

2. 医療 AI の研究開発を支える データ利活用環境の課題と展望

黒田 知宏 京都大学医学部附属病院医療情報企画部

医療データを取り巻く 社会の動き

人工知能 (AI) は、一般的に教師データと呼ばれる一定量のデータを用いた学習を経て、利用可能となる。医療 AI を開発するには、当然一定量以上の医療データが必要であり、これを収集する環境を求める声は、日増しに高くなっている。

一方、医療データは個々の患者の究極の個人情報の集合である。したがって、その取り扱いはいわゆる慎重に行われねばならない。近年、欧州を中心に、データの自己コントロール権であるプライバシー権を基本的人権とする考え方が広がってきており、欧州一般データ保護規則 (general data protection regulation: GDPR) を中心に、各国でプライバシー権を保障する個人情報保護法制度の整備が進められている。本邦の個人情報保護法においても、本人の同意なく個人情報の利用や第三者提供を行うことを固く禁じている。特に、医療データについては、2015年の改正で「要配慮個人情報」に指定され、慎重な取り扱いが求められている。

このような、医療データを巡る「提供者 (患者)」と「利用者 (研究者など)」の利害の衝突は、AI の登場以前から存在した。医療データは、AI が現れる以前から、医学研究や医療の発達に欠かせない人類の共有資産であって、情報技術の進展によって大規模集積と解析

が容易になったことで、「個」と「公共」の衝突が顕在化したにすぎない。世界医師会は、2016年に医療データベースとバイオバンクの倫理的取り扱いに関する台北宣言を発表したが、その検討が公式に始められたのは2002年のことである¹⁾。

本邦の医療ビッグデータ 基盤

わが国の医療制度は、国民皆保険制度とフリーアクセス制度からなっていることは広く知られている。どこに行っても (悉皆性) 同じ医療 (均てん性) を受けられることを保証するためには、個別の医療サービスに対する厳密なルールと詳細な値付けが必要になる。したがって、本邦の保険制度は、出来高支払制度 (fee for service) と呼ばれる、個別の医療行為の価格を定める仕組みを採っている。国民皆保険・出来高支払制度の下では、診療報酬の請求書 (レセプト) を収集すれば、国内で行われているすべての医療行為の記録を収集することができる。2015年度から診療報酬の電子請求が義務化されたことで、すべての医療記録を電子的に収集し、データベース化することが可能になった。2008年、高齢者の医療の確保に関する法律の下でレセプト情報・特定健診等情報データベース (以下、NDB) が構築され、閣議決定された「新たな情報通信技術戦略工程表」に基づいて、2013年には研究者向けの第三者提供が始まっている。

NDBは、きわめて詳細な「行為の記録」ではあるが、診療の結果である「アウトカム」は保存されていない。対象疾患によって求められるアウトカムは異なるため、さまざまな「疾患別レポジトリ」が学会などによって作られている。特に、がんについては、2016年1月1日より「がん登録等の推進に関する法律」の下で、国家事業としてデータ収集と分析が進められている。

NDBも疾患別レポジトリも、基本的には学術研究目的で提供されている。しかし、学術研究の結果として得られた医療 AI などを、臨床研究に適用する開発研究には適用できない。個人情報保護法は、憲法で保障された「学問の自由」を妨げないため、学術研究目的での個人情報の利用については、自主的統制を求めているもの、基本的には妨げていない。一方、開発研究目的の場合には、同意原則が適用されることになる。しかし、大量のデータを取り扱うことによって立つ医療 AI の場合には、そのデータ源となるすべての人に個別同意を取ることは不可能に近い。

そこで、個人情報保護法の特例措置として、政府機関による認定を受けた特定の事業者 (認定匿名加工医療情報作成事業者: 認定事業者) のみが、医療データを収集し、名寄せし、匿名加工した上で開発研究目的で提供できるようにする「医療分野の研究開発に資するための匿名加工医療情報に関する法律」、いわゆる「次世代医療基盤法」が制定された。図1に、次世代医療基盤法の全