

#### 資料4 「医療における機械学習 倫理的課題への取組み」 (PlosOne 誌・2018年)

最近イギリスで実施された調査によれば、成人の63%が医療の改善のために個人データが使われることに不快感があるとしており、また普段の診療業務を行っている医師や看護師に人工知能(AI)が取って代わることを好ましくないと考えていると報告された( )。

さらにドイツで実施された研究では、将来医師になる医学生の圧倒的多数(83%)が、AIが医療を向上させる見込みがあることを受け入れながらも、例えば画像検査などで最終診断を下すことにおいてはより懐疑的であることが分かった(56%が不同意)( )。AIに関する今後の展望についての質問では、米国の医療機関の意思決定者はAIが医療を向上させると確信しているが、しかしその約半数はAIが致命的な過誤をもたらす、適切に機能せず、現在高まっているような期待には応えられないと考えている( )。

このような調査データは、医療においてAIを取り巻いている倫理や規制に関する課題、特にプライバシー、データの公正性、説明責任、透明性、そして法的責任などの問題をよく捉えている。これらへの対処がうまくいけば、医療における機械学習(MLM、以下「医療機械学習」)の未来が発展し、医療に対しプラスの効果を与えるであろう。医療機械学習の倫理や規制の面での懸念は、それらの出現の仕方と場所によって3つの大きなカテゴリーに大別される。すなわち、医療機械学習で必要となる情報源、開発で用いられる方法、臨床診療への医療機械学習の配置である(仮説としてのケーススタディは図1を参照)。

#### 医療機械学習におけるデータの調達ではデータ保護とプライバシーに関する要件を遵守する必要がある

医療機械学習のアルゴリズムではプライバシー保護の対象となるデータが用いられ、開発者はデータ処理の各段階で倫理と規制に関する制約に細心の注意を払うことが求められる。データの出所および利用や再利用に関する同意は特に重要であり( )、量が多く幅広い種類のデータを要する医療機械学習では特にそれが言える。そのような異種のデータは使用条件が異なる、および/または異なる法的保護による制約を受ける可能性が非常に高い。

その顕著な例として新たに成立した欧州一般データ保護規則(GDPR)が挙げられるが、この規則はデータの使用に際して明確なインフォームドコンセントの要件を設定

しており、対象となる人々のデータの処理において尊重されなければならないいくつかの権利を保証している( )。さらにこの法は、データ処理が行われる場所にかかわらず、欧州連合域内の市民から得られたデータに適用される。アルゴリズムの開発に用いるデータは必要な利用許可を得なければならないが、所与の目的のためにどのデータの利用が許可されるのかを特定するのは容易な作業ではない。さらにデータの特定は、データの種類や法的管轄区域、利用目的、監督モデルといったものにも依存する。

医療データの倫理や規制に関する懸念についてはこれまで長らく議論されてきた経緯があり、GDPRなどの規制は、この新しいデジタルな課題に対処するための法的な動機付けとして象徴的である。しかしながら、これまで法的管轄区域ごとに異なる結論に至っており、また今後も異なる結論に至ることが予期される。例えば、GDPRと米国の「医療保険の相互運用性と説明責任に関する法律」(HIPAA)を比較すると、後者は患者のカルテから得られた医療データに重点を置く一方、生命保険会社や血糖値をモニタリングするスマホのアプリによって生成されたデータなど、対象とする団体や仕事上の関係者の外部で生成された医療データ等の他の医療データソースは対象としていない( )。

### 医療機械学習の開発では公正性に力を注ぐ必要がある

コンピューターサイエンスの格言に「ゴミを入れればゴミが出てくる」という格言がある。これは医療機械学習に特に当てはまる。というのは、医療機械学習モデルを学習させ、その有効性を確認する際に用いられるデータセットは、予測アルゴリズムが倫理的に使用されるようにするために重要な役割を担うものだからである。原データを不完全にしか代表しない学習用データセットでは、医療機械学習を学習させるアルゴリズムにバイアスが持ち込まれる可能性がある。

「バイアス」は問題をはらむ用語であるが、これには医療データによくみられる少なくとも2つの原型がある。1つ目は、データソースそれ自体が所与の集団内の実際の疫学を反映していない場合であり、例えば、アフリカ系アメリカ人を統合失調症であるとする凝り固まった過剰診断によってバイアスがかった集団データなどがある( )。2つ目は、所与の対象者の集団に十分な人数が含まれていないデータセットによりアルゴリズムに学習させる場合であり、例えば、ほとんどが高齢の白人男性から得たデータを用いて学習させたアルゴリズムなどである。このようなアルゴリズムでは、例えば若年の黒人女性に関する予測は不完全なものになるであろう。このような特徴を持ったデータセットで学習させたアルゴリズムが医療で採用されてしまえば、健康格差が悪化する可能性がある( )。

こうした落とし穴を避けるためには、科学分野の学会や規制当局がバイアスのかかった学習データセットの下流における影響を認識し最少化するための業界標準を開発する必要があり、一方で、治験審査委員会、倫理審査委員会、健康技術評価機関などの団体がそのような基準の遵守を確認する必要がある。

アルゴリズムにバイアスがかかるリスクを低下させる一つの格好の機会として、米国食品医薬品局 (FDA) がデジタルヘルス・イノベーション行動計画の一環として、品質も含めた5つの優れた基準に基づいて開発中の医療ソフトウェアを評価する事前許可制のパイロットプログラムを開始した例が挙げられる( )。バイアスが存在する学習セットによる悪影響を避けるために、FDAの品質基準、およびこれに似た他の規制基準を拡大し、学習用データセットにおけるバイアスのリスクを対象にすることもできるだろう。このような必要性に対して、いくつかの団体で取り組みが行われている。例えば、米国医師会 (AMA) が最近行ったAI (この場合は「拡張知能」) に関する政策提言では、データのバイアスを特定し最少化することを奨励している( )。さらにこれとは異なる方法は、米国での精密な医学コホート研究「All of Us」でみられるように、学習や有効性の検証に用いることが可能な原データをより良く代表するデータセットに資金提供することである( )。

規制によってアルゴリズムをバイアスから保護することに加え、規制当局が当面の問題に対する指針などの形式で最新の状況を反映した承認を行うことが、医療機械学習をベースにしたアルゴリズムの安全性と有効性の維持には必要である。このようなアルゴリズムは、原則的には使用を重ねることで継続的に改善させることが可能であり、リアルタイムに「更新」されるものだが、そのような更新を個々に治験で検査したり、医療を管轄する規制当局の典型的なタイムテーブルに沿って評価したりすることは恐らく不可能である。規制当局は市販後モニタリングのための効果的なメカニズムなど標準的な手続きを開発する必要があり、それによってソフトウェアの開発者は医療機械学習ソフトウェアの変遷を、透明性をもって記録することができる。

### **医療機械学習の導入では透明性の基準が満たされる必要がある**

医療機械学習は倫理的および法的に最も困難な問題や、現在の規制のあり方にも最も大きな課題を恐らく投げかけているが、その医療機械学習は、内部の理論が開発者にも隠されたままの説明不可能ないわゆるブラックボックス・アルゴリズムによって象徴されている。このような透明性の無さは、医療機械学習がベースとなる評価の機械的な解釈を不可能にしており、また信頼性も低下させている。さらに、医療についての基本的で意味のある詳細情報を患者に公開することは医療倫理における基本原則であるが、そのような情報公開を行う場合には、医師は自ら用いている機器の少なくとも内部の基

本的な働きを理解することが求められる。したがって医療機械学習が倫理的であるためには、開発者はエンドユーザーである医師に対し、医療機械学習をベースとした決定の背後にある一般的論理を伝える必要がある。また、前向き研究や無作為化試験などにおいては医療機械学習の臨床評価の正当性を証明するために、ある程度の説明ができることが求められる。医学的決定が完全に自動化された場合では、決定の手順に伴うリスクのレベルにより、患者のケアの管理に用いられている医療機械学習をベースとした技術の存在についての情報を患者に説明するのかどうか、またどのような方法で説明するのかが決まるだろう。医療機械学習という技術の使用を患者に伝えることで患者がより信頼し受容する可能性があるが、このことは既に述べた調査結果が示唆しており、引き続き関心が持たれることである。

診断や治療介入が医療機械学習に基づいて行われるようになるにつれ、患者の健康に関する決定プロセスにおける患者の匿名性や、意思決定が共有される可能性が損なわれる場合が出てくる。こうしたことが起こると思われる場合とは、例えば自動化された意思決定ツールへの依存により医療提供者と患者の間の意味のある対話が減る、または支払者が医療機械学習によって推奨されたことを払戻しの前提条件のように考え、自分の意思に反する医療機械学習の推奨内容には治療費の支払いを拒むなどの場合である。患者と医療専門家やケアの提供者との間の個人的な接触や意思の疎通の重要性を考えれば、医療提供者をサポートして力づけるような提供システムに医療機械学習を組み込むことが極めて重要である。第三者による監査も必要と証明される可能性があり、そのように倫理面で統合することが認定の前提条件となる可能性がある。

スマートフォンのアプリやウェアラブルデバイスなど、AIによって可能になる消費者向けの情報処理システムに特に注意を払う必要がある。機械学習の急速な進歩により健康関連の消費者向け製品が急増する可能性があり、それらの製品のための品質基準や認定システムが開発される必要がある。

最後に、医療機械学習の使用に関連した有害事象に対する責任の割当てやその根拠を明らかにする必要があるであろう。さまざまな種類の医療機械学習によって病院や医師、開発者がどの程度責任を負うことになるのかを徹底的に予想することが可能にならない限り、幅広い採用を実現するのは困難であろう。

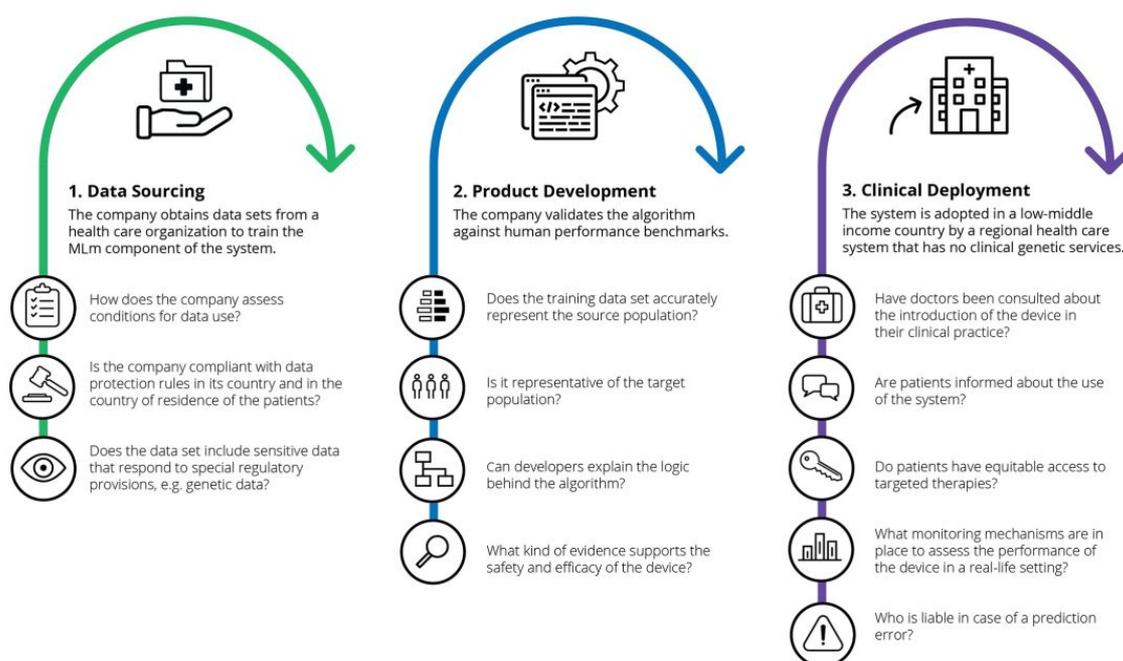
## 結論

医療機械学習を臨床で使用することにより、既存の医療提供の様式が変わる可能性がある。医療機械学習は医療専門家が臨床の環境で用いるようになり、モノのインターネット（IoT）を通じて高性能デバイスに組み込まれ、そして患者自身が臨床現場の外

で慢性疾患の自己管理のために用いるようになるだろう。医療機械学習への投資が急激に伸びていることは、研究が加速しており、より多くの製品が近い将来の市場参入を狙っていることを示している。医療機械学習が患者からの信頼と医療提供者による採用に値するためには、データ保護の要件に十分に合致し、バイアスによる影響が最少化され、規制が有効であり、透明性を達成している必要がある。このような倫理や規制に関する問題にできるだけ早く対処することは、医療機械学習のさらなる進展を妨げる不要なリスクや落とし穴を回避するのに不可欠である。

## 図 1. 機械学習に基づく機器を開発する医療ソフトウェア開発企業のイメージ

□ この機器は、癌患者の病理組織を示したスライドの完全に自動化した分析を行い、それらの画像のみに基づいて腫瘍の遺伝子変異を予測する。このようにして推測された遺伝子情報は予後の評価や候補となる標的療法を特定するために用いられる。しかし、アルゴリズムが画像のどのような特徴を変異遺伝子と関連付けたのか、またはそのような関連付けを生物学的にどのように解釈するのか、ユーザーが知ることはない。機器の販売提案は次のようなものである。機器は診断プロセスの早期に、価値ある遺伝情報を推論でき、またこの機器は、遺伝子検査を実施することができない状況下で、病理学者がクラウドベースのプラットフォーム上で共有する画像を分析することができる。



### 1. データの調達：企業が医療機関からデータセットを入手し、システム内の医療機械学習を構成する要素に学習させる。

- 企業はデータの使用条件をどのように評価するのか？
- 企業がその属する国や患者が在住する国のデータ保護規則を遵守しているか？

- そのデータセットは遺伝子データなど、特殊な規制条項に対応するデリケートなデータを含んでいないか？
2. 製品開発：人間作業の指標と比較しつつアルゴリズムの妥当性を実証。
- 学習用のデータセットは元の集団を正確に代表しているか？
  - そのデータセットは対象集団を代表しているか？
  - 開発者はアルゴリズムの背後にある論理を説明できるか？
  - 機器の安全性と有効性を裏付けるどのような種類のエビデンスがあるのか？
3. 臨床現場への配置：低 中所得国において臨床現場で遺伝子サービスを持たない地域医療体制において採用
- 医師は自らの臨床診療に機器を導入することについて相談を受けたことがあるか？
  - 患者はこのシステムの使用について知らされているか？
  - 患者には標的療法を利用する公平な権利があるのか？
  - 実際の生活環境での機器の動作を評価するためにどのようなモニタリング機構が整っているのか？
  - 予測に誤差があった場合は誰が責任を負うのか？

(仮訳：井上悠輔)

注記：文中の「 」は、該当箇所に参考文献があったことを示す。参考文献の詳細な文献情報は以下では省略した。

著者：Vayena E, Blasimme A, Cohen IG.

原題：Machine learning in medicine: Addressing ethical challenges.

出典：PLoS Med. 2018 Nov 6;15(11):e1002689.