

[教育講演4]

検査ビッグデータの利活用

浅野 薫

シスメックス株式会社

1. はじめに

ビッグデータというキーワードを目にする機会が増えてきている。ビッグデータに明確な定義はないが、単にデータの量が多いというだけではなく、さまざまな種類や形式が含まれる非構造(非定型)データの集合体であり、さらに、日々膨大に生成される時系列性・リアルタイム性があるものを指すことが多い。

そういう意味で、日々の診療・検査データが蓄積される医療・ヘルスケアデータはまさにビッグデータであり、それを解析することによって、医療・ヘルスケアプロセスの効率化や、診断・治療支援など、さまざまな新たな価値を創造することが期待される。

ただし、医療データは厳重に守られるべき個人

情報であり、取り扱いに細心の注意を要することは言うまでもない。

2. 米国での検査データ活用動向

医療データ、とくに検査ビッグデータの活用という点では、米国の臨床検査機関の動きが参考になる。Clinical Lab 2.0というコンセプトでは、臨床検査機関は、従来のような単に臨床サイドからの要求に対して正確な検査結果を返す(ClinicalLab 1.0)だけではなく、彼らが保有する検査ビッグデータを使って新たな価値提供を行うという取り組みが提唱されている(図1)。

具体的には、検査データ使って、
✓ポピュレーションヘルスへの活用

(ポピュレーションヘルスとは、予防から予

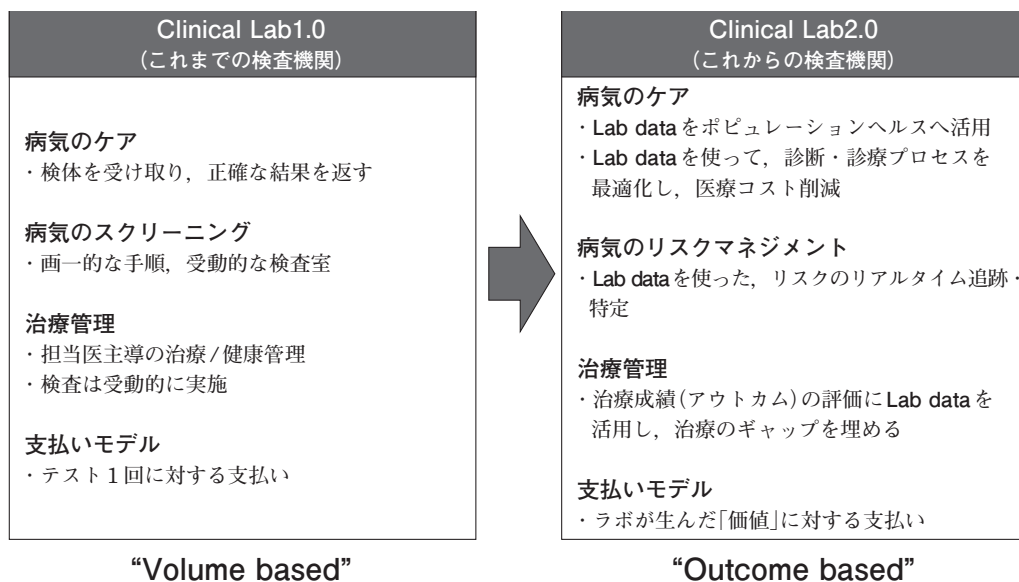


図1 Clinical Lab 1.0から2.0へ

後まで長期スパンで慢性疾患のリスクを低減する健康管理の仕組みのこと)

- ✓検査や診療プロセスの最適化による患者アウトカムの向上と医療費削減
- ✓医療プロセスにおけるリスクマネジメント
- ✓治療成績(アウトカム)の評価

などに活用し、検査の量をベースとしたビジネスモデルから価値提供をベースとしたビジネスモデルへの転換をはかろうとしている。

3. 検査ビッグデータの活用

それでは検査ビッグデータをどのように活用していくかについて述べたいと思う。

検査ビッグデータは、大きく分けて二種類ある(図2)。

一つは、不特定多数のデータの集合によるビッグデータで、この場合、平均値や標準偏差(SD)といった統計値が意味を持つ。たとえば、検査の基準値は、不特定多数の健常者の検査値分布の $\pm 2SD$ 範囲(95%の方が含まれる範囲)を言い、これを使って診断やスクリーニングが行われている。

もう一つは、個人ごとの時系列的データの集合で、個人ごとの変化の様子に意味がある。

前者の統計値は、予防のアウトカム評価に利用できる。たとえば、予防として、“ある施策”により生活習慣を変えることによって、特定の疾患に対する発症リスクを下げる取り組みがなされたとする。この“ある施策”の効果は、対象集団の検査値の平均値の変化により、定量化できる。

また、新型コロナ肺炎の例では、ウイルスに感染するとリンパ球が35%減少することが報告されているが、ビッグデータ解析により、このような知見が容易に得られるようになり、診断支援に役立つ。

個人ごとの時系列的データについては、治療(薬)の効果判定に使えることは言うまでもないが、さらに個人の健康管理や予防・先制医療につなげる可能性がある。

たとえば、個人の検査値のばらつきは、先に述べた検査の基準値範囲よりも狭い範囲に分布する(個人の基準値)と考えられる。検査値Xが、たとえ検査の基準値範囲に入っていたとしても、個人の基準値から逸脱したとすれば、従来よりも早期に異常をとらえられるかもしれない(図3)。さらに、時系列変化の動きから、AI手法を使って予兆をとらえ、“悪化(発症)”の方向へ向かっているのか?、“状態が維持されるのか?”それとも

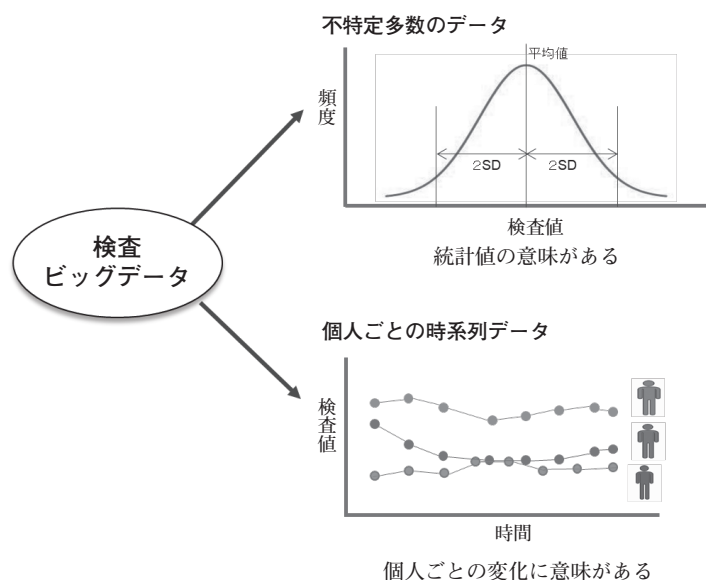


図2 検査ビッグデータの2つの種類

“改善の方向に向かうのか”を予測することが可能ではないかと考えている（図4）。

これは、従来の“リアクティブ型”医療から、“プロアクティブ型”医療への転換を意味する。

4. 最後に

献血者の血液検査データは、大規模かつ健常者

中心で、貴重なビッグデータである。

このビッグデータには、3で述べたような方法を使って、献血者個人の健康管理・予防・先制医療につなぐ解析が可能であると思われる。この解析結果を献血者へ適切に伝えることができれば、献血者にとって大きなインセンティブになると考えられる。

検査の基準値

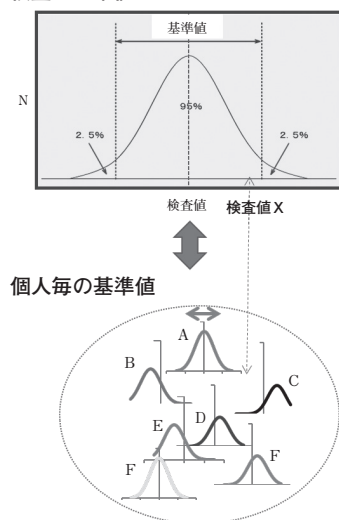


図3 検査の基準値と個人ごとの基準値

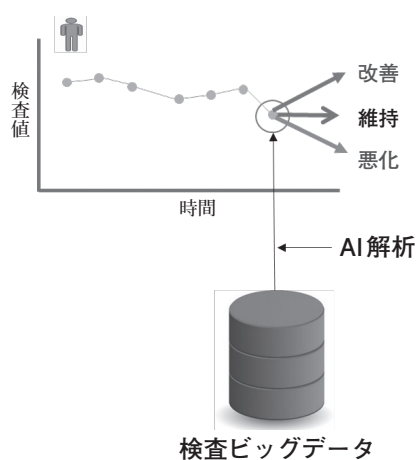


図4 “プロアクティブ型”医療の実現