

巻頭言

医学(医科学と医療)とビッグデータ

佐久間 貞行

最近よく目・耳にする言葉にビッグデータ(**big data**)がある。ビッグデータとは、通常、一般的に使用されているデータベース管理ツールやデータ処理アプリケーションで処理することが困難なほど巨大で複雑なデータ集合の集積物を表す用語とされている。

ビッグデータを冠した用語はビッグデータビジネス、ビッグデータ戦略、ビッグデータ革命などと称して、政経ビジネスにつなげるものが多い。

今ビッグデータ戦略が一番求められているのが医学の世界ではなかろうか。ビッグデータの医学応用に関しては、現在では医科学的データベースと医療データベースと入り口は二分化されているが、将来は本来の姿と思われる **big data** と機能的な **networking** による **reference man(virtual individual)** の構築とその医療応用にあると考える。

しかし進行しない一番の要因は、医科学的データベースでは特許問題であり、医療データベースではプライバシーの問題のように思われる。

この原稿を書いているときにノーベル賞の生理学・医学賞と化学賞の決定のニュースが入った。生理学・医学賞は細胞内輸送の問題である。受賞した内容より現在はもう少し研究が進んでいるようであるが、生体内ではこのような現象が無数に蠢いて生態を維持、構成しているわけである。ホメオスターシスもこの一環と考えても良いであろう。化学賞は生体内の化学反応のより正確なシミュレーション手法を創成したもので、京のようなスーパーコンピュータが出番である。こう並べて書くとビッグデータでは不足で人間に関するあらゆるデータの総てによる **Reference Individual** の創造の可能性がみえてくる。

医療記録のデジタル化が始まった当初から、個人記録のデータ保管についてはICカードの利用などが喧伝されたが、未だにごく一部の実行にとどまっている。東北大震災時には被災病院の診療記録の喪失で診療カードのIC化の必要性が再び論議されたり、震災用に用意されていた外科的医療材料よりも、日常診療で用いられる医薬品の不足が深刻だったり、データの集約と分析の不足が指摘されたりした。しかしこのような医療行政、医療経済の分析のためだけで

無く、日常診療に役立つビッグデータのフィードバックが望ましい。

医学的ビッグデータのうち医科学的データベースについては、ゲノムや、多くの一塩基多型の発見など、可成り整備され応用されやすくなりつつあるように思われる。しかし生化学的データベース、生理学的データベースなど分野毎の整備は認められるが、人体構成の微細構造、微細動態など全体像 **virtual individual** 構成の立場から見るとまだ十分とは言い難い。

放射線防護の基本量である実効線量の評価には、標準人(**reference man**, **ICRP Publ.23**)ボクセルファントム(**ICRP 国際放射線防護委員会**)を用いる。これは放射線量の測定を目的としたデータで構成された仮想人体(**virtual individual**)であるが、既知のあらゆる医学的データで構成される仮想人体があれば、照射や薬効の新しい解釈や発見もあり、臨床現場で大いに役立つであろう。

医学的ビッグデータでは従来のリレーショナルデータベース管理システム(**RDBMS**)だけでなく、新しいデータベース管理システムとより大量のデータ処理能力が要求される。新しいツールプラットフォームが、異種の大量のデータの様々な処理をするために開発される必要がある。ビッグデータは、大部分のリレーショナルデータベース管理システム、デスクトップ統計可視化パッケージのみでは処理が困難であり、その代わりに、数十台、数百台、数千台ものサーバ上で動く大規模並列化ソフトウェアが真の意味で実用化される事を期待する。2012年、ガートナーは、ビッグデータの定義を次のように変えたということである。「ビッグデータは、高ボリューム、高速度、高バラエティの情報資産のいずれか(あるいは全て)であり、新しい形の処理を必要とし、意思決定の高度化、見識の発見、プロセスの最適化に寄与する」。

大量にどんどん産出される「医科学データ」、「医療データ」のビッグデータ処理と人体を構成する全ての要素の代謝機能を表現する **reference man** 作り、その医学的応用を期待する。

(名古屋大学名誉教授、健康文化振興財団理事)