

口腔保健医療におけるビッグデータの利用価値

座長 宮崎秀夫 新潟大学大学院医歯学総合研究科予防歯科学分野

講師 Helen Whelton, Martin P. Ramsdale, Evangelos Zormpas リーズ大学歯学部

1. ビッグデータの定義

医療におけるビッグデータという用語は、デジタルの医療データセットとの関連性が高い。これらのデータセットは大きく複雑ではあるが、一般的なソフトウェアとハードウェアおよび通常のデータ管理ツールにより管理できるものである¹⁾。問題となるのはデータの量ではなく、多岐にわたるデータの種類であり、ビッグデータを操作するときに必要なとされるスピードである¹⁾。Boydら²⁾が示しているように、ビッグデータで重要なのは、収載できるデータの量ではなく、巨大なデータを検索、集約、相互参照・比較・分析する能力である。ビッグデータを異なる2つの観点から定義したBoydら²⁾によると、ビッグデータは技術的にはコンピュータ処理のスピードとアルゴリズムの正確さに関する概念であり、分析的には経済的、技術的、社会的、法的な還元を得るために、巨大なデータからいくつかのパターンを見つけ出すことである。医療におけるデータソースとしては、コンピュータによるオーダーエントリー(CPOE)システム、臨床的判断支援システム(医師による診療録や処方箋、医用画像、調剤記録、検査記録、保険および他の管理的データ)、電子カルテ、機器生成データ(バイタルサインのモニタリングなど)、ウェアラブル端末やスマホアプリからのデータ、ブログ³⁾、ソーシャルメディアへの投稿(ツイッターやフェイスブック、ウェブページおよび他のソーシャルメディアのプラットフォーム)⁴⁾、さらに患者の関与の低い救急医療のデータ、ニュースの記事や医学誌論文⁵⁾がある。McKinseyの報告によると、ビッグデータが医療の効率と質を改善するためにすべて有効活用されたとしたら、米国の医療業界に3,000億ドル以上の利益をもたらし、医療関連の支出は8%以上縮小されるとしている⁶⁾。

2. ビッグデータのソース

Miller⁷⁾は医療のビッグデータを2つの主要なソースによって分類している。1つはゲノミクス主導型のビッグデータで、遺伝子型決定やジーンシーケンシング(遺伝子配列解析)、遺伝子発現解析などがあり、もう1つは保険者主導型のビッグデータで、一連の電子カルテや処方箋、患者のフィードバックおよびレスポンス、さらには保険記録などが含まれる。Chenら⁸⁾は、長期間にわたって蓄積された年齢、性、教育歴など社会属性データや治療予後といった診療録(カルテ記録)や検査値からなる数十万規模の患者データとして保険者-医療機関のカテゴリーを記述している。Gelfand⁹⁾によると、医療関連のビッグデータか

らの情報を活用するにあたっては、健康関連の倫理的調査のためのプライバシー保護と信用保全の基盤を築くために、HIPAA(医療保険の相互運用性と説明責任に関する法律)およびIRB(治験審査委員会)が求める個々の要件を満たすことが前提条件である。プライバシーとデータ保護の課題に対して対応がなされたら、電子カルテなどの医療関連のビッグデータは、重要な臨床情報の収集や患者の疾患パターンの的確な把握に関して絶大な力を発揮すると思われる⁸⁾。たとえばLinら¹⁰⁾はある大規模病院からICD-09のコードを使って選択された約210万件の健康医療電子記録を使用した。その結果、臨床的に関連があり有効なSDT(症状-疾患-治療)において、7つの個別の健康状態(癌や感染症、慢性疾患など)に関連性があることが明らかになった¹⁰⁾。

3. 医療におけるビッグデータ解析の4つのV

1)量(Volume)

ビッグデータの解析には、3つの基本的な特徴が存在する。すなわち、量(volume)、速さ(velocity)、多様性(variety)である⁵⁾。膨大な量の医療データに含まれているものとしては、各患者のカルテ、エックス線画像、臨床試験データ、ヒトゲノム解析や集団ゲノム解析があげられ、最近では3D画像やゲノミクス、生体センサー測定値といった重いデータも含まれる⁵⁾。

2)速さ(Velocity)

ビッグデータ解析の速さとは、解析結果に基づいて意思決定を行う際に、データの獲得、保存、操作、分析、比較がリアルタイムにできることをいう⁵⁾。

3)多様性(Variety)

集められたデータは構造化、非構造化、半構造化されたものがあり、その形式は多岐におよぶのがデータの多様性の特質である。構造化および半構造化データとしては、医療機器の表示記録や、紙媒体の記録から電子媒体の健康医療記録へと段階的に変換されることで蓄積されたデータなどがある一方で、非構造化データには、院内保管のカルテ、医師や看護師による手書きの記録、紙媒体の処方箋やエックス線写真などが含まれる⁵⁾。健康医療電子記録(EHR)や電子カルテ(EMR)の構造化データでは、患者の氏名、生誕時の記録、住所、担当医の氏名、病院名、治療コードと償還コードの他に、自動化されたデータベースによって容易に管理できる関連情報の入力項目がある⁵⁾。医療におけるビッグデータの巨大な容量は、従来のデータ形式と新しいデータ形式を個人と集団の両方のレベルで結合すること

によって得られる。

4) 正確さ (Veracity)

臨床家や研究者によって導入されているビッグデータ解析の4つの特徴は正確さであり、データの確実性、信頼性、有効性が求められている⁵⁾。人が介入し治療方針を決定しなければならない分野では、正確で高品質な情報が必要なことは自明である。非構造化データセットでは低品質なデータがたびたび見受けられ、不完全な手書きの処方箋を不正確に“解釈”することでデータの内容が間違っていることがある⁵⁾。データの信頼性と品質を確保するために、生データは必ずしも公正、確実、正確ではないことを念頭に置き、データマイニングや統計処理、アルゴリズムおよび視覚化技術といったデータ解析モデルや解析技法により正確さを検討する必要がある⁵⁾。

ビッグデータの4つのVの他に視覚化 (Visualisation) と価値 (Value) が加えられ、理解を助けるためにビッグデータを視覚的に表示することの重要性が強調されている。現在入手できるビッグデータの量が重要な意味をもつのは、それが有用かつ信頼できるデータを提供できる場合に限られており、しかもそのデータは最終的には患者に価値を付加できるように分析・解釈されるものでなければならない。

口腔保健医療に関するデータ

口腔保健医療の分野でも、以下に示すようなさまざまなソースから大量の電子データが集積されている。

健康医療電子記録 (EHR)

2004年以降、米国における最近の法改正により、健康医療電子記録 (EHR) の採用とその有効利用が推し進められている¹¹⁾。その結果、米国では90%以上の病院と70%以上の地域医療サービスがEHRを活用し¹²⁾、米国の歯科関連のデータ蓄積量はこの期間で増大している¹³⁾。

歯科患者の電子記録は情報システムのデザインによって異なるが、一般的に従来の紙媒体で記録されていた氏名、住所、年齢、性別といった人口統計的情報と、医科的既往歴、現在および過去のデンタルチャート (残存歯、齲窩や修復の状況、要治療歯、歯肉の状態など)、さらにはエックス線写真、メモや覚書が含まれる。治療の詳細の他に、処方箋や歯科技工指示書も記録される。

有効かつ適切に構成された患者の電子記録は、医療の質と費用対効果を改善するのに有用な手段である¹¹⁾。

診療報酬請求書のデータ

診療報酬請求のために利用されるデータセットは、国あるいは地域の保険加入者のデータから入手したものから、特定の疾患を有する個人のデータに基づくものまで多岐にわたる¹⁴⁾。

診療報酬請求書のデータベースはきわめて多様であり、以下に関する事項の一部もしくはすべてを含んでいることが多い¹⁴⁾ (ICDのコード (ICD-CM) を利用した診断的情報、診療報酬のコード、患者および医療提供者の人口統計的情報と歯科技工指示書および処方箋の情報)。

研究と目標の立案のために診療報酬請求書のデータを活

用することは、保険制度の特性に左右される。1つの決定因子はそのシステムが医療機関の診療報酬のためにさまざまなレベルで要求するデータの範囲であり、支払いの質が保証されているかどうかである。たとえば、登録された患者につき医療機関に定額が支払われる人頭払い制では、治療内容ごとに支払われるシステム (行つたすべての治療の詳細を記載する必要がある) に比較して、診療報酬請求書に記載が求められる臨床データはより少ないと考えられる。

口腔保健におけるビッグデータ活用の実例

口腔保健医療の分野では、一般的に記録と診療報酬の両方の目的で電子データが収集されている。

口腔保健の調査では、確立された市場経済において過去60年間に口腔保健の改善が認められたことを報告している。多くの調査においてWHOの基準が測定指標として使用されている¹⁵⁾。このような調査は、臨床データや診療報酬請求書のデータからは得られないデータを収集するのに有用であり、患者の履歴や嗜好性を調べるのに役立ち、さらには疾患の関連性を特定し、仮説を立てることも可能である。調査の有効性は被験者の抽出法によって影響を受ける。母集団の中から代表的な被験者を無作為に抽出する必要があり、しかも被験者からは参加の同意を得なければならないので、その確保は容易ではない。

さまざまな調査が蔓延している現代においては、口腔保健の調査で高い返答率を得ることは困難である。さらに、歯科医院を訪れないような人々が口腔保健の調査に参加するとは考えにくい。一定の母集団における口腔保健を評価するための別の補助的な手段は、行政のデータベースや診療報酬請求書のデータベースからデータを引き出すことである。

アイルランドにおいてGuineyら¹⁶⁾は、あまり裕福ではない成人の被験者において、国の疫学調査で推定された治療ニーズを、その後同じ母集団に対して行われた治療のレベルと比較した (治療は第三者機関が出資した保険制度により提供されたものである)。Guineyらは、被験者が抽出された母集団に対して提供された治療に対する診療報酬請求書のデータベースを利用して、実際の治療レベルを算出した。その結果、調査では高い治療ニーズが見込まれたにもかかわらず、その後実際に行われた治療とはかなりの格差があったことが明らかになった。

別の分析においてGuineyら¹⁷⁾は、成人就業者の保険制度における1997~2008年の治療傾向について調査している。12年間の診療報酬請求書のデータベースを遡って調査した結果、患者1人あたりの平均治療数はこの期間で2%減少しており、残存歯数の増加と修復歯数の減少が認められ、行われた検査および予防処置は毎年増加していた。このようなデータは医療の“現実的な”評価基準となり、口腔保健医療の策定の変革が、提供される治療や処置にどのような影響を与えるかをモニタするのに有用であり、口腔保健の動向を示すものとして活用することが可能である。さらにこれらのデータから、今後の調査に役立つ仮説を立てることもできる。

英国国民健康保険の一般歯科診療の診療報酬請求書などのデータは、公正さをチェックする目的で収集されることもあり、Lucarottiら^{18,19)}とBurkeら^{20,21)}はこれらのデータを分析して、修復物の耐用期間に影響をおよぼす因子について調査した。その結果、患者が歯科医師を替えた場合や、患者の齲蝕リスクが高く歯科医院に頻繁に通った場合には、修復物の耐用期間が短くなることが明らかになった^{18,19,21)}。しかし、診療報酬請求書のデータを用いて治療結果を分析することには、問題がないわけではない。Lucarottiら¹⁸⁾は、患者の診療報酬請求書を一定期間追跡することには、以下のような限界が存在することを指摘している：追跡期間中の患者の氏名変更、自費診療の選択、患者の移住あるいは死亡、紙媒体の請求書からデータベースへ変換する際のミス、観察期間中における患者の転院もしくは通院中止。診療報酬請求書のデータ分析では、一定期間の治療結果を実証するのにレトロスペクティブな観察アプローチが利用されるが、治療結果に影響を与える他の要因については、データ解釈において考慮されない。Lucarottiら¹⁸⁾が指摘した前述の要素の例としては、歯科医師と医療提供者との相互作用のために生じる治療介入の選択、臨床的技能、口腔衛生状態や食事といった患者側の要素なども含まれる。

口腔保健医療のデータベースは、治療介入の影響が医療従事者の行動におよぼす影響を調べるためにも利用されている。Rindalら²²⁾はPBRN(診療に基づく研究ネットワーク)の参加者および非参加者の診療パターンの変化について調査した。彼らが注目した結果の1つは、PBRNへの参加が最新のエビデンスの導入に影響を与えるかどうかであった。歯科の電子カルテと行政のデータベースを組み合わせ使用したところ、PBRNへの参加は日常の診療にエビデンスを広く導入するのに効果的であり²²⁾、PBRNにより深くかかわっている医療従事者ほどその傾向が顕著であることが明らかになった。

口腔保健との関連性の調査を目的として、多くの研究がビッグデータ(あるいはビッグデータベース)を活用するために進められている。これらの研究は、Jeffcoatら²³⁾が指摘しているように、口腔の健康と全身の健康との関連性などに関する新たな仮説を立てるのに有効である。たとえばJeffcoatら²³⁾は、注意深く抽出されたコホートにおける口腔保健医療と全身保健医療の保険に関するデータをリンクさせ、2005～2009年の期間において、ベースライン時の歯周治療が、5つの全身疾患に関連する医療費および入院とどの程度関連性があるのかについて推定した。歯周治療群は、ベースライン時の2005年に歯周疾患と診断され4回の歯周治療を受けた群と定義された。対象となった全身疾患は、II型糖尿病、冠動脈疾患、脳血管障害、関節リウマチ、妊娠への影響である²³⁾。その結果、歯周治療群では、II型糖尿病、冠動脈疾患、脳血管障害、妊娠への影響において医療費および入院が有意に低減した($p < 0.05$)が、関節リウマチでは影響は認められなかった²³⁾。この分析は、2005年に全身疾患あるいは妊娠への影響のうち1つを診

断された患者と、2005年に歯周疾患と診断された患者(2004年に歯周治療を受けていない患者に限る)の338,891人の診療報酬請求書に基づいている。しかし、歯周治療群の試験対象患者基準により、治療介入群は全身疾患を有する患者の約1%に限定された。

しかし、Jeffcoatらの研究²³⁾は、健康的な生活態度や全身疾患の治療に対するコンプライアンスについての情報が不足しているため、改善された結果をもたらした真の交絡因子を除外できていないことを意味している。治療介入群と定義された1%の患者は、1年間に4回の歯周治療のための通院を完遂した患者であるため、みずからの病気をコントロールするにあたり生活態度や行動を改善するための動機づけや治療コンプライアンスがきわめて高い群であると考えられる。試験対象患者基準である歯周治療の完遂は、メンテナンス治療に対して良好なコンプライアンスを示す患者のマーカー(目印)となっていた可能性があり、これが医療費と入院の低減に寄与していたかもしれない。Jeffcoatらの研究²³⁾に基づく仮説を精査するには、さらなる研究が必要である。

Fellowsら²⁴⁾はDPBRN(歯科診療に基づく研究ネットワーク)において、572,606名のヘルスプランメンバーの電子保険請求書を利用し、顎骨壊死とビスホスホネート(BP)製剤使用およびその他のリスク因子との関連性について調査した。コホートのうち21,164名が経口BP製剤の治療を受けた。全群の中で23例が顎骨壊死を発症し、そのうち6例(35.3%)が経口BP製剤を服薬していた。単変量解析によると、経口BP製剤使用患者は非使用患者より顎骨壊死の罹患率が15.5倍高かった(信頼区間6.0～38.7)²⁴⁾。これらの結果により、経口BP製剤による顎骨壊死の絶対リスクは低いことが明らかとなった。

これらの結果を解釈するには注意が必要であり、関連性を明確にするためにはさらなる研究が必要と思われる。というのは、著者らも指摘しているように、顎骨壊死症例が少ないために回帰分析の結果が限定され、適応による交絡を排除することが困難だからである。また彼らは、将来的な研究は2007年の新しいICDコード“無菌性骨壊死、顎骨”を導入することで促進されるであろうと述べている。この新しいコードは顎骨壊死症例を特定するのに有用であるが、将来的には発症年月日、リスクへの曝露、治療期間の情報管理も顎骨壊死の決定因子に関する理解を深めるために重要となる。

後者の2つの例では、口腔保健と全身保健の保険データの潜在的価値と現時点での限界がわかりやすく示されている。関連性と因果関係を調べる際にはこれらのデータセットの限界を明確にすることが、将来のデータベースの設計と開発に役立つはずであり、疾患の決定因子と治療の影響に関する重要な疑問に答えを出せるようになることが期待される。

モバイル機器のデータを歯科患者の電子カルテと組み合わせて利用することで、患者の診療を改善できると考えられている¹²⁾。近年では歯科機器メーカーが電動歯ブラシ

と連携したアプリを開発しており、ブラッシング時のデータを活用してブラッシングの技術を向上させることができ²⁵⁾、患者のケアの改善にきわめて有用であることが示されている。このようなデータの活用は、患者に直接フィードバックすることにより、口腔衛生の予防策を導入するプロセスを支援すると思われる。

口腔保健のデータが幅広く電子化されると同時に、膨大なデータベースの処理能力が向上している現状を背景として、現在使用可能な口腔保健システムのデータを活用して歯科疾患の予防に焦点を当てて口腔保健の改善をはかるためのプロジェクトが、ヨーロッパの研究者によって展開されている。

ADVOCATE プロジェクト

ADVOCATE (Added Value for Oral Health) プロジェクトは EC のホライゾン 2020 プログラムの助成を受けており、より患者本位でエビデンスに基づいた予防的な口腔保健医療システムへ移行するためのストラテジーを創出することを目的としている²⁶⁾。このプロジェクトは EU 加盟諸国との協力関係を基盤として、各大学、民間保険会社、ドイツ・デンマーク・ハンガリー・オランダ・アイルランド・英国の公立医療機関、スコットランドのバイオ医療情報科学会社である Aridhia との共同研究により、患者への動機づけとなるフィードバックや経済的評価および政策提言を通じて、各加盟国における口腔保健医療システムを評価し、口腔保健医療の成果の評価基準を提示し予防的アプローチを促進することを目的としている²⁶⁾。

ADVOCATE プロジェクトにおけるビッグデータの活用

プロジェクトの全体的な目標は、予防を重視した患者中心の口腔保健システムへいかにして移行するかを検討することであり、そのための手段として、2つのデータソースの調査と分析が必要である。1つは行政の健康保険のデータであり、もう1つは歯科診療時に携帯電話やタブレットのアプリを使って患者が自己申告する電子質問票のデータである²⁶⁾。この分析により、提供された歯科医療とその成果が評価され、GDP へのダッシュボードとして可視化されることで、他の患者と比較してどのような医療がどの程度提供されたかがフィードバックされ、予防を重視した処置を行って患者の口腔保健を改善するために、患者の内的動機づけを高めることができる²⁷⁾。EU 加盟 6 か国の健康保険データの別の分析により、外的動機づけの理論を踏まえて、予防的な口腔保健医療の医療機関支払い制度と患者自己負担制度が評価されるようになると考えられる²⁶⁾。この考え方は、さまざまな経済的報酬や支払い制度(診療内容ごとの料金や人头払い制など)を最も効率よく組み合わせ、医療の質を向上させ効率的・効果的な疾患予防を推進する方法を明らかにしようとする試みである^{28,29)}。ADVOCATE プロジェクトの核心は歯科医師にもっと予防的な処置を患者に提供しようとする試みなので、EU 加盟 6 か国との共同研究を通じて口腔保健医療に長期的な好影響を与えることが見込まれている。それは健康増進のみならず、より低侵襲の治療によってもたらされ

るものであり、患者中心の総合的な口腔保健システムのための有益な提言が導き出され、歯科医師、為政者、患者、保険会社、さらには社会全体にとってプラスとなる影響をおよぼすことが期待されている。

ADVOCATE プロジェクトは EC のホライゾン 2020 研究・イノベーションプログラムの助成を受け委託研究契約 635183 (<http://advocateoralhealth.com>.) に基づいて実施されている。

文献

- 1) Frost & Sullivan : Drowning in Big Data? Reducing Information Technology Complexities and Costs for Healthcare Organizations. Frost and Sullivan. USA.[Online].
- 2) Boyd D, et al : Critical Questions for Big Data. Information, *Communication & Society*, [Online]. 15(5) : 662-679, 2012.
- 3) Raghupathi V, et al : An Overview of Health Analytics. *J Health Med Inform*, [Online]. 2013.
- 4) Bian JT, et al : Towards Large-scale Twitter Mining for Drug-related Adverse Events. *SHB. Proceedings of the 2012 international workshop on Smart health and wellbeing*. [Online]. 25-32, 2012.
- 5) Raghupathi W, et al : Big data analytics in healthcare : promise and potential. *Health Inf Sci Syst*, 2(3) : 1-10, 2014.
- 6) Manyika J, et al : Big Data : The next frontier for innovation, competition, and productivity. McKinsey and Company. [Online]. 2011.
- 7) Miller K : "Big Data Analytics in Biomedical Research." *BCR*, [Online]. 2012.
- 8) Chen H, et al : Business Intelligence and Analytics : From Big Data to Big Impact. *MIS Q*, [Online]. 36(4) : 1165-1188, 2012.
- 9) Gelfand A : Privacy and Biomedical Research : Building a Trust Infrastructure. *BCR*, [Online]. 23-28, 2011.
- 10) Lin Y, et al : Data Mining Large-Scale Electronic Health Records for Clinical Support. *IEEE Intell Syst*, [Online]. 26(5) : 87-90, 2011.
- 11) DiGangi RDH : Dental Electronic Health Records : Meaningful and Useful. *Academy of Dental Therapeutics and Stomatology*, [Online]. 94-99, 2012.
- 12) Basole RC, et al : Data and Analytics Challenges for a Learning Healthcare System. *ACM J Data Inf Qual*, [Online]. 6, article no : 10 [no pagination], 2015.
- 13) Liu K, et al : Using electronic dental record data for research : a data-mapping study. *J Dent Res*, [Online]. 92(7 Suppl) : 90S-96S, 2013.
- 14) Stein JD, et al : use of health care claims data to study patients with ophthalmologic conditions. *Ophthalmology*, 121(5) : 1134-1141, 2014.
- 15) World Health Organisation : Oral Health Surveys Basic Methods. World Health Organisation. [Online]. 2013.
- 16) Guiney H, et al : Comparing epidemiological estimated treatment need with treatment provided in two dental schemes in Ireland. *BMC Oral Health*, [Online]. 12(31) : 1-10, 2012.
- 17) Guiney H, et al : Analysis of a payments database reveals trends in dental treatment provision. *J Dent Res*, 92(7 Suppl) : 63S-69S, 2013.
- 18) Lucarotti, PS, et al : Analysis of an administrative database of half a million restorations over 11 years. *J Dent*, 33(10) : 791-803, 2005a.
- 19) Lucarotti PS, et al : Outcome of direct restorations placed within the general dental services in England and Wales (Part 3) : variation by dentist factors. *J Dent*, 33(10) : 827-835, 2005.
- 20) Burke FJ, et al : Outcome of direct restorations placed within the general dental services in England and Wales (Part 4) : influence of time and place. *J Dent*, 33(10) : 837-847, 2005.
- 21) Burke FJ, et al : Outcome of direct restorations placed within the general dental services in England and Wales (Part 2) : variation by patients' characteristics. *J Dent*, 33(10) : 817-826, 2005b.
- 22) Rindal BD, et al : Practice change toward better adherence to evidence based treatment of early dental decay in the National Dental PBRN. *Implement Sci*, 9 : 177, 2014.
- 23) Jeffcoat MK, et al : Impact of periodontal therapy on general health : evidence from insurance data for five systemic conditions. *Am J Prev Med*, 47(2) : 166-174, 2014.
- 24) Fellows JL, et al : ONJ in two dental practice-based research network regions. *J Dent Res*, 90(4) : 433-438, 2011.
- 25) Oral B : Oral B website, [Online]. 2016.
- 26) Leggett H, et al : Toward More Patient-Centered and Prevention-Oriented Oral Health Care : The ADVOCATE Project. *JDR Clin Trans Res*, [Online]. 2016.
- 27) DePasque S, et al : Effects of intrinsic motivation on feedback processing during learning. *Neuroimage*, 119 : 175-186, 2015.
- 28) Birch S, et al : The economics of oral health and health care. Munich (Germany) : Max Planck Institute for Social Law and Social Policy. Discussion Paper No. 7, 2015.
- 29) Brocklehurst P, et al : The effect of different methods of remuneration on the behaviour of primary care dentists. *Cochrane Database Syst Rev*, 6(11) : CD009853, 2013.