

特集：ビッグデータと保健医療の未来**ビッグデータが支える保健医療の未来：
人工知能と共存する健康行動がもたらす価値**

田倉 智之*

抄 録

機械学習などの人工知能（Artificial intelligence: AI）は、医療介護系のビッグデータとの相乗効果が大きいツールであり、Dx（Digital transformation）などの導入が進む保健医療分野における発展も期待されている。近年の医師の働き方改革や医療介護を取巻く社会経済の潮流を眺めると、今後のヘルスケアサービスの提供には、より効率的な仕組みが不可欠であると考えられる。そのためにも、情報（ビッグデータ）の共有化や流通化により、生産性の向上と供給量の拡大、品質の維持・改善が希求される。このような中、AIによる画像診断補助ソフトが公的保険で評価されるようになり、またビッグデータとAIを応用したアドヒアランスによる臨床経済的な長期予測モデルの開発がなされるなど、当該領域においても幾つかの胎動が垣間見られる。一方で、データサイエンスを保健医療分野で普及させるためには、複数の課題も挙げられる。例えば、学術研究や保険制度で利用が進むリアルワールドデータや深層学習に着目すると、「情報の品質担保」の課題を中心に一定の社会リスクも想像される。その対策として、「医療ビッグデータの品質評価ガイドランス（標準化など）」の整備などが考えられる。当該領域は、その発展により大きな社会経済的な価値を創造すると推察されるため、社会全体の理解（受容）の促進や必要な環境整備を進めつつ、適切に育まれることが望まれる。

キーワード：データベース研究，機械学習，データサイエンス，アドヒアランス，医療価値，予測モデル，保健事業，診断補助，品質保証

社会保障研究 2024, vol.9, no.1, pp.100-114.

I データサイエンス（ビッグデータ、AI）を活用した保健医療の将来像

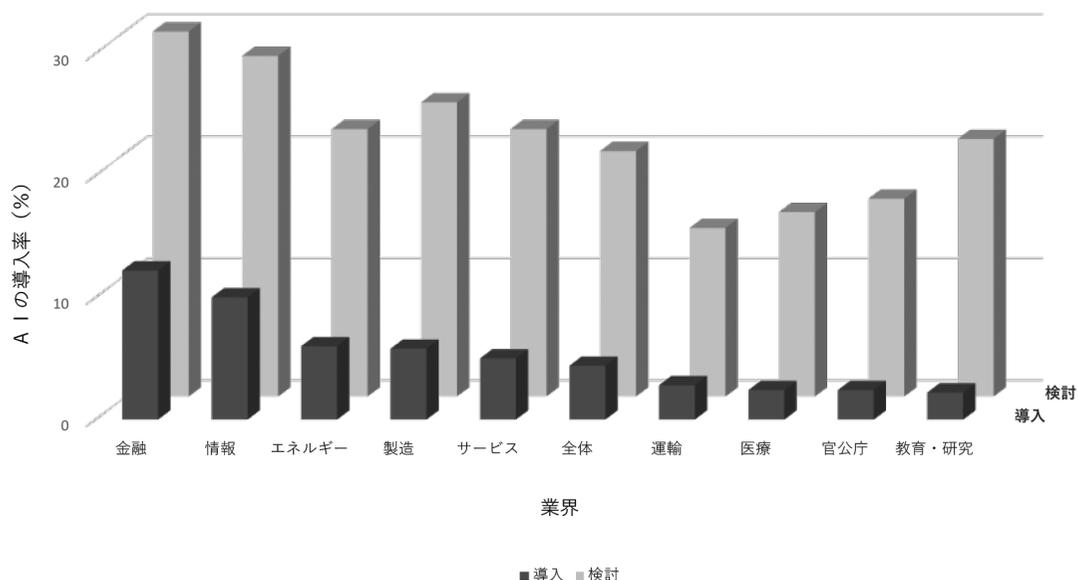
本章では、データサイエンス（ビッグデータ、AI）について、最初に技術的な動向を中心に整理をする。続いて、医療介護分野におけるそれらの導入や普及、さらには議論の概況について解説を

する。最後に、データサイエンス（ビッグデータ、AI）を活用した保健医療の将来像について、主要な潮流を踏まえつつ参考イメージを提供する。

1 データサイエンスの進展

データサイエンス（Data science）とは、データから有益な知見や価値を得るための学問分野である〔Smyth Padhraic et al. (2017)〕。この学問分野

* 日本大学 医学部社会医学系 医療管理学分野
東京大学大学院医学系研究科 医療経済政策学講座



出典：企業の人工知能（AI）導入実態調査。MM総研。2018。

図1 人工知能技術の導入の割合（国内における業種別）

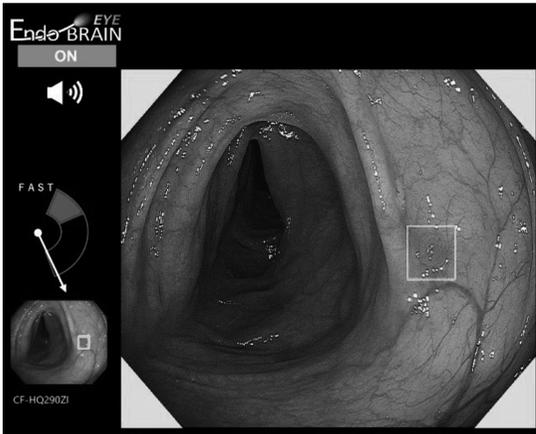
は、医学統計や物理統計、情報工学、計算科学、いわゆるアルゴリズムなどのさまざまな分析手法・解析手法を応用することで、学際領域のみならず社会経済、創薬開発、病院運営などの課題の解決や成長・発展に役立てられることが期待される。データサイエンスは、近年のコンピューター能力の向上や環境の整備に伴い、大規模なデータセットの前処理が可能になったことで、より広く利用され独立した位置付けとなっている。

特に、機械学習などの人工知能（Artificial intelligence: AI）については、医療介護系のビッグデータとの相乗効果が大きいツールであり、Dx（Digital transformation）などの導入が進む保健医療分野での発展も期待されている（図1）。一方で、AIを含むデータサイエンスの定義などについては、専門家の間でも多様な意見があり、論述の視点や対象の領域（テーマ）によって異なる場合もある。すなわち、データサイエンス自体は、さらなる進展の過程にあるとも見なされる。また、これは多くの領域へ横断的に活用が可能であり、社会システムに対して幅広く貢献が期待される分野と推察される。

例えば、PER（Prioritized experience replay）技術を用いた強化学習は試行錯誤を通じて学習するものであるが、ゲーム理論、資源最適化などの多岐に渡る問題に対処する能力が高く、予防医学や社会医学などの分野へ応用し発展が期待される〔Hannes H Loeffler et al. (2024)〕。また、大量で複雑な構造のため処理負荷が高い医療ビッグデータの社会実装においては、ハイパーパラメータチューニングなどの一部を自動化する自動機械学習の導入も望まれる。

2 医療のデータサイエンス

保健医療分野においても、世界的にデータサイエンスへの期待は高く、ビッグデータを中心に応用が進んでいる。ただし、その多くはまだ学際的な側面が強く、例えばビッグデータはデータベース研究として盛んに利用されているが、政策支援や製品開発に対して恒常的かつ本格的な運用はこれからと想像される。また、AIについても診断支援の製品（医療機器プログラム：SaMDを含む）が散見しつつあるが、まだ広く普及するに至ってはいない状況にある。



出典：サイバネットシステム株式会社プレスリリース資料：
<https://www.cybernet.co.jp/company/about/news/press/2024/20240214.html>

図2 大腸内視鏡診断支援AIプログラムの例（診療報酬上の加算対象）

一方で、不正画像や剽窃行為のチェックも含む学術誌における査読システムについては、パターン認識の進歩が著しいAIや大量に蓄積された科学関連のデータが社会実装として段階的に応用され始めている〔Douglas Heaven (2018)〕。このような中、米国ではAIによる脳卒中や糖尿網膜症の診断支援が2020年度よりメディケアなどの保険償還の対象となっている〔Michael D Abràmoff et al. (2022)〕。わが国においても、データサイエンスは臨床や予防の分野で徐々に認知が進んでいる。

例えば、AIによる画像診断補助ソフト（図2）が2022年度より公的保険における「画像診断管理加算」で評価されるようになり、またビッグデータとAIを応用したアドヒアランス（健康関連行動）による臨床経済的な長期予測モデルの開発がなされる〔Tomoyuki Takura et al. (2021)〕など、当該領域においても幾つかの胎動が垣間見られる。なお、診断補助については、癌や難病の症例を対象にAIによる遺伝子解析の補助に対する保険適用の議論も将来的に想像される。また、疾病予防分野のデータサイエンスによる長期予測モデルについては、保健事業への実装により社会経済的な価値の創出も期待される。その詳細については、次章

で紹介を行うことにする。

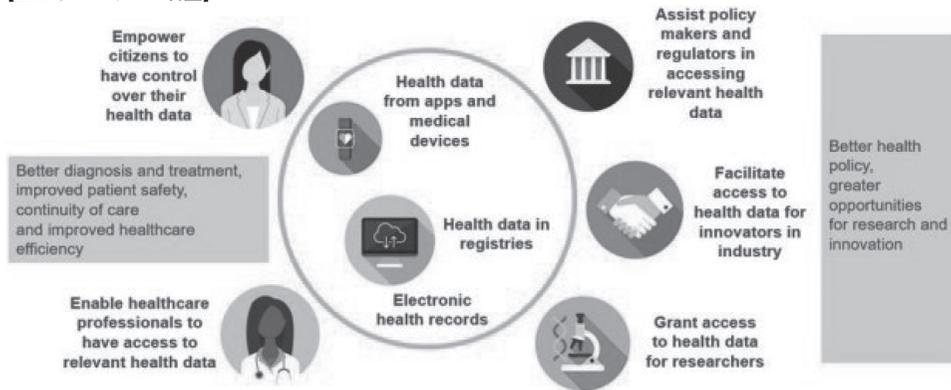
3 データサイエンスを活用した保健医療の未来

データサイエンスは、それ自体の定義が定まらずさらなる変貌も予想される中、技術革新の動向から長期的にはより一層の発展が期待される。そのため、現在のそれらの概念をもとにした保健医療の未来を論じるのは難しいのも事実である。ただし、データサイエンスを活用した保健医療の未来は、より健康的で安心・快適な社会を目指す世間一般の動機付けとベクトルが同じであるため、政府などが同様のテーマで検討を進めた将来ビジョンなどが参考になるとと思われる。例えば、「Society 5.0」〔内閣官房 (2023)〕の「ビッグデータ連携がもたらす未来社会像」が、一つのイメージになるとと思われる。

保健医療の特性から多少肉付けを行うならば、キーワードとして「安心の醸成」「疾病を予測」「行動の変容」「自然で簡便」の4つが挙げられる。すなわち、ビッグデータとAIを駆使して健康管理（リスク評価も含む）を自然体で恒常的に行い、必要に応じてメディカルチェックや受診勧奨（受診方法の推奨も含む）を促す仕組みが想像される。また、自分や家族の心身不調や不安に対して、個々人の置かれた環境やリテラシーなどにそっていつでもどこでも相談や助言を受ける機会の充足も想像される。

さらに、意思決定や行動変容のサポートは、ライフスタイルや住居・就業の環境にそって利便性の改善や各負担の軽減を図りつつ、最適化される仕組みも想像される。これらは、プライバシーやインセンティブにも配慮されつつ、利用・参加のハードル（手間やストレス）が低くなるように社会的にデザインがなされ、選択性や継続性に優れた形態で実現がなされることも想像される。なお、このような潮流は、北欧などで先行的なケースが散見する。例えば、ビッグデータの活用目的を絞るならば、欧州連合（EU）のEHDS（European Health Data Space）構想は、前述の未来像の礎として参考イメージになると考えられる

【EHDS ユーザーの利益】



解説：個人は自分のヘルスデータをより細かく管理できるようになる。希望すれば、選択した医療専門家に迅速かつ簡単に透明性のある共通の形式でデータを提供することができる。これにより、不要な検査の数とコストが削減され、医療の安全性が向上する。ヘルスデータにアクセスし、分析し、共有する能力は、医療をより効率的にし、より良い医療上の決定をサポートし、その結果、健康上の成果を向上させる。EHDS は、2030 年までに EU のデジタル変革に関する委員会のビジョン、「100% の市民に医療記録とデジタル原則宣言 40 へのアクセスを提供することを目的としたデジタルコンパス 41 (Digital Compass)」の達成を支援する。これは欧州委員会の提案に基づいており、European Digital Identity Framework と Wallet を利用することで、市民は携帯端末から自分のヘルスデータに国境を越えて信頼できる方法でアクセスできるようになる。

出典：COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL: A European Health Data Space: harnessing the power of health data for people, patients and innovation. EUROPEAN COMMISSION. 2022.

図3 EUにおけるEHDS (European Health Data Space) 構想 (ヘルスケアデータの利活用にかかわる包括的・体系的な制度のあり方)

(図3)。

II ビッグデータとAIを応用したアドヒアランスの臨床経済の予測モデル

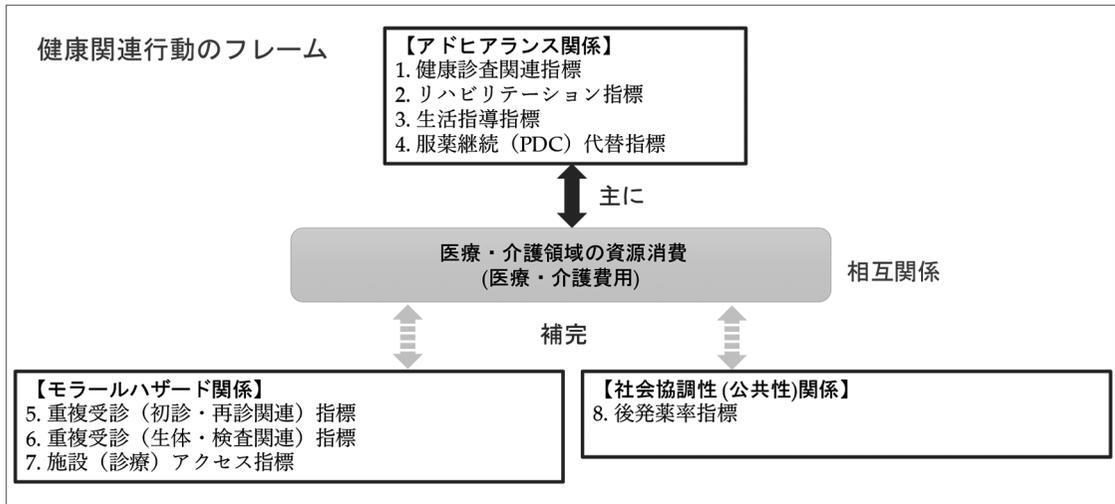
医療介護分野を取り巻く社会的な潮流や保健事業の役割と経営などを背景に、最近、アドヒアランスのコントロールが注目をされている。そこで本章では、保健事業などへの展開を目的に、広義のアドヒアランスを定量化するとともにその長期の臨床経済効果を検証し、かかわる予測モデルを開発した我が国の研究例を紹介する〔Tomoyuki Takura et al. (2021)〕。

1 アドヒアランスによる臨床経済アウトカムの長期予測モデル

将来の医療介護システムの発展に不可欠な視点として、医療価値の認識醸成（その見える化、関

係者間の共有）と健康関連行動（アドヒアランス）の向上が挙げられる〔田倉智之（2023a）〕。一方で、アドヒアランスが臨床経済に及ぼす影響の評価は、多様で複雑な因子が絡み合うため、通常の臨床試験での評価が困難である。

そこで、広義のアドヒアランス（図4）をスコア化して、将来の死亡と費用との関係を長期的に予測するデータサイエンス研究が実施されている。具体的には、医療ビッグデータとAI（機械学習：ニューラルネットワーク）を応用することで、アドヒアランスが長期（48か月間）の医療・介護費用や生命予後、ほかの臨床指標に与える影響を48,456人（循環器領域）のコホートで検証しつつ、予測モデル（10水準のASHROスコア：低い＝良い）を開発している。このスコアは、対象者（被保険者）の将来の臨床経済的なりリスクを予見するため、行政者は保険財政（医療・介護）の管理に、医療者は重症化予防の促進に活用することで、疾



出典：文献 [5]。

図4 広義のアドヒアランス（健康関連行動）の基本構造（複数要因による統合指標：その他に、医学的な患者特性等の要因を含む）

病負担の改善と社会保障の発展に貢献すると期待される。

この後ろ向きのコホート研究は、日本の医療保険制度、介護保険制度、特定健康診査制度の公的事業情報から構成される大規模なデータ（国保データベース：KDB）を研究ソースの中心としている（研究プラットフォームはTheBD：倫理委員会承認済）。観察期間は、2014年4月から2018年3月である。この研究では、保険財政や臨床成果に影響をおよぼす複合指標を統合したアドヒアランスを説明変数（独立変数）に、医療・介護費用を目的変数（従属変数）とする因果関係の検証と予測モデルの開発がなされている。この予測モデルの構築は、ランダムフォレストなどの機械学習により説明変数を調整のうえ、重回帰分析を基本評価としつつ、ロジスティック回帰分析にて補完している。

このような研究では、複雑系の事象の大規模なサンプルについて多変量の解析を行うため、経験則に基づく通常の統計学的なアプローチには、物理的な側面や結果の網羅性において限界がある。そこで研究グループは、データサイエンスの一環でAIを説明変数の選択と統合、および重みづけの

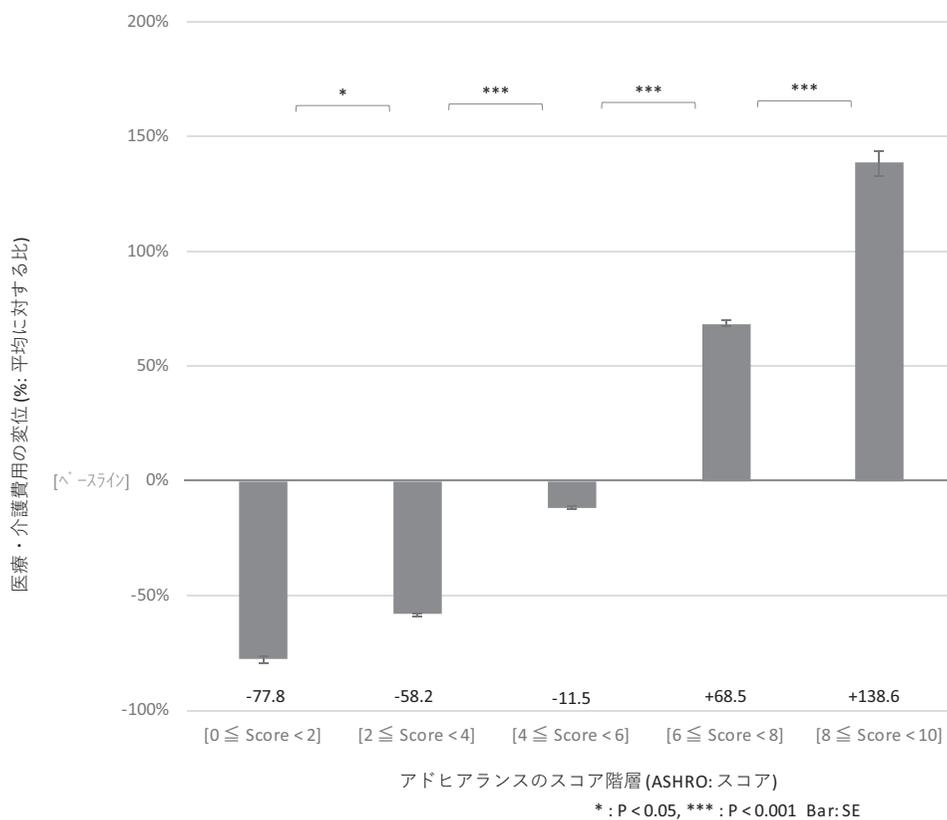
設定に活用している。この利点として、ランダムフォレストの長所は、オーバーフィットの問題を最小限に抑えることにあるうえ、医療ビッグデータなどの数千の入力変数を持つ大規模サンプルに対して効率的に実行できる利点もある。また、異なるデータスケール（例えば、血圧とGFRは異なる正規値を持つ）に対応できること、無関係な変数の包含に対する堅牢性があることなどが挙げられる。

さて、広義のアドヒアランス（ASHROスコア）は、収縮期血圧、LDLコレステロール、HbA1c、eGFRなどのリスク因子とも有意な相関関係を担保しつつスコア化されている。危険因子を揃えた予測モデルの全死亡に対する感度の検証の結果、ASHROスコアの低い群と高い群の間には、3年以上後の累積死亡率に統計学的有意な差が認められている（2% vs. 7%, $p < 0.001$ ：表1）。また、生命予後（全死亡）に対するASHROスコアのオッズ比は、1.860（95%CI: 1.740-1.980, $p < 0.001$ ）である。その予測性（DiscriminationとCalibration）が検証された結果、48か月後の医療・介護費用の変位は、集団の平均を基準に $-77.8\% \sim +138.6\%$ に分布している（図5）。例えば、アドヒアランスが最も

表1 アドヒアランス（ASHROスコア）の長期生命予後への感度（傾向スコア法で患者背景を調整）

指標	単位	ASHRO_50% カットオフ		p value
		低い群	高い群	
サンプル	人	6,154	6,154	
男性（割合）	人（%）	4,307（70%）	4,279（70%）	0.596
年齢	歳	69.2 ± 7.1	69.1 ± 6.2	0.397
BMI	kg/m ²	23.5 ± 3.4	23.5 ± 3.5	0.783
収縮期血圧	mmHg	132.1 ± 14.8	132.2 ± 15.5	0.629
中性脂肪	mg/dL	125.3 ± 74.3	125.3 ± 74.9	0.965
HbA1c	%	6.0 ± 0.8	6.0 ± 0.8	0.764
血清クレアチニン	mg/dL	0.9 ± 0.8	0.9 ± 0.8	0.490
喫煙（ダミー）	喫煙有:1, 非喫煙:0	0.2 ± 0.4	0.2 ± 0.4	0.796
飲酒（頻度）	回	2.2 ± 0.8	2.2 ± 0.9	0.899
死亡数（率）	人（%）	123（2%）	430（7%）	< 0.001

出典：文献 [5]。



出典：文献 [5]。

図5 ASHROスコアと医療・介護費用変位の関係（将来予測モデル：48か月後）

悪い群（スコア10）は、平均（スコア5）に対して、一定の精度のもとで将来の医療介護の累積費用が140%以上増加することを示している。

2 アドヒアランスを活用する新たな保健介入システムの研究

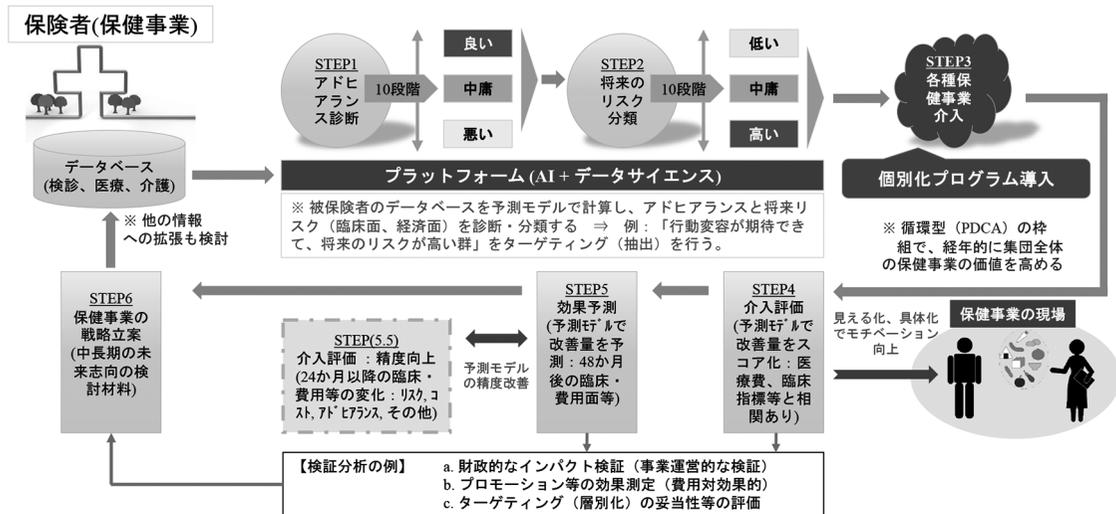
本節では、前節で紹介した広義のアドヒアランスの臨床経済的な意義にかかわる研究成果をさらに発展させるべく、最近立ち上がった科学研究費助成事業などにおけるプラットフォームの整備を紹介しつつ、アドヒアランス応用にかかわる今後の新たなアイデアや仮説検証も概説する〔田倉智之(2023b)〕。なおその仮説は、ビッグデータとAIを用いてアウトカムとの因果推論を検証しアドヒアランスを生成したことに起因しており、研究者個人の経験則による一般の研究アプローチでは気づきにくい観点と考えられる。

保健事業分野の普遍的な課題として、集団検診や標準的なプログラムのみならず、特定保健指導においても要因把握と判断基準の幅広さ、習慣や環境に対する感度が不十分なために、予防介入などの効果が小さく持続しない群が存在することが挙げられる〔杉田由加里他(2014)〕。この背景に

は、多様で複雑なアドヒアランスの存在が考えられる。これは、測定とその要因分解が難しいうえ、その相互干渉や影響度が不明な点が特徴である。また、保険者の予算や体制などの経営面も課題の一つとして想像される。

以上のように、健康寿命の延伸や医療財政の管理には、経済性にも配慮しながら、アドヒアランスを改善させる理論や手法の研究開発が重要になる。これらを背景に、関連する問題意識の解消に向けて、幾つかの萌芽的な取り組みがみられる。医療介護分野では、アドヒアランスに影響を及ぼす要因として、従来から「患者（内的）の因子」「疾病・診療の因子」「社会・経済の因子」の3分類が挙げられている〔勝木達夫(2012)〕。

現在、これらの因子を考慮しつつ、アドヒアランスの構成要素別にアウトカムへの影響を分析し最適化プログラムの開発を目指す、東京大学（大学院医学系研究科）と協会けんぽ（埼玉支部）による共同研究や科学研究費助成事業が進行中である。この研究開発では、得られた知見のフィードバック（機械学習やPDCA）を繰り返しつつ、自己成長型の保健事業評価のプラットフォームも議論されている（図6）。

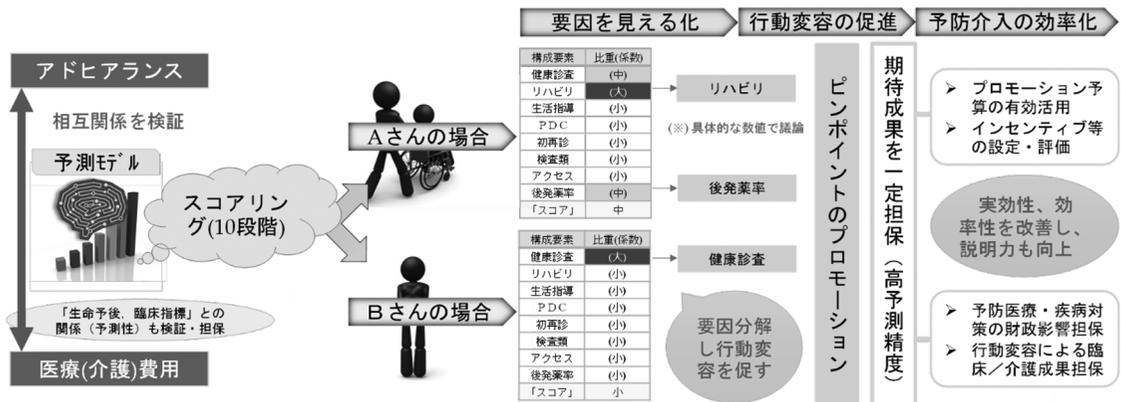


出典：文献 [5]。

図6 新たな健康介入に資する自己成長型のプラットフォームの概念 (AIなどで「予測（ターゲティング）⇒介入⇒学習（検証）⇒反映」の循環型の保健事業の改善モデルを試行)

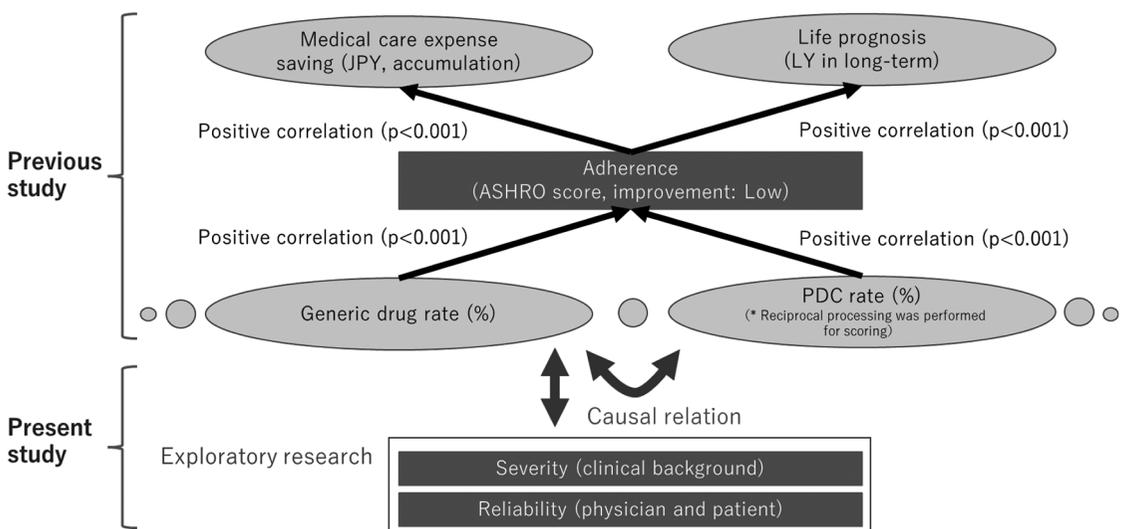
この研究の要は、前述の研究成果（ASHRO）のアドヒアランスの定量化手法と予測モデルを用いて、被保険者のアドヒアランスの診断と将来リスク（臨床面と経済面）の評価を行い、対象者個々に最適な介入プラン（ターゲティング）を優れた

費用対効果で展開することにある。すなわちASHROは、要因を見える化しつつ効果が期待できる要因に傾注する、ピンポイントのプロモーション（Precision medicine的）が潜在的に可能と思われる（図7）。



出典：文献 [5]。

図7 ASHROスコアを用いた要因分解とピンポイントのプロモーションの概念（対象者個々に最適な介入プランを優れた費用対効果で展開する）



(Annotation) Regarding the adherence score (ASHRO score), the lower values indicate greater clinical and economic improvement (scale: 0 [superior] - 10 [inferior]).

注：*：Proportion of days covered (PDC) rate data were based on the actual results, and a PDC score was used to incorporate it into the Adherence Score for Healthcare Resource Outcome (ASHRO) score, each of which had a different meaning in terms of scale and high/low. 出典：文献 [10]。

図8 患者の重症度と患者のアドヒアランス、医師患者の信頼関係が薬剤選択に影響（初期仮説的）

アドヒアランスにかかわるデータサイエンスも含むこれらの知見や環境は、群衆行動（Herding behavior）の一般理論などへ活かすことで、被保険者を効果的かつ効率的に適切な集団特性へ移行させることが期待される。つまり、人間は横並び意識を大なり小なり有しており、所属集団、それも自分と同じ立場の人間の成功に対する感度は高いため、ベンチマークなどを提供することで、アドヒアランスの向上が期待される訳である。

このような視点による研究の潮流として、前節で紹介した研究の続報も期待される。例えば、後発薬のスイッチングに及ぼす影響因子として、患者の重症度とともにアドヒアランスの水準も示唆されている。また、アドヒアランスを背景とした患者と医師の信頼関係も、後発薬のスイッチングに影響を与えるという仮説検証も進められている（図8）〔田倉智之（2023b）〕。以上のように、アドヒアランスにまつわる探求は、医療政策などにおいて大きな意義を有していると推察される。

Ⅲ データサイエンス（ビッグデータ、AI）の発展に必要な社会環境整備

本章では、保健医療の分野へビッグデータやAIの導入を進める意義として社会保障などの潮流を中心に論述しつつ、データサイエンスの普及に向けた課題も整理する。今後のデータサイエンスの保健医療への導入においては、ここで示めされる懸念に対して環境整備などを継続的に進めていくことが望まれる。

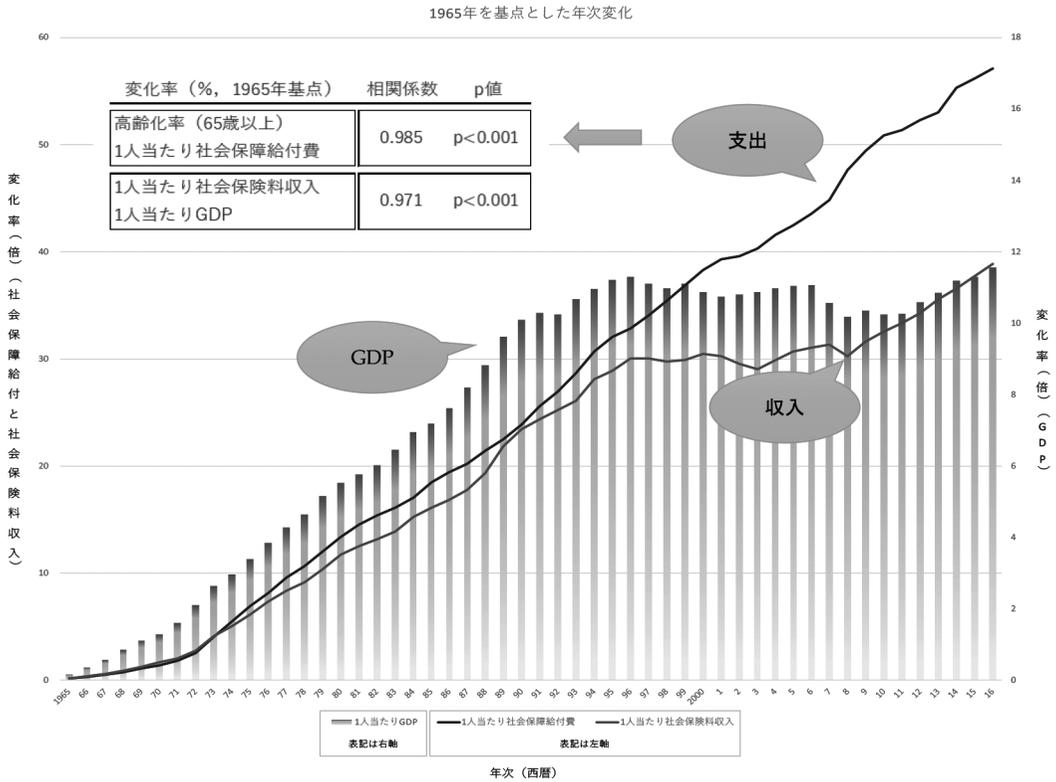
1 ビッグデータやAIの積極導入が望まれる保健医療を取巻く潮流

近年注目を集めるユニバーサルヘルスカバレッジ（UHC）の推進には、社会経済的な要因が大きな影響を及ぼすことも明らかとなっている〔Tomoyuki Takura, Hiroko Miura. (2022)〕。すなわち、医療介護の環境整備において、臨床と経済の調和が望まれている訳である。また、医療介護分野の進歩に影響をおよぼす各種イノベーションも、社会経済的なメカニズム（バリューチェーン）

をとおして、UHCとの関係は深いことが示唆されている〔Tomoyuki Takura et al. (in press)〕。この背景として、次のような医療介護分野を取り巻く潮流が挙げられる。一つ目は、言うまでもなく社会保障財源のひっ迫である（図9）。この背景はしごく簡単で、医療や介護の需要増加（高齢化の進展）と実体経済の伸び悩み（国内総生産〔GDP〕などの停滞）が大きな割合を占める〔田倉智之（2021）〕。二つ目は、診療技術のイノベーションとその超高額化である。この両者を合わせることで、数量増加と単価上昇により、国民医療費などが膨らんでいくことは容易に想像される。

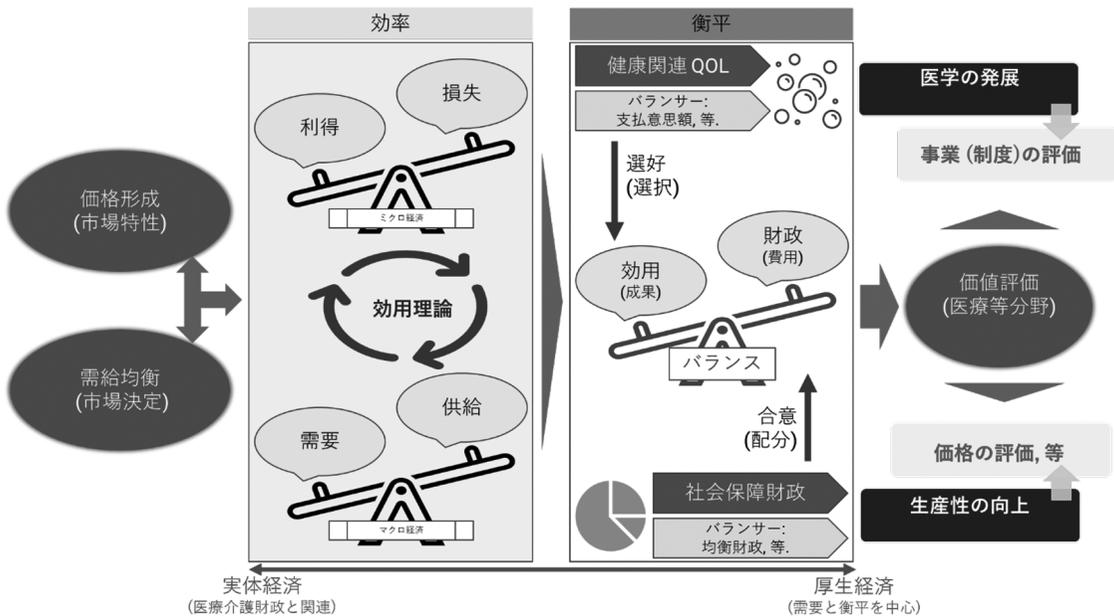
一方で、現役世代の人口減少と経済負担の許容水準から、保険財源の収入が追い付いていないのが実情である。結局のところ、受益と負担のバランス低下が前述の根源にあると考えられる。これらを俯瞰すると、将来の医療介護システムの発展に不可欠な視点は、おのずと明らかになってくる。そのキーワードは、医療価値の認識醸成（その見える化、関係者間の共有）と健康関連行動（アドヒアランス）の向上である。将来の医療システムの安定供給は、アドヒアランスを制することにかかっていると考えられる。すなわち、医療制度において国民の最も重要な財産（価値）を安定供給と見なした場合、アドヒアランスの向上が不可欠になる訳である。

このアドヒアランスの改善の臨床経済の意義およびデータサイエンスの貢献の可能性については、長期予測モデルの開発例を紹介しつつ前章で示した。そこで本節では、医療価値の評価の重要性とそれに対するデータサイエンスの必要性を述べる。公益性の高い医療分野においては、効率性も考慮しつつ衡平性（well-beingなどのバランス）の視点を取り入れ、患者の希望（選好、支払意思）と政府の負担（財政収支、所得再配分）の調和を念頭に公益的な価値を論じる必要がある。したがって医療の価値は、厚生経済学なども背景に、個人と社会の関係をも織りまぜながら、健康プログラム単位あたりの効用（健康成果）と費用（資源消費）の変位のバランスを検討することになる（図10）〔Tomoyuki Takura (2022)〕。



(資料) 社会保障給付費/社会保障料収入：社会保障費用統計 (国立社会保障・人口問題研究所)，GDP：World Economic Outlook Databases (IMF)，高齢化率：我が国の推計人口 (総務省) より作成
出典：文献 [13]。

図9 社会保障の財政収支の影響因子 (高齢化率と実体経済の因子)



出典：文献 [14]。

図10 厚生経済学を考慮した効用理論と費用対効果にもとづく医療の価値評価の概念

この医療価値評価は、費用対効果、すなわちパフォーマンスの向上の視点が含まれている。わが国の政府も、「統合イノベーション戦略(2023)」などのもとで医療Dxなどのヘルスケア分野におけるデータサイエンスの推進を進めているが、この背景の一つとして医療の生産性の向上が位置付けられている。2024年度から開始されている医師の働き方改革、および前述の医療介護分野を取巻く社会経済の潮流を眺めると、医療介護サービスの提供においても、効率性の向上が不可欠であると理解できる。そのためにも、情報(ビッグデータ)の共有化と流通化による生産性の向上、および供給量の拡大、品質の維持・改善が希求される。また、AIや通信技術(オンライン診療)の活用による医療職や介護職の労働負担の軽減、および患者・家族のセルフメディケーションの支援なども、注目度の高い重要なテーマである。これらの展開の結果として、実体経済や健康寿命が増長し、我が国全体の医療経済的な価値も拡大することが期待される。

2 保健医療へビッグデータやAIの導入を進めるにあたっての主な課題

前節では、保健医療分野においてデータサイエンスを活用する意義を述べたが、本節では、想定される課題や今後望まれる環境整備について俯瞰してみる。かかわる主な課題としては、大きく次の4つ挙げられる。一つ目は、個人情報の保護や情報の品質担保などのビッグデータの流通における制約である。二つ目は、データサイエンスの利用を促すためのインセンティブの仕組み(ビジネスモデルなど)である〔Melissa M Chen et al. (2021)〕。三つ目は、医療介護分野を中心に存在するAIなどに対する関係者の不安や抵抗感である。四つ目は、ヘルスケア分野に明るいデータサイエンス関連の技術者や研究者などの人材育成である。いずれもデータサイエンスを保健医療に普及させるために重要なポイントと思われるが、それらの中から特に配慮すべき点について、以降で簡単に解説を試みる。

まず一つ目に挙げた課題の中から、「情報の品

質担保」について整理を行う。仮に医療介護制度などの一翼としてビッグデータ利活用が社会活動の基盤となった場合、そのソースデータの品質や信頼性に問題があると、新たなサービスの展開にかかる時間やコストに悪影響をおよぼすことになる。特に、臨床現場で活用する関連サービスについては、健康上の懸念や対策への負担は大きなものになる。つまり、ヘルスケア分野においてビッグデータの品質保証を論じることの意義は、他産業以上に重要になる訳である。一方で、黎明期ともいえる現状において、政府レベルではデータ品質よりも情報活用(流通と規制、権利:個人情報の取扱など)の議論が多い傾向にある〔内閣官房(2023)〕。このように、最初は普及に向けた環境整備が優先されるものの、データ利活用が進むことにより、将来的にはデータ品質にかかわる議論も増すことが予想される(図11)。

例えば、学術研究や保険制度で利用が進むリアルワールドデータ(RWD)や深層学習に着目すると、幾つかの社会リスクも想像される。Lancetなどの一流医学誌が2020年、新型コロナウイルス感染症(COVID-19)に関する論文を相次いで撤回したのは記憶に新しいところである〔Mandeep R Mehra et al. (2020)〕。これらは、対象患者のビッグデータを分析したものだったが、第三者から疑問の声が上がり、検証を求められた著者がデータの信頼性を証明できなかった。この問題の論点は、RWDを利用した研究が広がりつつある中、こうした意図的なデータのねつ造を見抜けるかどうかということである。さらに未来志向のもとに考察を行うならば、AIなどを用いて巧妙かつリアルに大量の操作がなされると、研究者個人レベルでは見抜くのが難しい場合も想像される。

最近では、RWDの利用は学際領域に留まらず、レギュレーションにおける利用も広がりつつある。例えば、我が国の医療保険制度(費用対効果評価など)における審査プロセスでは、NDB(National Data Base)のみならず商業データベースの直接活用が拡大している。また、診療ガイドラインに基づく審査の議論も進む中、医療ビッグデータに基づく研究報告が診療ガイドラインに引用される場

合、データの品質水準は、間接的に影響をおよぼす可能性も想像される。以上を踏まえ、保健医療分野でデータサイエンスを普及させるにあたり、

ゆくゆくは次に述べる3つの対策（環境整備）も不可欠になると推察される（図11）。まずは「医療健康分野に対する専用の認証機関（公益性担

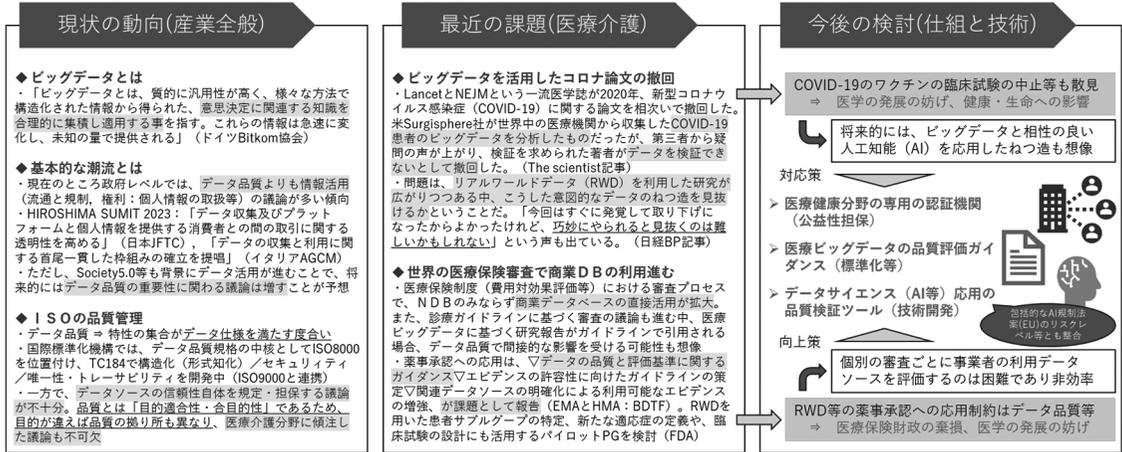
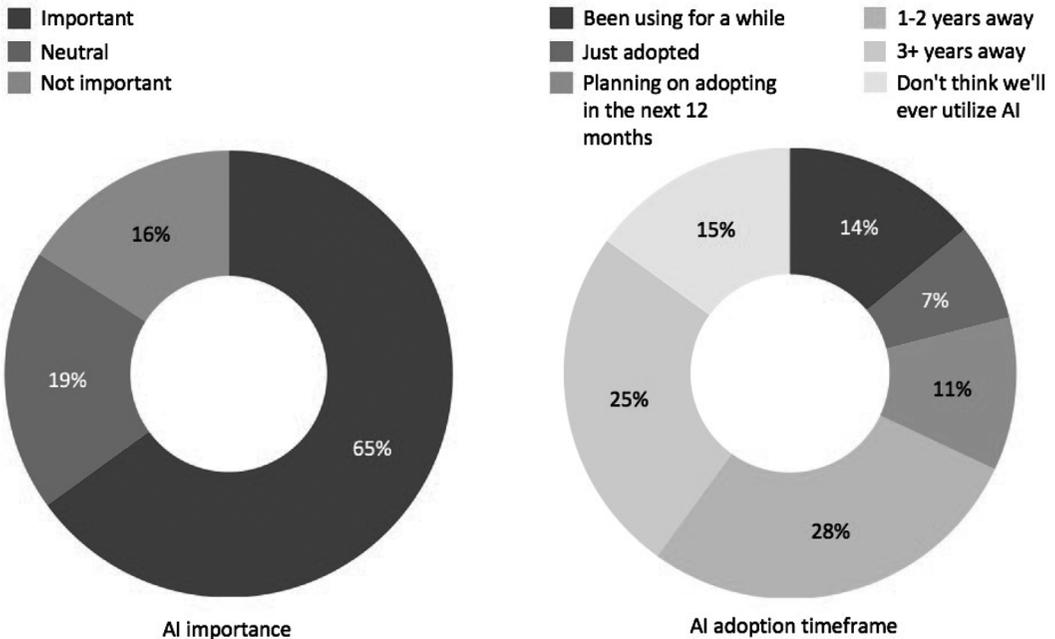


図11 ヘルスケア分野のビッグデータの品質保証の意義

US Providers Recognize AI's Value But Have Been Slow To Adopt



出典：US PROVIDER RECOGNIZE AI'S VALUE BUT HAVE BEEN SLOW TO ADOPT. business insider. 2018: <https://ainow.ai/2018/12/13/158185/>

図12 AIを導入した医療に対する理解と導入の状況 (米国)

保)」の検討, 続いて「医療ビッグデータの品質評価ガイドランス (標準化など)」の整備, さらに「データサイエンス (AIなど) 応用の品質検証ツール (技術開発)」の推進, が挙げられる。

冒頭に挙げた三つ目の課題とも重なるが, 実際のところ, 医療者や患者においては, AIなどに対して不安を感じる方々も少なくないようである。海外の報告ではあるが, 医療へのAIの導入の重要性を理解している医療者の割合が65%であるのに対して, すでに導入している割合は14%と大きな開きがある (図12)。この傾向は, 1章で解説した我が国におけるデータサイエンスの導入の傾向と類似している。その背景の一つとして, AIという技術の特性上, その結果にたどり着いた思考プロセスが分からない「ブラックボックス問題」が想像される。また, 一つ目の課題にも関係するが, AIを活用した技術や結果に対する責任の所在が, 法律や社会のなかで不明なことに起因する抑制も考えられる。さらに, 希少疾患などでデータ数量が不足することによる診断の精度 (正確性) の低下への対策には, ビッグデータを応用しつつも, アルゴリズムの発展が不可欠である。

生成AIシステムが産業や人々の生活に浸透する中, EUの立法機関である欧州議会は, 2024年3月に, 世界初となるAIの包括的な規制法案を可決している。この背景として, 誤情報やフェイクニュース拡散への懸念の高まりも挙げられており, 影響の大きい汎用AIモデルとリスクの高いAIシステムについては, 特定の透明性を満たす義務と, EUの著作権法の順守が必要となっている。我が国においても今後は, このような世界の潮流にも傾注しつつ, ヘルスケア分野のデータサイエンス (ビッグデータ, AI) について社会実装の推進が必要と考えられる。

参考文献

- Smyth Padhraic, Blei David M (2017) “Science and data science,” *Proceedings of the National Academy of Sciences*, Vol.114, No.33, pp.8689-8692.
- Hannes H Loeffler, Jiazhen He, Alessandro Tibo, Jon Paul Janet, Alexey Voronov, Lewis H Mervin, Ola Engkvist (2024) “Reinvent 4: Modern AI-driven generative

- molecule design,” *J Cheminform*, Vol.16, No.1, p.20.
- Douglas Heaven (2018) “AI peer reviewers unleashed to ease publishing grind: A suite of automated tools is now available to assist with peer review but humans are still in the driver’s seat,” *Nature*, Vol.563, pp.609-610.
- Michael D Abràmoff, Cybil Roehrenbeck, Sylvia Trujillo, Juli Goldstein, Anitra S Graves, Michael X Repka, Ezequiel Zeke Silva Iii (2022) “A reimbursement framework for artificial intelligence in healthcare,” *NPJ Digit Med*, Vol.5, No.1, p.72.
- Tomoyuki Takura, Keiko Hirano Goto, Asao Honda (2021) “Development of a predictive model for integrated medical and long-term care resource consumption based on health behaviour: application of healthcare big data of patients with circulatory diseases,” *BMC Med*, Vol.9, No.1, p.15.
- Society 5.0 (2023) 「ビッグデータ連携がもたらす未来社会像」, 内閣官房. https://www8.cao.go.jp/cstp/society5_0/index.html
- 田倉智之 (2021) 『医療の価値と価格—選択と決定の時代へ』, 医学書院。
- (2023a) 「健康関連行動 (アドヒアランス) の医療費用 (医療財政) に与える影響」, 『健康保険』, Vol.77, No.6, pp.12-17。
- (2023b) 「循環器医と医療経済学から見た服薬アドヒアランスの重要性」, 『ランチョンセミナー53』, 第87回日本循環器学会学術集会。
- 杉田由加里, 山下留理子 (2014) 「特定保健指導の展開過程における課題と対応方法」, 『千葉大学大学院看護学研究科紀要』, Vol.37, pp.47-56。
- 勝木達夫 (2012) 「患者アドヒアランスをいかに高め, 維持するのか」, 『心臓』, Vol.44, No.3, pp.286-290。
- Tomoyuki Takura, Hiroko Miura (2022a) “Socioeconomic Determinants of Universal Health Coverage in the Asian Region,” *Int. J. Environ. Res. Public Health*, Vol.19, No.4, p.2376.
- Tomoyuki Takura, Naotsugu Ichimaru, Atushi Aikawa (in press) “Health Economics of Renal Replacement Therapy,” *Updates on Renal Replacement Therapy*, IntechOpen. (ISBN 978-1-83769-175-3)
- Tomoyuki Takura (2022) “Socio-Economic Considerations of Universal Health Coverage: Focus on the Concept of Health Care Value and Medical Treatment Price,” *Health Insurance*, IntechOpen, pp.1-34.
- 統合イノベーション戦略 (2023) 『閣議決定』, 内閣官房. https://www8.cao.go.jp/cstp/tougosenryaku/togo2023_honbun.pdf
- Melissa M Chen, Lauren Parks Golding, Gregory N Nicola (2021) “Who Will Pay for AI?,” *Radiol Artif Intell*, Vol.3, No.3, e210030.

Mandeep R Mehra, Sapan S Desai, Frank Ruschitzka, Amit N Patel (2020) "Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis," *The*

Lancet Journal. doi.org/10.1016/S0140-6736 (20) 31180-6.

(たくら・ともゆき)

The Future of Health Care Supported by Big Data: Health Value of Adherence Coexisting with Artificial Intelligence

TAKURA Tomoyuki*

Abstract

Artificial intelligence (AI), such as machine learning, is a tool that has excellent synergy with big data in the medical and long-term care fields and is expected to be developed in the healthcare field, where digital transformation (Dx) and other technologies are being introduced. Considering the reforms to be implemented in the scheduling of physicians work in the future and the socioeconomic trends surrounding the medical care field, it is essential to improve the efficiency of medical care service provision. Therefore, there is a need to increase productivity, expand supply, and maintain and improve quality through the sharing and distribution of information (big data). In this context, we anticipate some promising signs in the field of healthcare, such as the development of AI-based diagnostic imaging software that is now being evaluated by public insurance and the development of a long-term clinical and economic forecasting model based on adherence to big data and AI. Several issues must be addressed to promote the use of these data sciences in healthcare. When we look at real-world data and deep learning, which are increasingly used in academic research and insurance systems, we can imagine a certain level of social risk centered on the issue of information quality assurance. One possible countermeasure would be the development of quality assessment guidance (e.g., standardization) for medical big data. Because the development of this field is expected to create significant socioeconomic value, it is desirable to promote the understanding (acceptance) of society as a whole, develop the necessary environment, and appropriately nurture the field.

Keywords : Database research, Machine learning, Data science, Adherence, Medical value, Predictive modeling, Health services, Diagnostic assistance, Quality assurance

* Department of Health Care Services Management, Nihon University School of Medicine/ Department of Healthcare Economics and Health Policy, Graduate School of Medicine, The University of Tokyo