



ビッグデータ・AIによる医学・医療の第3次革命と未来の医学

田中 博 Hiroshi TANAKA

医学・医療の分野も、今や「ビッグデータ医療時代」に突入し、今後数十年続く「第3次革命」(第1次革命: 抗生物質による細菌性感染症の激減, 第2次革命: 分子生命科学による分子医学) を起こしつつある。従来の医学は「仮説駆動型」医科学で、得られた知識は穴だらけで全体性が欠如している。未来の医学研究は「データ駆動型」医科学で、全体被覆性のある網羅的データから人工知能の「教師なし学習」によって知識を自動発見する。生命情報学はそれがもたらす分子生命科学に基づいた「深い知識」によって、次世代ビッグデータ・AI 医学・医療を支える。

医科学における大変革

まず、医学のビッグデータ革命が起こした、医学分野における研究方法論における変革、すなわち「データ駆動型 (data-driven)」医学の方法的インパクトについて論じよう。

これまでの医学では、医学研究者の関心や社会の要請の下に、一定の医学の対象領域が選ばれ、これまでの知識体系に基づいて解明すべき不明の機序に対して仮説が構築され、それを実験を通して検証し、その結果として医学の知識体系が新しく順次築かれてきた。このような研究方式は、一般に「仮説駆動型 (hypothesis-driven)」医科学と呼ばれる。

これに対して、ビッグデータ医学時代では、まずは生命の網羅的データが収集され、そのデータから機械学習や人工知能の帰納的方法を活用して、新しい知識を構成する。すなわち、「データ駆動型」医科学である。

この相違は、通常考えられている以上に大きい。「仮説駆動型」医科学では、ヒトが仮説を立てるために、関心の高い分野、社会的要請のある分野は、非常に深く究明されたが、隣り合った、研究関心も少なく社会

的要請もない分野は、全く探求されずに放置されていることも多い。医学の体系は様々な深さの知識の集まり、つまり「穴だらけ」の孤発的知識の集合である(図1参照)。これに対して「データ駆動型」究明は、網羅的データの収集から知識を生成するため、知識の深度は一樣で、人間が立てた分野の境界に煩わされることはない。

ただ、データに基づいてだけ知識を生成するといっても、どんなデータでもよいのではない。そのデータの収集性の範囲において、一樣な深さを持ち「全体被覆性」を持つ網羅的データでなければならない。生命科学においては、この全体被覆性を持つ「網羅的データ」は、ヒトゲノム解読計画以降に現れた。全遺伝子の発現状況を観測できるDNAチップや、次世代シーケンサによるゲノム配列や様々なオミックス情報の計測などは、一樣な深さと全体被覆性を持つ「網羅的生命データ」であり、これらデータの「ゲノム・ワイド性」が知識生成の1つの成立条件となる。

さらにもう1つの成立条件は、これらビッグデータから知識を生成するときに、これまでの仮説に惑わされずに知識を生成する方法である。生命ビッグデータの規模の増大速度が加速している現状に適した方法として登場したのが、機械学習・人工知能である。特にニューラルネットワークの一種であるディープラーニングは、ヒトが与える知識なしに、いわゆる「教師なし学習」によってデータの持つ構成的特徴量をみずから自動的に抽出する。このような「教師なし学習」の発見は、「データ駆動型」医科学の成立条件を満たすもう1つの柱である。

たなか・ひろし
東京大学医学部衛生学 客員研究員、東京医科歯科大学 名誉教授

【経歴】1981年東京大学医学系大学院修了、医学博士。83年同大学工学系研究科より工学博士を受ける。82年同大学医学部講師、90年米国マサチューセッツ工科大学 (MIT) 客員研究員、91年東京医科歯科大学教授。2007年同大学大学院生命情報科学教育部部長、大学評議員。15年同大学名誉教授、東北大学東北メガバンク機構長特別補佐、23年東京大学客員研究員。



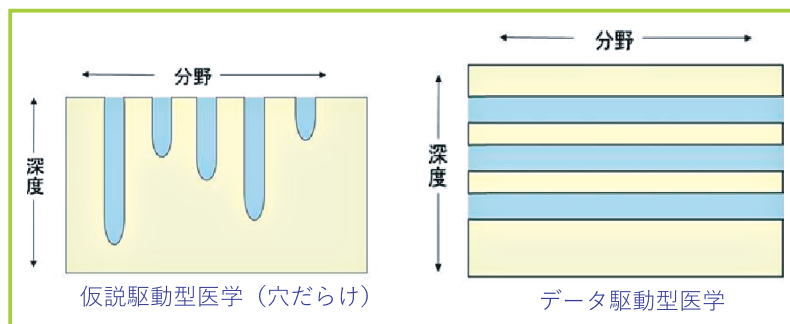


図1 仮説駆動型医科学とデータ駆動型医科学

それでは、ビッグデータ・AIは、研究面だけでなく、医療の実践においてはどのような変革をもたらすのであろうか。筆者は、最近刊行した成書「医療ビッグデータ」¹⁾、「次世代生命情報医学」²⁾でこの点を詳しく論じたが、ここでは、その要点を述べよう。

医療ビッグデータの現在における主なものは以下の4種類である。

- 1) 次世代シーケンス革命によって急速に発展したゲノム・オミックス分子配列の生命情報
- 2) 欧州を中心に急速に普及したゲノム・コホート型のバイオバンク情報
- 3) スマートメディアとウェアラブルセンサーによるモバイル・ヘルス情報
- 4) 電子カルテの普及によるリアルワールドデータである。そのそれぞれがどのような未来の医療をもたらすか論じてみよう。

(1) ゲノム医療による個別化医療の発展

2007年のシーケンス革命の3年後の2010年にウィスコンシン医科大学での5歳のニコラス君の原因不明の腸炎症疾患に対して、全エキソーム・シーケンスでその原因遺伝子を同定し、臍帯血骨髄移植でその一命を救った。この出来事の報告は全米を震撼させ、ゲノム医療時代が到来したことを広く認識させた。米国ではこの成功以来、怒涛の勢いでゲノム医学の臨床実装が広がった。これらは、同じ先天的稀少疾患の原因遺伝子の「臨床の現場」での同定を目指すものである。

また、同じ2010年からバンダービルト大学医療センターで薬剤の代謝酵素の多型性の診断が行われ、薬理ゲノム学が臨床応用された。さらに2012年からはダナハーバーがんセンターをはじめとして、昨年我が国で

も導入された100余りのがん遺伝子のパネル診断が始まった。これは、患者のがんのドライバー変異遺伝子の同定とそれに対する分子標的薬の選択という新しいがん医療を普及させた。

これらのゲノム・オミックス情報の臨床応用は、従来の医学では、同じ名前で括られていた疾患が、相異なる成因よりなる様々な「内在的サブタイプ」よりなること、そのそれぞれが予後も治療方法も異なることが明らかとなり、「個別化医療(individualized medicine)」の概念が出現した。これまでの平均的病態像に基づく「ポピュレーション医療」すなわち、one size fits for allの医療はもはや成り立たないことが明らかとなった。

(2) 大規模ゲノム・コホートの欧州における普及

翻って欧州では、社会福祉国家の理念の下、社会全体のヘルスケアの向上を目指して、「集合的な遺伝情報」の収集を国家政策として推進する流れが進展している。これは大規模な住民準拠(population-based)な健康・疾患発症追跡コホートを、生活習慣・健康情報だけでなく住民のゲノム情報も収集して長期間実施するもので、ゲノム・バイオバンクとも呼ばれている。UKバイオバンクが「大規模前向き population-based ゲノム・コホート」という研究形式を確立して2007年に50万人の長期追跡コホートとして実行された。ゲノム・コホートは疾患の成立機序と発症リスク解析のために、不可欠な医療ビッグデータである。

これまでの医療は、病気になってからの対応主義的な医療というべきものであるが、ゲノム・コホートは、life-course-orientedな視座からの疾患把握で、疾患の発症をリスク予測し、疾患が本格的に発症する前にこれを治療する「予知制御医療」を可能にする枠組み

をもたらした。

(3) 生理的変量のIoT連続測定とモバイル・ヘルス

もう1つの医学・医療の新しいタイプの「ビッグデータ」は、近年の「モバイル・ヘルス (mHealth : mobile health)」が生み出す「医療ビッグデータ」である。

これは、①近年のスマートメディアの急速な発展による「デジタル通信技術の革新的な展開」と、②有機伝導繊維などによる非拘束型生体モニタリング技術の発達、さらには、③「患者中心主義」、「患者参加主義」運動の拡大——という合体した流れによって生み出されたものである。

2000年代の初め頃から、「Quantified Self運動」が米国西海岸から始まった。現在ではモバイル・ヘルス医療は、生理的状態の情報に基づいて、治療に関与する段階に達し、「デジタル治療法 (digital therapeutics : DTx)」と世界的に呼ばれている。例えば、Blue Star というシステムは、2型糖尿病の患者の血糖値・服薬状況・生活情報を基に、患者の治療を支援し、2017年米国のFDAに薬事承認された。

(4) 電子カルテによる「リアルワールドデータ」

臨床研究を「科学」にするために、確立された概念である「Evidence-based Medicine (EBM)」や「ランダム化比較試験 (RCT)」が、「医療ビッグデータ」の出現と「Real World Data」の急速な普及によって、その有

効性に対して疑義がもたれるようになった。特に実際のRCT臨床試験が医療実践の外側で行われていることに、多くの批判を集めるようになった。実際に米国の治験を見ると、被検者は白人の成人ばかりで、高齢者、妊婦、黒人青年は含まれていない (図2)。

今や、「ビッグデータ」時代であり、サンプル収集費用が低廉化し、大規模なデータを収集できる。サンプリング方法に過剰に注意を払って標本化する必要はなく、言わば、母集団に近い大規模な Real World Data を直接収集できる状況にある。

未来の医学：次世代生命情報医学

このように医学・医療ビッグデータと人工知能は、新しい第3次革命を医学・医療の分野にもたらしつつある。しかし現状のままでは、まだ真の第3次革命には不十分である。

まずは、モバイル・ヘルスで利用している生体情報は、脈拍や血糖値などの古典的生理変量である。最近では、liquid biopsy など血中に流れる浮遊DNAやがん浮遊DNAを測定する技術が発達してきている。モバイル・ヘルスがこれら「深い生命情報」を積極的に利用するPM (Precision Medicine) モバイルヘルスへと発展することによって「予知制御医学」へつながる。

さらには現在の real world data の利用は、古典的電子カルテの大規模集積でしかなく、近年発展の著しいゲノム電子カルテ、がん治療分子医学電子カルテの大規模な利用には至っていない。がん治療分子医学電子カルテの臨床実践を集積した精密医療DB (PM real world data base) は、今や臨床治験の対照群として薬効検定に用いられるまでに至っている。このようなゲノム電子カルテやがん分子治療電子カルテを収載したPM real world dataこそ、新しい次世代生命情報医学の知識基盤である。PMモバイルヘルスとPM real world dataによる基盤の情報支援の下に医学・医療の第3次革命は、次世代生命情報医学を展開するであろう。

- 1) 田中 博, 医療ビッグデータ, 羊土社, 2020.
- 2) 田中 博, 次世代生命情報医学, コロナ社, 2023.
- 3) Y. Asahina, *J. Nutr. Health Aging* 2014, 18, 520.

© 2024 The Chemical Society of Japan

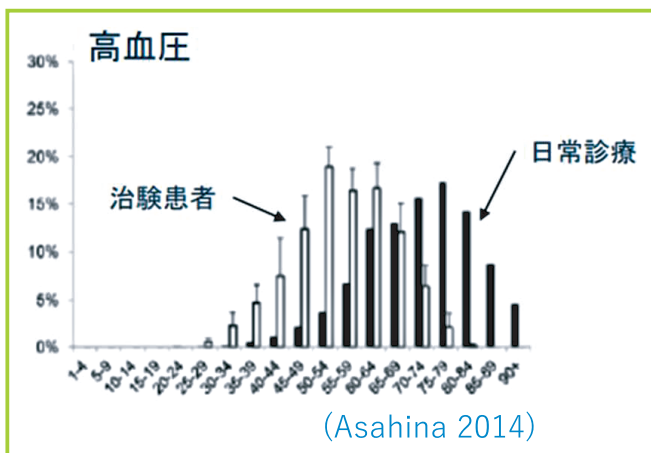


図2 治験集団と日常診療：高血圧患者年齢分布³⁾