

II AIと医療を考える—インタビュー

1. 情報科学の視点から考えるAIと医療の未来

膨大なゲノム情報を扱う
がん医療に不可欠なAI個別化医療を実現するためにも
ビッグデータの蓄積が必要

宮野 悟氏 東京大学医科学研究所ヒトゲノム解析センターセンター長



医療において最も人工知能 (AI) の活用が進んでいるのは、ゲノム変異を突き止めて最適な治療法を見つけなければならないがん医療の分野である。膨大なゲノム解析データを基に数千万に及ぶ文献情報・特許情報から治療法を提示するAIは、これまでのがん医療のあり方を大きく変えようとしている。IBM社のコグニティブ・コンピューティング「Watson」を用いた最先端のがん医療を支援する研究に取り組んでいる東京大学医科学研究所 (医科研) ヒトゲノム解析センターのセンター長である宮野 悟氏に、情報科学の視点から見たAIについてインタビューした。

■がん医療はビッグデータとの格闘

医療の中でも、がん医療では、“ビッグデータとの格闘”が行われています。がん医療では、非常に膨大なデータを扱うため、スーパーコンピュータやAIといった技術の活用が重要になっています。がんは、喫煙や放射線被ばく、ウイルス感染といった環境因子と加齢などによりゲノムが変異し、遺伝子欠損や挿入、塩基置換が起こることによって発生します。このがん発生の原因であるゲノムの変異を調べるために、次世代シーケンサーを用いてゲノムを解析し、さらにスーパーコンピュータを用いて変異を探索します。医科研では、ゲノム変異を検出するソフトウェア“Genomon”を開発し、スーパーコンピュータ「SHIROKANE」シリーズと組み合わせて、研究を進めています。

しかし、ゲノムの変異を発見できても、その解釈・翻訳には、人海戦術で膨大なデータベースや文献などを検索する必要があり、がん医療のボトルネックになっています。このボトルネックを解決する手段として、医科研ではIBM社のWatsonを導入することにしました。

■2000万件以上の論文要旨と
1500万件以上の特許情報を学習したWatson

米国のメモリアル・スローン・ケタリングがんセンター (MSKCC) のプレジデント兼CEOのCraig B. Thompson氏は、「がんの理解は、もはや一人では不可能である」と述べています。これまでの基礎研究からも、がんの理解は人知を超

えていて、想定外に複雑だということが明らかになっています。

がんに関するこれまでの研究成果は、膨大な量に及んでいます。医学・生物系の論文要旨のデータベースであるPubMedには、2016年までに2600万件の論文が登録されており、これを印刷して積み重ねると富士山よりも高い4000mに及びます。さらに、2016年だけでもがんに関する論文は20万報あり、このままでいけば、2050年には大気圏外である高度100kmに及ぶと予測されます。また、特許件数も膨大な数に上っており、2015年には1500万件以上となっています。これらを一人の人間がすべて読むのは不可能です。

一方で、ゲノムの変異も非常に複雑であり、これを理解するのは困難です。例えば、BRCA1遺伝子のエクソンは365個の変異が報告されています。しかし、BRCA1遺伝子に変異を持つ女性1万9591人のうち、乳がん発症が9052人(46%)、卵巣がん発症が2317人(12%)、両方発症が1041人(5%)となっていますが、37%に当たる7171人は70歳までがんを発症していません。つまり、ゲノムの変異は多様であり、がんの発症との関係を簡単に解明することはできません。現状では、がんの発症メカニズムのすべてを理解できておらず、一部のゲノム変異の情報から、治療を行っているにすぎません。

このような現状の中でWatsonの存在を知り、2015年7月に“Watson Genomics Analytics (WGA)”を導入し、IBM社との共同研究を開始しました。WGAの導入は、MSKCCやメイヨークリニックなど北米の医療機関・研究機関以外では初めてです。まずNew York Genome Centerにおいて、MEDLINEにある2000万件以上の医学・生物学論文要旨と1500万件以上の特許情報、COSMIC、米国国立衛生研究所のClinVar、米国国立がん研究所のパスウェイデータベースなどの情報を学習しています。

ところで、医科研では当初、WGAを2014年12月に導入する予定でした。それが翌年7月にずれ込んだのは、プログラムとデータの組み合わせであるAIが、薬機法における医療機器プログラムの対象となるかが明確に規定されていなかったためです。私たちの経験では、Watsonは運用しているうちにデータ量が増え、さらに学習しているため、1か月使用した