujiu K, Matsuoka S, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Node K, Yasunaga H, Komuro I	Association of Life's Simple 7 with incident cardiovascular disease in 53,974 patients with cancer	European Journal of Preventive Cardiology	29(18)	2324-2332	2022
Kaneko H, Yano Y, Lee HH, Lee H, Okada A, Itoh H, Morita K, Fukui A, Fujiu K, Suzuki Y, Matsuoka S, Nakamura S, Michihata N, Jo T, Takeda N, Morita H, Yokoo T, Nishiyama A, Node K, Viera AJ, Muntner P, Oparil S, Kim HC, Yasunaga H, Komuro I	Medication-Naïve Blood Pressure and Incident Cancers: Analysis of Two Nationwide Population-Based Databases	American Journal of Hypertension	35(8)	731-739	2022
Kaneko H, Yano Y, Suzuki Y, Okada A, Itoh H, Matsuoka S, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Node K, Viera AJ, Lima JAC, Oparil S, Lam CSP, Carey RM, Yasunaga H, Komuro I	Reduction in blood pressure for elevated blood pressure/ stage 1 hypertension according to the ACC/AHA guideline and cardiovascular outcomes	European Journal of Preventive Cardiology	29(14)	1921-1929	2022

Karakawa R, Konishi T, Yoshimatsu H, Fuse Y, Hashimoto Y, Matsui H, Fushimi K, Yano T, Yasunaga H	Comparison of short-term outcomes between pedicled- and free-flap autologous breast reconstruction	Breast Cancer	29(6)	1067-1075	2022
Kodama S, Jo T, Yasunaga H, Michihata N, Matsui H, Kumazawa R, Shirota Y, Fushimi K, Toda T, Hamada M	Outcomes of gastrointestinal cancer surgeries in Parkinson's disease patients: a nationwide study	Parkinsonism and Related Disorders	96	45-49	2022
Konishi T, Fujiogi M, Michihata N, Kumazawa R, Ohbe H, Matsui H, Fushimi K, Ogita M, Tanabe M, Seto Y, Yasunaga H	Interstitial lung disorders following postoperative radiotherapy with concurrent or sequential hormonal therapy for breast cancer: a nationwide database study in Japan	Breast Cancer	29(4)	688-697	2022
Konishi T, Fujiogi M, Michihata N, Matsui H, Tanabe M, Seto Y, Yasunaga H	Association between body mass index and incidence of breast cancer in premenopausal women: A Japanese nationwide database study	Breast Cancer Research and Treatment	194(2)	315-325	2022
Konishi T, Fujiogi M, Michihata N, Morita K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Association of operative day of the week with length of stay and total hospitalization costs in patients with partial mastectomy: a nationwide database study in Japan	JMA Journal	5(3)	319-327	2022
Konishi T, Fujiogi M, Michihata N, Ohbe H, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Cost-effectiveness analysis of trastuzumab monotherapy versus adjuvant chemotherapy plus trastuzumab in elderly patients with HER2-positive early breast cancer	Japanese Journal of Clinical Oncology	52(10)	1115-1123	2022
Konishi T, Fujiogi M, Sato A, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Short-term outcomes following breast cancer surgery with and without neoadjuvant chemotherapy: a nationwide administrative database study in Japan	Annals of Surgical Oncology	29(5)	3000-3010	2022
Konishi T, Fujiogi M, Shigemi D, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Risk factors for postoperative bleeding following breast cancer surgery: a nationwide database study of 477,108 cases in Japan	World Journal of Surgery	46(12)	3062-3071	2022
Konishi T, Fujiogi M, Shigemi D, Nishioka K, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Surgical and obstetric outcomes of breast cancer surgery during pregnancy: A nationwide database study in Japan	Breast Cancer Research and Treatment	195(3)	289-299	2022
Konishi T, Goto T, Fujiogi M, Michihata N, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using a Japanese nationwide inpatient database	Surgery	171(4)	1036-1042	2022
Konishi T, Sakata A, Inokuchi H, Kumazawa R, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Treatments and outcomes of adult parapharyngeal and retropharyngeal abscess: 1,882 cases from a Japanese nationwide database	American Journal of Otolaryngology	44(2)	103770	2022
Konishi T, Takamoto T, Fujiogi M, Hashimoto Y, Matsui H, Fushimi K, Tanabe M, Seto Y, Yasunaga H	Laparoscopic versus open distal pancreatectomy with or without splenectomy: A propensity score analysis in Japan	International Journal of Surgery	104	106765	2022
Konishi T, Yoshimoto T, Fujiogi M, Yamana H, Tanabe M, Seto Y, Yasunaga H	Validity of operative information in Japanese administrative data: a chart review-based analysis of 1221 cases at a single institution	Surgery Today	52(10)	1484-1490	2022
Kumazawa R, Jo T, Matsui H, Fushimi K, Yasunaga H	Direct oral anticoagulants versus warfarin for secondary prevention of cerebral infarction and bleeding in older adults with atrial fibrillation	Journal of the American Geriatrics Society	70(7)	2029-2039	2022
Kurakawa KI, Okada A, Bessho K, Jo T, Ono S, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yamaguchi S, Yamauchi T, Nangaku M, Kadowaki T, Yasunaga H	Major complications after percutaneous biopsy of native or transplanted liver in pediatric patients: A nationwide inpatient database study in Japan	BMC Gastroenterology	22(1)	395	2022
Kutsuna S, Ohbe H, Matsui H, Yasunaga H	Effectiveness of fluoroquinolone antimicrobials in addition to tetracyclines for Japanese Spotted Fever: A retrospective analysis using a national inpatient database	International Journal of Infectious Diseases	123	70-75	2022
Kutsuna S, Ohbe H, Matsui H, Yasunaga H	Delayed Tetracycline Initiation Increases Mortality Risk in Patients With Japanese Spotted Fever: Retrospective Analysis Using a National Inpatient Database	Open Forum Infectious Diseases	9(11)	ofac573	2022
Matsuda K, Jo T, Matsui H, Fushimi K, Yasunaga H, Sugimoto K	Institutional factors associated with early mortality of newly diagnosed acute promyelocytic leukemia	Blood Cancer Journal	12(12)	167	2022
Matsuda K, Jo T, Shimura A, Honda A, Taoka K, Masamoto Y, Matsui H, Fushimi K, Yasunaga H, Kurokawa M	Risk of febrile neutropenia in very elderly patients aged ≥ 80 years receiving their first cycle of R-CHOP regimen: a nationwide real-world study in Japan	British Journal of Haematology	197(3)	e37-e41	2022

Matsui H, Yamana H, Fushimi K, Yasunaga H	Development of deep learning models for predicting in-hospital mortality using an administrative claims database	JMIR Medical Informatics	10(2)	e27936	2022
Matsuoka S, Kaneko H, Kamon T, Suzuki Y, Yano Y, Okada A, Itoh H, Morita K, Fukui A, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Yokoo T, Nishiyama A, Node K, Yasunaga H, Komuro I	Incorporation of Retinal Arteriosclerosis into Risk Stratification of Blood Pressure Category According to the 2017 ACC/AHA Blood Pressure Guideline	Journal of Atherosclerosis and Thrombosis	29(10)	1487-1498	2022
Matsuoka S, Kaneko H, Okada A, Fukui A, Yano Y, Itoh H, Morita K, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Nakamura S, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I	Association between Proteinuria and Incident Colorectal Cancer: Analysis of a Nationwide Population-Based Database	BMJ Open	12(4)	e056250	2022
Matsuoka S, Kaneko H, Okada A, Itoh H, Suzuki Y, Fujiu K, Michihata N, Jo T, Takeda N, Morita H, Nakamura S, Node K, Yasunaga H, Komuro I	Association of retinal atherosclerosis assessed using Keith-Wagener-Barker system with incident heart failure and other atherosclerotic cardiovascular disease: Analysis of 319,501 individuals from the general population	Atherosclerosis	348	68-74	2022
Matsuoka S, Kaneko H, Okada A, Morita K, Itoh H, Michihata N, Jo T, Takeda N, Morita H, Fujiu K, Nakamura S, Node K, Yasunaga H, Komuro I	Age Modified Relationship between Modifiable Risk Factors and the Risk of Atrial Fibrillation	Circulation: Arrhythmia and Electrophysiology	15(1)	e010409	2022
Michihata N, Shigemi D, Yamana H, Matsui H, Jo T, Yasunaga H	Safety of Japanese Herbal Kampo Medicines for the Treatment of Depression During Pregnancy	International Journal of Gynecology and Obstetrics	159(3)	865-869	2022
Michihata N, Suzuki T, Honda A, Oka I, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H	Clinical features of Kawasaki disease with atlantoaxial rotatory fixation	Pediatric Infectious Disease Journal	41(8)	626-630	2022
Michihata N, Suzuki T, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H	Association between intravenous immunoglobulin dose and outcomes in patients with acute Kawasaki disease	European Journal of Pediatrics	181(10)	3607-3615	2022
Miyachi H, Konishi T, Kumazawa R, Matsui H, Shimizu S, Fushimi K, Matsui H, Yasunaga H	Treatments and outcomes of generalized pustular psoriasis: a cohort of 1516 patients in a nationwide inpatient database in Japan	Journal of the American Academy of Dermatology	86(6)	1266-1274	2022
Miyachi H, Konishi T, Shigemi D, Matsui H, Shimizu S, Fushimi K, Matsue H, Yasunaga H	Association between psoriasis and short-term outcomes of acute myocardial infarction: a matched-pair cohort study using a nationwide inpatient database in Japan	JAAD International	8	21-30	2022
Miyakawa T, Michihata N, Morita K, Matsui H, Honda M, Yasunaga H	Ambient temperature and hospital admissions for acute cholecystitis: A nationwide inpatient database study in Japan	HPB	24(3)	398-403	2022
Miyamoto Y, Iwagami M, Aso S, Uda K, Fushimi K, Hamasaki Y, Nangaku M, Yasunaga H, Doi K	Postoperative outcomes of cancer surgery in patients with and without kidney failure with dialysis therapy: A matched-pair cohort study	Clinical Kidney Journal	15(6)	1137-1143	2022
Miyamoto Y, Matsuyama T, Goto T, Ohbe H, Kitamura T, Yasunaga H, Ohta B	Association between age and neurological outcomes in out-of-hospital cardiac arrest patients resuscitated with extracorporeal cardiopulmonary resuscitation: a nationwide multicentre observational study	European Heart Journal: Acute Cardiovascular Care	11(1)	35-42	2022
Mizuno S, Ono S, Takano A, Yasunaga H, Iwase H	Dental characteristics associated with methamphetamine use: Analysis using forensic autopsy data	BMC Oral Health	22(1)	141	2022
Nagata S, Ohbe H, Jo T, Matsui H, Fushimi K, Yasunaga H	Glucocorticoids and rates of biphasic reactions in patients with adrenaline-treated anaphylaxis: A propensity score matching analysis	International Archives of Allergy and Immunology	183(9)	939-945	2022
Nishimoto Y, Ohbe H, Matsui H, Nakajima M, Sasabuchi Y, Goto T, Morita K, Fushimi K, Sato Y, Yasunaga H	Predictive ability of the sequential organ failure assessment score for in-hospital mortality in patients with cardiac critical illness: a nationwide observational study	EHI Acute Cardiovascular Care	11(4)	312-321	2022
Ogasawara H, Obe H, Yasunaga H	Oral food challenge management in Japan: A retrospective analysis of health insurance claims data	Clinical & Experimental Allergy	52(7)	898-900	2022
Ogawa Y, Yamana H, Noda T, Kishimoto M, Yoshihara S, Kanaoka K, Matsui H, Fushimi K, Yasunaga H, Kawaguchi M, Imamura T	Severe complications after general anesthesia versus sedation during pediatric diagnostic cardiac catheterization for ventricular septal defect	Journal of Clinical Medicine	11(17)	5165	2022
Ohbe H, Goto T, Matsui H, Fushimi K, Yasunaga H	Associations of government-issued ICU admission criteria with clinical practices, outcomes, and ICU bed occupancy	Annals of the American Thoracic Society	19(6)	1013-1021	2022

Ohbe H, Goto T, Nakamura K, Matsui H, Yasunaga H	Development and validation of early prediction models for new-onset functional impairment at hospital discharge of ICU admission	Intensive Care Medicine	48(6)	679-689	2022
Ohbe H, Matsui H, Kumazawa R, Yasunaga H	Intensive care unit versus high dependency care unit admission after emergency surgery: a nationwide in-patient registry study	British Journal of Anaesthesia	129(4)	527-535	2022
Ohbe H, Matsui H, Kumazawa R, Yasunaga H	Postoperative intensive care unit admission following major elective surgery: a nationwide inpatient database study	European Journal of Anaesthesiology	39(5)	436-444	2022
Ohbe H, Matsui H, Yasunaga H	ICU versus high-dependency care unit for patients with acute myocardial infarction: a nationwide propensity score-matched cohort study	Critical Care Medicine	50(6)	977-985	2022
Ohbe H, Sasabuchi Y, Kumazawa R, Matsui H, Yasunaga H	Intensive care unit occupancy in Japan, 2015–2018: a nationwide inpatient database study	Journal of Epidemiology	32(12)	535-542	2022
Ohbe H, Sasabuchi Y, Matsui H, Yasunaga H	Impact of the COVID-19 pandemic on critical care utilization in Japan: a nationwide inpatient database study	Journal of Intensive Care	10(1)	51	2022
Ohbe H, Tagami T, Ogura T, Matsui H, Yasunaga H	Low-flow duration and outcomes of extracorporeal cardiopulmonary resuscitation in adults with in-hospital cardiac arrest: a nationwide inpatient database study	Critical Care Medicine	50(12)	1768-1777	2022
Ohbe H, Tagami T, Uda K, Matsui H, Yasunaga H	Incidence and outcomes of in-hospital cardiac arrest in Japan 2011–2017: a nationwide inpatient database study	Journal of Intensive Care	10(1)	10	2022
Ohisa K, Yamana H, Morita K, Matsui H, Fushimi K, Yasunaga H	Association between subcutaneous steroid injection for extravasation of vesicant anticancer drugs and skin ulcers requiring surgery	European Journal of Oncology Nursing	58	102119	2022
Okada A, Hashimoto Y, Goto T, Yamaguchi S, Ono S, Ikeda Kurakawa K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T	A machine learning-based predictive model to identify patients who failed to attend a follow-up visit for diabetes care after recommendations from a national screening program	Diabetes Care	45(6)	1346-1354	2022
Okada A, Yamana H, Yamaguchi S, Kurakawa K, Michihata N, Matsui H, Fushimi K, Nangaku M, Yamauchi T, Yasunaga H, Kadowaki T	Effect of branched-chain amino acid infusion on in-hospital mortality of patients with hepatic encephalopathy and endstage kidney disease: a retrospective cohort study using a national inpatient database	Journal of Renal Nutrition	32(4)	432-440	2022
Okada A, Yasunaga H	Prevalence of non-communicable diseases in Japan using a newly developed administrative claims database covering young, middle-aged, and elderly people	JMA Journal	5(2)	190-198	2022
Okubo Y, Nishi A, Uda K, Miyairi I, Michihata N, Kumazawa R, Matsui H, Fushimi K, Yasunaga H	Financial Incentives for Infection Prevention and Antimicrobial Stewardship to Reduce Antibiotic Use: A Japan's Nationwide Observational Study	Journal of Hospital Infection	131	89-98	2022
Ono S, Ishimaru M, Yokota I, Konishi T, Okada A, Ono Y, Matsui H, Itai S, Yonenaga K, Tonosaki K, Watanabe R, Hoshi K, Yasunaga H	Risk of post-extraction bleeding with direct oral anticoagulant compared with warfarin: Retrospective cohort study using large scale claims data in Japan	Thrombosis Research	222	24-30	2022
Sakamoto T, Fujiogi M, Ishimaru M, Matsui H, Fushimi K, Yasunaga H	Comparison of postoperative infection after emergency inguinal hernia surgery with enterectomy between mesh repair and non-mesh repair: a national database analysis	Hernia	26(1)	217-223	2022
Sato M, Yamana H, Ono S, Ishimaru M, Matsui H, Yasunaga H	Trends in prophylactic antibiotic use for tooth extraction from 2015 to 2018 in Japan: an analysis using a health insurance claims database	Journal of Infection and Chemotherapy	28(4)	504-509	2022
Shibahashi K, Konishi T, Ohbe H, Yasunaga H	Cost-effectiveness analysis of termination-of-resuscitation rules for patients with out-of-hospital cardiac arrest	Resuscitation	180	45-51	2022
Shibahashi K, Ohbe H, Yasunaga H	Association between intensive care unit admission practices and outcomes in patients with isolated traumatic subarachnoid hemorrhage: A nationwide inpatient database analysis in Japan	Neurocritical Care	37(2)	497-505	2022
Shibahashi K, Ohbe H, Yasunaga H	Urban-rural inequalities in care and outcomes of severe traumatic brain injury: A nationwide inpatient database analysis in Japan	World Neurosurgery	163	e628-e634	2022
Shibahashi K, Ohbe H, Yasunaga H	Adjuvant oral tranexamic acid and reoperation after burr-hole surgery in patients with chronic subdural hematoma: Propensity score-matched analysis using a nationwide inpatient database	Journal of Neurosurgery	138(2)	430-436	2022

Shigemi D, Tabuchi T, Okawa S, Yasunaga H	Association between health literacy and COVID-19 prevention behaviors among pregnant and postpartum women	The Journal of Maternal-Fetal & Neonatal Medicine	35(25)	9971-9977	2022
Sugawara C, Yamana H, Sasagawa E, Yonezawa K, Hikita N, Morita K, Matsui H, Yasunaga H, Haruna M	Factors associated with surgical treatment in postpartum women with mastitis or breast abscess: A retrospective cohort study	Breastfeeding Medicine	17(3)	233-238	2022
Sugihara T, Kamei J, Yasunaga H, Sasabuchi Y, Fujimura T	Prescription of Choreito, a Japanese Kampo medicine, with antimicrobials for treatment of acute cystitis: a retrospective cohort study	Antibiotics	1.736111		2022
Suzuki J, Sasabuchi Y, Hatakeyama S, Matsui H, Sasahara T, Morisawa Y, Yamada T, Fushimi K, Yasunaga H	Additional effect of azithromycin over β -lactam alone for severe community-acquired pneumonia-associated acute respiratory distress syndrome: a retrospective cohort study	Pneumonia	14(1)	1	2022
Suzuki T, Michihata N, Yoshikawa T, Saito K, Matsui H, Fushimi K, Yasunaga H	Low- versus high-concentration intravenous immunoglobulin for children with Kawasaki disease in the acute phase	International Journal of Rheumatic Diseases	25(5)	576-583	2022
Suzuki Y, Kaneko H, Okada A, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I	Impact of Glucose Tolerance on Incident Proteinuria in Prediabetes: Analysis of a Nationwide Population-Based Dataset	American Journal of Nephrology	53(4)	307-315	2022
Suzuki Y, Kaneko H, Okada A, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Yano Y, Kamiya K, Matsunaga A, Ako J, Fukui A, Yokoo T, Nishiyama A, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I	Risk for Proteinuria in Newly Defined Hypertensive People Based on the 2017 American College of Cardiology/American Heart Association Blood Pressure Guideline	American Journal of Cardiology	168	83-89	2022
Suzuki Y, Kaneko H, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I	Comparison of cardiovascular outcomes between SGLT2 inhibitors in diabetes mellitus	Cardiovascular Diabetology	21(1)	67	2022
Suzuki Y, Kaneko H, Okada A, Itoh H, Morita K, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Kamiya K, Matsunaga A, Ako J, Fukui A, Nishiyama A, Yokoo T, Node K, Yamauchi T, Nangaku M, Yasunaga H, Komuro I	Change in Cardiovascular Health Metrics and Risk for Proteinuria Development: Analysis of a Nationwide Population-Based Database	American Journal of Nephrology	53(2-3)	240-248	2022
Suzuki Y, Kaneko H, Okada A, Matsuoka S, Itoh H, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yamaguchi S, Node K, Yamauchi T, Yasunaga H, Komuro I	Prediabetes in Young Adults and Its Association with Cardiovascular Health Metrics in the Progression to Diabetes	Journal of Clinical Endocrinology & Metabolism	107(7)	1843-1853	2022
Suzuki Y, Kaneko H, Okada A, Nangaku M, Yasunaga H, Komuro I	Kidney outcomes in patients with diabetes mellitus did not differ between individual sodium-glucose cotransporter-2 inhibitors	Kidney International	102(5)	1147-1153	2022
Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Kamiya K, Matsunaga A, Ako J, Node K, Yasunaga H, Komuro I	Glycemic status and the association of change in blood pressure with incident cardiovascular disease	American Heart Journal	254	48-56	2022
Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Matsunaga A, Node K, McEvoy JW, Lam CSP, Oparil S, Yasunaga H, Komuro I	Association of Cardiovascular Health Metrics with Risk of Transition to Hypertension in Non-Hypertensive Young Adults	American Journal of Hypertension	35(10)	858-866	2022
Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Matsuoka S, Fujii K, Yamaguchi S, Michihata N, Jo T, Takeda N, Morita H, Node K, Kim HC, Viera AJ, Oparil S, Yasunaga H, Komuro I	Age-Dependent Relationship of Hypertension Subtypes with Incident Heart Failure	JAHA	11(9)	e025406	2022
Suzuki Y, Kaneko H, Yano Y, Okada A, Itoh H, Ueno K, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Yokota I, Node K, Yasunaga H, Komuro I	Dose-dependent relationship of blood pressure and glycemic status with risk of aortic dissection and aneurysm	European Journal of Preventive Cardiology	29(18)	2338-2346	2022

Tagami T, Yamakawa K, Endo A, Hayakawa M, Ogura T, Hirayama A, Yasunaga H	Japanese Multicenter Research of COVID-19 by Assembling Real-world Data: A Study Protocol	Annals of Clinical Epidemiology	4(3)	92-100	2022
Taguchi R, Shigemi D, Yasunaga H	Trends in antihypertensive prescription for pregnant women with hypertension and their peripartum outcomes before and after label and guideline revisions in Japan	Hypertension Research	45(11)	1823-1831	2022
Tahara S, Hattori Y, Aso S, Uda K, Kumazawa R, Matsui H, Fushimi K, Yasunaga H, Morita A	Trends in Surgical Procedures for Spontaneous Intracerebral Hemorrhage in Japan	Journal of Stroke and Cerebrovascular Diseases	31(9)	106664	2022
Takeuchi Y, Iwagami M, Ono S, Michihata N, Uemura K, Yasunaga H	A post-marketing safety assessment of COVID-19 mRNA vaccination for serious adverse outcomes using administrative claims data linked with vaccination registry in a city of Japan	Vaccine	40(52)	7622-7630	2022
Ueno K, Kamiya K, Kaneko H, Okada A, Itoh H, Fujii K, Takeda N, Morita H, Michihata N, Jo T, Yasunaga H, Komuro I	Acute-phase initiation of cardiac rehabilitation for improvement in activities of daily living in patients hospitalized for acute heart failure	Journal of Cardiovascular Development and Disease	9(4)	97	2022
Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Node K, Yamauchi T, Yasunaga H, Komuro I	Relationship of normal-weight central obesity with the risk for heart failure and atrial fibrillation: Analysis of a nationwide health check-up and claims database	European Heart Journal Open	2(3)	oeac026	2022
Ueno K, Kaneko H, Kamiya K, Itoh H, Okada A, Suzuki Y, Matsuoka S, Fujii K, Michihata N, Jo T, Takeda N, Morita H, Ako J, Yasunaga H, Komuro I	Clinical utility of simple subjective gait speed for the risk stratification of heart failure in a primary prevention setting	Scientific Reports	12(1)	11641	2022
Urushiyama H, Jo T, Hasegawa W, Yokoyama A, Ando T, Sakamoto Y, Kumazawa R, Uda K, Michihata N, Awano N, Hiroki M, Fushimi K, Yasunaga H, Nagase T	Effect of nintedanib on acute exacerbations of fibrosing interstitial lung diseases: A national database study in Japan	ERJ Open Research	8(4)	00209-2022	2022
Wada Y, Takahashi H, Matsui H, Yasunaga H, Fujiwara H, Sasabuchi Y	Adhesion barriers and intraperitoneal or uterine infections after cesarean section: A retrospective cohort study	Surgery	172(6)	1722-1727	2022
Yamana H, Sasabuchi Y, Matsui H, Michihata N, Jo T, Yasunaga H	Maoto plus neuraminidase inhibitor versus neuraminidase inhibitor alone for reducing hospitalisation in older adults with seasonal influenza	Annals of Clinical Epidemiology	4	120-128	2022
Yamana H, Tsuchiya A, Horiguchi H, Morita S, Kuroki T, Nakai K, Nishimura H, Jo T, Fushimi K, Yasunaga H	Validity of a model using routinely collected data for identifying infections following gastric, colorectal, and liver cancer surgeries	Pharmacoepidemiology and Drug Safety	31(4)	452-460	2022
Yamazaki R, Ohbe H, Matsuda Y, Kito S, Shigeta M, Morita K, Matsui H, Fushimi K, Yasunaga H	Early electroconvulsive therapy in patients with bipolar depression: a propensity score-matched analysis using a nationwide inpatient database	Journal of Affective Disorders	312	245-251	2022
Zhang L, Michihata N, Matsui H, Fushimi K, Yasunaga H, Tsuda Y, Tanaka S, Kobayashi H	Preoperative Arterial Embolization and Wound Complications After Resection of Malignant Bone Tumor in the Pelvis: A Nationwide Database Study	Japanese Journal of Clinical Oncology	52(10)	1176-1182	2022

厚生労働大臣 殿

機関名 国立大学法人東京大学

所属研究機関長 職名 学長

氏名 藤井 輝夫

次の職員の令和4年度 厚生労働行政推進調査事業費補助金 の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究
- 研究者名 （所属部署・職名）大学院医学系研究科・教授
（氏名・フリガナ）康永 秀生・ヤスナガ ヒデオ
- 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	東京大学	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容：)

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣 殿

機関名 京都大学

所属研究機関長 職 名 医学研究科長

氏 名 伊佐 正

次の職員の令和4年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究
- 研究者名 （所属部署・職名） 大学院医学研究科・教授
（氏名・フリガナ） 中山健夫・ナカヤマタケオ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容：)

（留意事項） ・該当する口にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和5年3月8日

厚生労働大臣 殿

機関名 国立大学法人筑波大学

所属研究機関長 職名 学長

氏名 永田 恭介

次の職員の令和4年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究
- 研究者名 （所属部署・職名） 医学医療系・教授
（氏名・フリガナ） 田宮菜奈子・タミヤナナコ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査の場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容：)

（留意事項） ・ 該当する□にチェックを入れること。
・ 分担研究者の所属する機関の長も作成すること。

令和5年4月4日

厚生労働大臣 殿

機関名 国立大学法人東京大学

所属研究機関長 職名 学長

氏名 藤井 輝夫

次の職員の令和4年度厚生労働行政推進調査事業費補助金の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 診療現場の実態に即した医療ビッグデータ（NDB等）を利活用できる人材育成促進に資するための研究
- 研究者名 （所属部署・職名）大学院医学系研究科・特任准教授
（氏名・フリガナ） 笹淵 裕介・ササブチ ユウスケ
- 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称：）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容：)

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。