

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／特定教授

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課題と可能性について検討を進める。ここでは 2025 年度の活動の概況を報告する。

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、バイジャンネットワーク等)、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

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。2023 年度はコースを実施せず、過去 3 年間を総括するとともに、コースで得られた知見について、日本薬剤疫学会（2023 年 11 月）、医療情報学連合大会（2023 年 11 月）、21 世紀 先端医療コンソーシアム 医療マネジメント部会 特別プログラム（2023 年 1～3 月）にて報告した。2024 年度は 3 社、3 名が参加した。

2025 年は 3 社 4 名が参加し、前年度に引き続きオンデマンド動画を活用した実質的なコース期間短縮を継続し、病院見学や実習・演習の分量を維持したこともあり、参加者の満足度は保たれていた。加えて、参加団体のうちのひとつであり、自前で保有する大規模医療データの活用が課題となっている社会保険診療報酬支払基金（以下、支払基金）と何度かコンテンツについての意見交換を行い、有用性が高いと思われる事項について複数回の出張講義を行った。講義で用いたスライドの一部を図 1, 2 に示す。出張講義では、医療現場から生み出される各種データの有用性や、“データの綾”について、より多くの職員が理解を深める機会を提供することができた。2026 年度は今年度の実績をもとに、出張形式での講義を充実させる予定である。

さらに、支払基金と同様に自前で保有する大規模医療データを保有し、その活用を課題としている全国健康保険協会（協会けんぽ）と国民健康保険中央会とも意見交換する機会があり、前述の出張

講義の内容の一部を各団体のニーズに合わせて改変し、各団体でデータヘルスを担当する多くの職員に向けて講義を行った。講義の内容は、支払基金への出張講義の内容をもとに、より汎用性のある内容とした。

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。

・人材育成プログラムに参加することで、その者が通常業務に就いていたことで得られたはずの派遣元への利益を上回る貢献を派遣元にもたらす可能性を、できる限り可視化することの重要性。

・人材育成プログラムに参加した者が、派遣元に戻った後も継続的に医療データサイエンティストとして活躍できるために、医療データに関連する具体的な研究事業が派遣元に存在することの重要性。

・少人数を対象とした数ヶ月にわたる集中的な教育コースだけでなく、多人数を対象とした短期コースにより、各団体のニーズに合わせた内容をピンポイントで提供することの重要性。

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

- 1: Shimamoto T, Tateyama Y, Kobayashi D, Yamamoto K, Nishioka N, Takahashi Y, Ueshima H, Sasaki K, Kiyohara K, Nakayama T, Iwami T. Observational Study on Actual Cancer Screening Participation and Outcomes Among Patients with Lung Cancer Based on Linkage of Cancer Registry and Kyoto City Integrated Database Data from 2014 to 2018. Int J Environ Res Public Health. 2025 Oct 21;22(10):1595.
- 2: Takahashi N, Takahashi Y, Murase K, Nagata K, Nakatani Y, Hamada S, Sunadome H, Togawa J, Hirai T, Chin K, Nakayama T, Sato S. Association

- between nonrestorative sleep status and sleep apnea syndrome: A cross-sectional and longitudinal study using health check-up and health insurance claims data in Japan. *Respir Investig*. 2025 Sep;63(5):771–779.
- 3: Takahashi N, Takahashi Y, Murase K, Nagata K, Nakatani Y, Hamada S, Sunadome H, Togawa J, Hirai T, Chin K, Nakayama T, Sato S. Associations among nonrestorative sleep status, sleep apnea syndrome, and major adverse cardiac and cerebrovascular events: health check-up and claims data in Japan. *Sleep*. 2026 Feb 10;49(2):zsaf290.
- 4: Tatsumoto M, Hirata K, Nakayama T, Yamato K, Nakamichi N, Inuyama L. Association between seasons with substantial atmospheric pressure change and migraine occurrence: a retrospective cohort study using Japanese claims data and meteorological data. *Front Neurol*. 2025 Sep 10;16:1600822.
- 5: Takeshima T, Nakayama T, Sano H, Koga N, Matsukawa M. Association of Migraine Comorbidities with Quality of Life, Work Productivity and Daily Activities: Survey and Medical Claims Data in Japan. *Adv Ther*. 2025 Aug;42(8):3839–3860.
- 6: Morimoto K, Moriwaki K, Nishikawa Y, Maeda T, Toyama M, Agatsuma N, Kojima S, Utsumi T, Takahashi Y, Shimosuma K, Nakayama T. Comparative Cost-Effectiveness Analysis of Multiple First-Line Treatments for HER2-Negative Unresectable Advanced or Recurrent Gastric Cancer in Japan. *Pharmacoecon Open*.
- 7: Tabara Y, Shoji-Asahina A, Sato Y, Kawaguchi T, Nakayama T, Matsuda F. Prognostic Significance of the Estimated Small Dense Low-Density Lipoprotein Cholesterol Levels in Predicting the Incidence of Coronary Artery Disease: The Nagahama Study and Shizuoka Kokuho Database Study. *J Atheroscler Thromb*. 2026 Mar 6.
- 8: Agatsuma N, Utsumi T, Inoue T, Tanaka Y, Nishikawa Y, Horimatsu T, Nakanishi Y, Nikaido M, Seta T, Hoshino N, Takahashi Y, Nakayama T, Seno H. Development and Validation of a Claims-Based Algorithm for Identifying Incident Colorectal Cancer and Determining Progression Phases. *JCO Clin Cancer Inform*. 2025 Oct;9:e2500107.
- 9: Ishikawa T, Kako A, Sato J, Hattori J, Yoshimoto H, Kitsuregawa M, Goda K, Takahashi Y, Nakayama T, Mitsutake N. Impact of the COVID-19 pandemic on continuity of medical treatment for patients with chronic diseases in Japan: a retrospective cohort analysis. *BMC Health Serv Res*. 2025 May 19;25(1):721.
- H. 知的所有権の取得状況
なし

レセプト解析結果の解釈に生じる揺れの原因

要素	項目	根拠となる 主な論理	解釈に揺れが生じるプロセス
請求項目	① 物品	・ 使用の有無	・ 使用した医薬品や、医療器材に関する請求が相当する ・ 最も信頼度が高いが、査定による影響を受けることも
	② 基本料 加算 管理料等 (技術及びサービス)	・ 診療現場 での技術や サービスの 提供	・ 診療報酬の算定要件に合致していることが「大前提」 ・ 「算定がない＝サービスを提供していない」とは必ずしもならない
診療報酬 請求の根拠	③ 傷病名	・ 臨床医療	・ 多くの場合、「医療従事者の側の論理」で傷病名が付与される ・ 査定を見越した「レセプト病名」が付与される場合も
①～③への 解析結果の 解釈	④ 一般常識	・ 日常の社会 生活	・ 上記①～③は、診療行為を「診療報酬請求」の文脈で表現された情報となるが、その情報の分析結果を解釈する際、あくまでもデータを二次利用しているという認識が十分でない場合、誤った解釈に至る可能性がある ・ ①～③のいずれにも、かかってくる

KYOTO UNIVERSITY

17

図 1 支 払 基 金 出 張 講 演 ス ラ イ ド 1

文献抄読のポイント

論文に何が書かれているか？（書かれていないか？）

Method と Results: 必要な事項が書かれているか（結果を再現できるか）

Discussion: 言い過ぎていないか、研究の限界が適切に書かれているか

内的妥当性（≡ バイアスの小ささ）→ 因果いえそう？

測定したデータから推定した因果関係のバイアスがどれだけ小さいか

1. 選択バイアス: 対象者選択は適切か？
2. 情報バイアス: レセプトを用いた操作的定義と、実際の定義に相違はないか？
3. 交絡: 測定されていない交絡（未測定交絡）は何か？

外的妥当性（≡ 一般化可能性の大きさ）→ 誰に適用できそう？

NDBなどの大規模レセプトデータ: 一般的に高い

KYOTO UNIVERSITY

15

京都大学

図 2 支 払 基 金 出 張 講 演 ス ラ イ ド 2