

「ヘルステックの情報社会論」研究会 第2回レポート

講演：「医療分野におけるリアルワールドデータ活用の現状と課題」

講師：石川ベンジャミン光一

(国立研究開発法人国立がん研究センター 社会と健康研究センター 臨床経済研究室長)

パネリスト：宮田俊男 (日本医療政策機構 エグゼクティブディレクター)

庄司昌彦 (国際大学 GLOCOM 主任研究員・准教授)

日時：2016年3月31日(木) 19:00～21:00

会場：国際大学グローバル・コミュニケーション・センター

主催：国際大学グローバル・コミュニケーション・センター (GLOCOM)

メドピア株式会社

医療におけるリアルワールドデータは、多様な種類の大量のデータを迅速に分析することを可能にする。中でも DPC データは患者の基礎情報・検査・治療の実施記録、医療施設の情報がセットになっており、マイクロレベル、マクロレベルで様々な分析を可能とする。本講演では、DPC データの整備および分析の第一人者である石川ベンジャミン光一氏が、その意義や分析事例を紹介した。DPC データの活用は、医療費・医療事務の効率化、医療サービス提供の実態把握、地域医療計画の策定などに貢献する。さらに他の医療関連データベースとの連携強化が進めば、医療や保険だけではなく、臨床研究・政策立案などの点で非常に利益の大きなものになるだろう。

講師の石川ベンジャミン光一氏が所属する「国立がん研究センター社会と健康研究センター」では、ビッグデータを用いた病気の予防、社会的文脈における医療提供のあり方を研究している。今回は石川氏に、医療におけるリアルワールドデータとは何か、そしてそこから見えてくるマイクロレベルでの医療提供の実態とマクロレベルでの地域医療の実態、さらに今後の課題等を伺った。

■ 医療におけるリアルワールドデータとは何か

ビッグデータは、大量かつ多様な情報の迅速な処理を可能にする。これまでの医学研究は、非常に限られたデータと狭い環境で行われてきた。しかし、リアルワールド（実世界）データは、従来のデータがカバーしない広範な領域を扱えるビッグデータである。自然発生するデータを悉皆的に捉えようとするものであり、どのタイミングでどのような検査が必要だったのか、結果的にどのような状況になったのかなど多様なデータを集め、そこから様々な知識を得ることができる。

医療の中で使われる情報には様々なものがある。患者の来院から医療サービスの提供に際して発生するデータは医療施設ごとの情報システムに蓄積されている。近年では SNS やスマホを活用した社会的な情報もある。様々なデータの関連性を評価しながら結びつけ、新しい知見を見つけていくことができると、今までは見落とされてきたデータをきちんと体系的に扱えるようになると期待される。

医療サービスの提供の中でデータはどのように出てくるのか。まず、患者が受診すると、その患者の状態について必要な検査をする。医師による検査の指示（オーダー）、次に様々なスタッフによる検査の実施、そして得られた結果の記録の 3 つのステップがある。次の段階では、これらをもとに医師が診断を確定し、実際に患者の治療にあたる。この場合も必要な治療の指示、その実施、結果の記録の 3 ステップがある。こうした情報は、それぞれ「オーダー」「実施記録」「カルテ（＝結果。医師の所見を含む）」としてそれぞれの病院の中の情報システムに記録される。

この中で一番分析しやすいのは、オーダーである。いつ誰にどのような検査・治療を行うかという 1 つのオーダーに対して 1 つのアクションが対応している。しかし、その検査・治療を中断したということもありえるが、このデータだけではわからない。そのため、患者に対して実際に行ったことを抜き出すのが次の段階である。ただし、実施記録では各検査を誰がどのように行ったかが細かく記録されているが、医療機関を超えた分析は難しい。さらに、患者の年齢や検査結果の医学的意味を求めるならば、カルテの情報まで必要となる。医療のリアルワールドデータを見ていくときに、通常は、まず何をやったかを評価できればいいという段階があり、さらにその意味づけのために臨床情報を付加していくという段階がある。

オーダー、実施記録、カルテ以外には、医療施設が対価を請求するための保険請求情報もある。保険請求システムは最も古くから導入されている。次にオーダーの記録システムが導入され、業務効率化のための実施記録・業務支援システムが導入され、約 10 年前にようやくカルテのシステム化が進んだ。しかしカルテは標準化が進んでおらず、分析・活用が難しい。最も整備されている診療報酬請求システムでさえ分析に使づらい。

そこで石川氏らの研究チームは、患者の基礎情報、検査・治療の実施記録、医療施設情報をまとめた **DPC データ**を開発した。開発の背景には政策的事情もある。DPC データは「急性期の入院医療の費用を支払うための包括評価制度」の中で開発・普及された。通常、一人の患者に対して行った検査や治療をレセプトにし、診療報酬を請求するが、必要ない治療も請求可能で医療費や請求事務の増加につながる。

それに対し DPC (Diagnosis Procedure Combination=診断と治療内容の組み合わせに基づく患者分類) では、ある診断に対して行われる診療行為を分析しグルーピングすることで、その病気の患者に対して払うべき医療費の算出と事務の効率化ができる。平成26年時点では2900を超える病院(全7439病院の39%)で年間1166万件(全退院数1441万件の81%)のデータを集めるまでになった。DPCデータは医療実態の情報公開のために厚労省への提出が義務付けられている。その際、データは匿名化され住所は郵便番号に変えられている。厚労省はこれらを元に、2年に1度診療報酬改定を行う。また、最新の2014年度のDPCデータから全国の2900の病院のどこでどのくらいの患者を受け入れ、どのような治療がなされているのか、などがわかるようにしている。これらのデータは一般社団法人診断分類研究支援機構によって運用され、倫理指針に基づいて、データのコピーが研究活動での利用に開かれている。

DPCデータとして各病院で収集されるのは、以下のような情報である。まず、患者基礎情報として保険者(保険請求先)、ID(匿名化)、性別、生年月日、住所(郵便番号)、入退院経路である。また、診療記録(カルテ)として病名、身長・体重、がんなど一部の病気の重症度である。これらはまとめて「様式1(退院サマリ)」と呼ばれている。細かな検査結果は膨大な量になり、またどの段階の情報を評価するかによって結果が異なるため扱いが難しく集められていないが、治療・検査のための細かなオーダー(診療指示)については「E/Fファイル」にまとめられている。ほかのリアルワールドデータであるレセプト(診療明細)と比べ、DPCデータは患者の基本的な情報と検査・治療の結果がすべて含まれているため、様々な形での悉皆的な分析が可能になる。また、全国がん登録と比較すると、患者の基礎情報が含まれるのは同じだが、がん登録は頭名のままである。またがん登録は初回受診時の診療内容が記録されるのみで、がん患者の生存期間が近年長くなる傾向にあるものの、その後の転移・再発・再入院等に関する情報がなく、一定期間後の生存情報に限られる。

このように、分析に使える医療データは大きくDPC、レセプト情報、全国がん登録の3つがあるが、それぞれの特質がある。中でもDPCは患者(入院した者のみ)の基本情報と、検査・治療の実施内容の詳細が悉皆的に把握できるという利点がある。

■ DPCデータのマイクロレベルでの利用事例

石川氏はDPCデータ活用の事例として、頭蓋内腫瘍摘出術を取り上げた。例えば、医療施設ごとの脳腫瘍の症例数のデータからは、3年間で症例数が年間10例未満の施設が半数以上を占め、ほとんどの患者が非常に症例数の少ない施設で治療を受けていることがわかる。単一施設のシステム内のデータをもとに症例検討などを行うことが、いかに少ないデータに基づかざるを得ないかがわかる。

さらに、1日単位での全診療内容がわかるので、いつどのような薬剤が使われているか、かかっているコストはどの程度か、そのシェアなども把握できる。例えば、脳内留置用剤ギリアデルは一人当たり100万円かかるもので、全体で4億2000万円以上の医療費が投入されているが、実際に使われている患者は6.7%にすぎない。また、新しい薬や技術の活用状況も分かる。例えば、アラベルは脳腫瘍の場所を示す診断薬であり、レザフィリンは脳腫瘍の場所を示すと同時にレーザーを当てるとその部分の腫瘍が死滅するという新しい薬である。前者はおよそ1割の手術で使われており、費用は約9万円である。一方、後者はそのための設備等にも費用が掛かり、38.7万円シェアはまだ0.5%にとどまっている。しかしこのような医療の進歩をデータとして扱えるようになることで、例えばレザフィリンの使用がアラベルの

何倍になった時に、医療費はどのくらいかかるようになるのかを予測することができる。

このように、DPC データからは、どのような薬・治療がいつどのくらい使われ、その費用やシェアがどれくらいか、どのくらいの入院が必要かといったことがわかる。これをもとに、医療費の試算や、医師と製薬会社の取り分のバランスなど、様々な分析が可能になるのである。

■ DPC データを利用した地域医療の分析

こうしたデータは厚生労働省のウェブサイトで公開されている。現在のわが国における入院医療については最も活用できるデータであり、地図上で分析すれば地域社会の医療実態が明らかになるため、地域医療保健医療計画・地域医療構想の策定に向けた重要な資料となっている。

例えば、二次医療圏における早産・切迫早産が年 10 例以上の施設に関するデータでは、宮城県内では仙台赤十字病院や仙台医療センターなどが症例数・病床数ともに多く、早産・切迫早産、子宮がんなどを診療することができる病院については「車で 30 分以内」の条件で人口の 70%以上がカバーされていることがわかる¹。

ここには私立病院や診療所は含まれておらず、それらも含めれば、県人口のほとんども緊急時も含め医療が行き届いていることがわかる。過疎地域も含めてすべての地域で病院設立が必要というわけではなく、既存の病院を生かし、また妊産婦自身や救急が自ら病院を把握しておく等によって、より多くの人口に医療を届けることが可能になる。利用目的や方法が明確な地域医療へのデータ活用といえる。

また、人口推計と入院患者の受療率のデータを使うと、地域ごとの入院患者数の期待値を算出できる。例として、急性心筋梗塞に関する宮城県内の地域別医療対策を示す。仙台圏で全国平均並みに急性心筋梗塞の患者が発生すると、2015 年の段階では約 45 人の患者が予測される。DPC データではすでに 35 床の病床がこの地域にあり、地域で患者の 80%近くをカバーできるということになる。一方、石巻・登米・気仙沼地域では、予想される患者数が 15 人ほどであるのに対し、病床数は 3 床にすぎず、カバー率は 22%である。これをもとに、広域医療を整備するのか、地域医療を整備するのかといった計画を立てる際のスタート地点となる現状を知ることができる。

こうした DPC を用いた分析をもとに、地域医療構想では、仙台で何床の病床が必要なのか、仙台で不足する分は周辺地域でどのくらいカバーしなければならないのかといった医療計画が検討されている。

■ DPC データの利点と欠点

DPC データを分析する上での利点は、大きく分けて 2 点ある。第一に、ミクロレベルの分析の観点から、入院中の診療プロセスの詳細な把握に最も役に立つ。DPC 調査に協力している範囲が明確なので、調査・分析を行う際に母数とその属性が非常に明確であり、病院ごと、術式別、薬剤の利用状況などの詳細な分析・比較が可能になる。第二に、マクロレベルの分析においては、DPC データの二次利用によって病院の機能と診療圏／需給状況から、地域の医療カバー率等の分析に役立てることができる。

一方、DPC データだけでは不十分な点も 2 点指摘できる。第一に、治療成績の評価である。DPC データは基本的に一医療機関での診療内容を把握するだけなので、当該病院に入院中の検査・治療の記録は収集することができるが、退院後や転院前後の状況は、死亡退院を除いては限定的にしか把握できず、

¹ ここでは全人口を対象として中・大規模病院のカバー率を算出している。出産可能年齢の女性に限って、個人入院などを含めれば、カバー率は変動することになる。

長期的な治療実態はつかめない。第二に、詳細な臨床病期・重症度、部位別の状態、あるいは抗がん剤治療後の副作用などの細かな臨床データは、逐一医師の手によって記録することが難しい。

これらを踏まえて、データの連携が活用の鍵となる。DPC データを参照することで、どの病院にどのような病気の患者がどのくらいいて、どのような治療がされているかを知ることができる。これをもとにデータを分析し、さらに必要なデータの追加・補完を依頼しながら、分析の効率化・精緻化が可能になるだろう。現状では不十分なことも多いが、次のステップのための手掛かりとするため活用することが可能になってきている。データの深度を増すためには今後のデータの連携の強化が重要である。

■ 国による医療（ビッグ）データの整備状況

国による医療に関する電子データ収集の枠組み整備は進んできている。DPC データ、全国がん登録、ナショナルレセプトデータベースなどのほか、介護保険総合データベース、副作用報告データベース、死亡者情報票データベースなどがある。しかしその実態は、収集者が異なっていたり、フォーマットが異なっていたり、匿名化の有無が異なっていたりという状況で、それぞれのプロセスが個別に行われ、業務が効率化されず、各データベースが孤立している状況である。これらは相互参照が可能な情報も重複していたりし、ビッグデータとしてのメリットを発揮できていない。

これらを横断的に参照し、ビッグデータとして扱うために、石川氏が提案するのは医療用データを国に提出する際に用いる「汎用匿名化 ID」の導入である。匿名化しないままで連結するのは不可能だが、共通 ID が付されてあれば、後から各データベースのデータを連結して用いることは可能である。近年の個人情報保護法改正やマイナンバー制度の導入では、医療用 ID をどのように制度化するかについてはまだ議論が進んでいない。導入したとしても運用の安全性や、診療記録の標準化によるデータ管理業務の簡易化など、乗り越えなければならない課題は多い。その解決のために必要なのは、医療データの「エコシステム」、つまり診療報酬請求・政策利用・研究などのデータの高度利用を前提とした第三者機関の管理・運営による医療データの流通基盤の整備である。こうした現状と課題を踏まえたうえで、医療ビッグデータ、リアルワールドデータが広範に活用されるようになれば、その利益は非常に大きいだろう。