

調査と情報—ISSUE BRIEF—

No. 1005 (2018. 5. 8)

医療ビッグデータの利活用

—医療情報の二次利用に向けた政府の取組—

- | | |
|------------------------------|----------------------|
| はじめに | 2 DPC データ |
| I 医療ビッグデータの概要 | 3 その他のデータベース |
| 1 医療ビッグデータとは | III 利活用の促進に向けた政府の取組 |
| 2 我が国の医療を取り巻く環境と医療ビッグデータ | 1 既存データベースの一元化への取組 |
| 3 医療情報の一次利用と二次利用 | 2 様々な医療情報の利用の活性化への取組 |
| II 既存データベースの二次利用の現状 | おわりに |
| 1 レセプト情報・特定健診等情報データベース (NDB) | |

- 医療ビッグデータ（人の健康、病気、治療等に関する膨大な量のデータの集まり）は、医療の質の向上や効率化、医療・健康分野の研究開発等に向けた二次利用において、大きな効果が期待されている。
- 診療報酬明細書（レセプト）等のデータを蓄積したデータベース（DB）が既に複数運用されている。しかし、これらの DB は連結されておらず、利用できる者も限定されているなど、利用効果は必ずしも十分なものとなっていない。
- 政府は、既存 DB の一元化や未利用の医療情報の活用のための基盤整備によって、二次利用促進を図っている。人の病歴等の機微に関わる医療情報の取扱いに細心の注意を払いつつ、医療ビッグデータの利活用が進むことが期待されている。

国立国会図書館 調査及び立法考査局

社会労働課 おかぼやし くすひろ 岡林 楠博

第 1005 号

はじめに

近年、様々な分野において、膨大な量のデータの集まりを指す「ビッグデータ」が注目を集めている。医療・健康分野においても例外ではなく、人の健康、病気、治療等に関するビッグデータは、医療の質の向上や効率化、医療・健康分野の研究開発等に資するものとして期待されている。

また、政府は、「健康・医療・介護」を経済成長の重点分野の一つとしている。具体的には、現在、別々に管理・運用されている医療・健康分野等のデータベースを連結して、保険者や研究者によるデータ分析を促進させるとともに、国の認定を受けた機関が、医療機関が保有する様々な医療情報を収集し、民間企業や研究機関に提供する仕組みの制度化を行うなど、医療ビッグデータの活用推進のための基盤整備を行っている。

本稿では、医療ビッグデータを蓄積する主な既存データベースを紹介するとともに、二次利用の促進によって医療ビッグデータの利活用を目指す政府の取組について整理する。

I 医療ビッグデータの概要

1 医療ビッグデータとは

ビッグデータとは、前述したように膨大な量のデータの集まりのことであり、一般に、データが多量（volume）、データの種類・形式が多様（variety）、データの発生・更新速度が迅速（velocity）という三つの特徴を有しているとされる。「ビッグデータ」という呼称の背景には、情報量が増えすぎて、コンピュータによる分析が困難であるという事情があるとされており¹、実際、従来のビッグデータは、巨大すぎて管理や分析が困難であるため活用されてこなかった²。しかし、近年は、情報通信技術（Information and Communication Technology: ICT）の進歩に伴って、その管理、分析が可能となり、その利活用によるメリットが認識されるようになっていく。例えば、日々生み出される消費者の購入履歴やインターネット上の検索履歴等のデータが蓄積・分析され、市場予測や景気調査等にかかされている³。

医療・健康分野でも、主に事務処理の効率化を目的として電子化が進み、現在では、日々膨大な量の医療情報が電子的に生成、蓄積されるようになっていく。こうした医療情報の例として、診療報酬明細書データ（以下「レセプトデータ」という。）や電子カルテなどに記録された診療データ、特定の疾患に関する臨床データ、薬局における調剤データ、健康診断データなどが挙げられる（以下、人の健康、病気、治療等に関するビッグデータを「医療ビッグデータ」という。）。また、比較的新しいタイプの医療ビッグデータとして、ウェアラブル機器等（リストバンド型の活動量計など）により人の行動をつぶさに記録した情報（例えば、心拍数、血

* 本稿は平成30（2018）年4月24日までの情報を基にしている。インターネット情報への最終アクセス日も同日である。

¹ ビクター・マイヤー＝ショーンベルガー、ケネス・クキエ（斎藤栄一郎訳）『ビッグデータの正体—情報の産業革命が世界のすべてを変える—』講談社、2013、p.17。（Victor Mayer-Schonberger and Kenneth Cukier, *BIG DATA*, 2013.）

² 中山健夫「医療ビッグデータ」福原俊一編集代表『医療レジリエンス—医学アカデミアの社会的責任—』医学書院、2015.3、p.66.

³ 城田真琴「ビッグデータを押さえる」『読売新聞』2017.8.15.

圧、歩数) や、遺伝子情報 (ゲノム情報) などがあり、近年、利活用が期待されている⁴。

2 我が国の医療を取り巻く環境と医療ビッグデータ

我が国の医療は、厳しい財政状況の下、超高齢社会の到来や医療技術の進歩等による医療費の増大という課題に直面している。医療を取り巻く環境が年々厳しさを増している中、政府は、これまでのように、単なる負担増と給付削減によって医療制度を維持するのではなく、質の高い医療サービスをいかに効率よく国民に提供するかという視点に立ち、新たな医療システムの構築を図っていくとしている⁵。

また、政府の成長戦略である「日本再興戦略」改訂 2015⁶においては、国民の利便性の向上という観点から、平成 32 (2020) 年までの 5 年間で「集中取組期間」として、医療等分野における ICT 化 (ICT の活用) を徹底的に推進することとされた⁶。その後、成長戦略として新たに策定された「未来投資戦略 2017」でも、医療ビッグデータの利活用については、同様の考え方が引き継がれている⁷。厚生労働省は、平成 29 (2017) 年 1 月に「データヘルス改革推進本部」を設置し、保健医療データを十分に活用することにより国民の健康的な生活を実現する「データヘルス改革」⁸を推進することとしている。このように、医療ビッグデータの利活用は、政府の成長戦略を進める上での重点施策の一つとされている。

3 医療情報の一次利用と二次利用

医療情報の利用には、大きく分けて、一次利用と二次利用がある。一次利用とは、患者の診療等によって得られた情報を本人の治療のために使用するなど、取得した本来の目的で使用するものである。一方、二次利用とは、取得した本来の目的以外の目的で使用するものである。医療情報の一次利用と二次利用のそれぞれの目的、利用主体等は、表 1 のとおり整理される。

個人が生涯にわたり自分自身に関する医療・健康情報を収集・保存して活用できる仕組みであるパーソナルヘルスレコード (Personal Health Record: PHR)⁹や、地域における複数の医療機関の間で患者の診療情報をネットワークで共有する地域医療情報ネットワークは、患者本人の診断や治療を目的としているため一次利用に該当する。一方で、治療の過程で収集した臨床

⁴ 山田亮「ライフサイエンス領域におけるビッグデータの利活用」『週刊医学界新聞』No.3107, 2015.1.5, p.8. <http://www.igaku-shoin.co.jp/paperDetail.do?id=PA03107_04>

⁵ 厚生労働省は、今後の医療政策の考え方として、「単なる負担増と給付削減による現行制度の維持を目的とするのではなく、新たな価値やビジョンを共有し、イノベーションを取り込み、システムとしての保健医療の在り方の転換をしなければならない時期を迎えている」(「保健医療 2035」策定懇談会「保健医療 2035 提言書」2015.6, p.9. 厚生労働省ウェブサイト <http://www.mhlw.go.jp/seisakunitsuite/bunya/hokabunya/shakaihoshou/hokeniryou2035/assets/file/healthcare2035_proposal_150609.pdf>)としている。

⁶ 「「日本再興戦略」改訂 2015—未来への投資・生産性革命—」(平成 27 年 6 月 30 日閣議決定) pp.20-22. 首相官邸ウェブサイト <<http://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/dai1jp.pdf>>

⁷ 「未来投資戦略 2017—Society 5.0 の実現に向けた改革—」(平成 29 年 6 月 9 日閣議決定) pp.39-44. 同上 <https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/miraitousi2017_t.pdf>

⁸ 「データヘルス改革—ICT・AI等を活用した健康・医療・介護のパラダイムシフトの実現—」(第 7 回未来投資会議資料 5) 2017.4.14. 同上 <<http://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/dai7/siryou5.pdf>>

⁹ 現状、個人の医療・健康情報は、個人、医療機関、自治体等がそれぞれ所有しており、紙資料やデータの形でばらばらに存在している。総務省は、これらの情報を一元化し、本人によるアクセスや持ち運びを可能とするためのアプリケーションの開発や、情報連携基盤の構築を推進している(総務省「総務省における医療 ICT 利活用基盤の構築等について」(未来投資会議構造改革徹底推進会合 第 6 回「医療・介護—生活者の暮らしを豊かに」会合資料 3) 2017.3.9. 同上 <https://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/suishinkaigo_iryokaigo_dai6/siryou3.pdf>)。

データを新たな治療法の開発のために活用したり、調剤データや診療データを組み合わせることで薬の副作用調査に活用したりすることは二次利用に該当する。医療情報の二次利用を促進する上では、情報のビッグデータ化が有効となる。

医療情報の二次利用は多くのメリットがある一方で、個々の医療情報は、個人の病歴や健康状態等の機微に関わるものであるため、漏えい、悪用等が起こると本人に対する不当な差別や偏見等が生じるというデメリットもある。こうした性格を踏まえ、これまで、①二次利用する際の患者本人からの同意取得の方法、②個人が特定されないようにする匿名化の方法など、個人情報保護の観点から、様々な検討がなされてきた。近年では、医療情報の有用性を損ねることのないよう、個人情報保護とのバランスの取れた技術開発、制度設計、組織的運用等の必要性が指摘されている。¹⁰

表 1 医療情報の一次利用と二次利用

	一次利用	二次利用
目的	患者本人の診断や治療	医療・健康分野の研究や開発、薬の副作用調査、行政における統計作成や政策立案
利用主体	医療機関	研究機関、行政機関、企業 など
個人情報の扱い	個人情報（患者本人の情報）を含む	個人を特定されないように匿名化された情報の形で利用する。
例	パーソナルヘルスレコード（Personal Health Record: PHR）、地域医療情報ネットワーク など	レセプト情報・特定健診等情報データベース（National Data Base: NDB）やDPCデータの研究目的での利用、全国がん登録 ^(注1) 、臨床データベース（National Clinical Database: NCD） ^(注2) など

(注1) がんの診断や治療を受けた患者の情報を登録するデータベース。

(注2) 全国の外科手術・治療情報を登録するデータベース。

(出典) 小尾高史ほか「プライバシーを考慮した医療情報の活用とその実現に向けた課題」『電子情報通信学会誌』98(3), 2015.3, p.208, 表1を基に筆者作成。

II 既存データベースの二次利用の現状

データを要素ごとに整理して蓄積したデータベースの構築は、ビッグデータの一元管理及び利活用を可能とする。医療情報を蓄積したデータベースは、我が国において、既にいくつか運用されている。そのうち、行政主導で整備が進められ、具体的に活用が進められている代表的なものとして、レセプト情報・特定健診等情報データベース（ナショナルデータベース（National Data Base: NDB）。以下「NDB」という。）がある。患者を、病名と行われた診療行為などから分類する診断群分類（Diagnosis Procedure Combination: DPC）に用いるデータ（以下「DPCデータ」という。）も、収集、蓄積が進んでいる。本章では、これら二つの特徴を整理し（表2参照）、二次利用の拡大に向けた現状と課題について述べる。

¹⁰ 小尾高史ほか「プライバシーを考慮した医療情報の活用とその実現に向けた課題」『電子情報通信学会誌』98(3), 2015.3, p.211; 山本隆一「日本における医療ビッグデータの利活用の現状と課題」『公衆衛生』79(9), 2015.9, pp. 615-618.

1 レセプト情報・特定健診等情報データベース（NDB）

(1) NDB の概要

NDB は、厚生労働省保険局が管理・運用しているデータベースであり、レセプトデータと特定健診・特定保健指導¹¹のデータが蓄積されている。医療分野ではデータ蓄積量が最大規模を誇るデータベースである。平成 18（2006）年の医療制度改革において、「老人保健法」（昭和 57 年法律第 80 号）が改正され、「高齢者の医療の確保に関する法律」（以下「高齢者医療確保法」という。）として平成 20（2008）年 4 月に施行された。高齢者医療確保法の中で、国及び都道府県の医療費適正化計画の作成等に資するため、保険者等がレセプトや特定健診等のデータを国に提供することとされたこと（高齢者医療確保法第 16 条第 2 項）に伴い、平成 21（2009）年から NDB の運用が開始された。

医療機関や特定健診実施機関で生成されたレセプトデータと特定健診等のデータは、レセプトの審査等を行う社会保険診療報酬支払基金等の審査支払機関を経由して、NDB のレセプト情報サーバ及び特定健診等情報サーバにそれぞれ提供され、蓄積される。提供の際には、氏名や被保険者証の番号などの個人を特定できる情報は削除され、代わりに、識別子（ID）が割り振られることで匿名化処理がなされる¹²。

現在のデータ数は、レセプトデータが約 128 億件（平成 21 年 4 月～28 年 12 月）、特定健診等データが約 1.9 億件（平成 20～27 年度）となっており、毎年、各々、約 20 億件及び約 0.3 億件が新たに蓄積されている¹³。

(2) 二次利用の現状

NDB は、国及び都道府県における医療費適正化計画や地域医療構想の作成等のため、各地域の必要病床数の推計や、患者の受療動向の分析の根拠データとして利用されている¹⁴。高齢者医療確保法という国によるデータ収集の根拠規定を持ち、専ら行政による政策立案のために運用される NDB は、情報の悉皆性（しっかいせい）が高く、データの蓄積が容易である。このことから、NDB のデータは、二次利用による効果が期待されており、政府において、二次利用の仕組みが検討されてきた¹⁵。政府の検討を経て、平成 23（2011）年度からは、必要な審査手続を経ることを条件として、研究者等に対して NDB のデータの提供が認められることとなった¹⁶。

一方で、二次利用をする上での問題点も指摘されている。NDB は、保険請求事務に用いるレセプトデータの集合であるため、医療行為についての情報が主であり、患者の臨床情報は最低

¹¹ 生活習慣病の予防のために、40 歳から 74 歳までを対象として行われているいわゆるメタボ健診と、同健診の結果から、専門スタッフ（保健師、管理栄養士など）が生活習慣を見直すために行う保健指導のこと。

¹² 岡本和也「ナショナルデータベース（NDB）の解析—胃癌治療の実態調査を目的としたレセプト情報データウェアハウスの構築—」『医学のあゆみ』259(7), 2016.11.12, p.756.

¹³ 厚生労働省保険局医療介護連携政策課保険システム高度化推進室「第三者提供の現状について」（第 37 回レセプト情報等の提供に関する有識者会議資料 2）2017.5.17, p.1. <<http://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000165139.pdf>>

¹⁴ 石川ベンジャミン光一「医療ビッグデータの価値を高めるには—生成・蓄積・利用の 3 局面から考える—」『Monthly IHEP』266 号, 2017.8, p.19.

¹⁵ 検討の内容については、「「医療サービスの質の向上等のためのレセプト情報等の活用に関する検討会」報告書」2008.2.7. 厚生労働省ウェブサイト <<http://www.mhlw.go.jp/shingi/2008/01/dl/s0130-16a.pdf>> を参照。

¹⁶ 「高齢者の医療の確保に関する法律第 16 条第 2 項の規定に基づき保険者及び後期高齢者医療広域連合が厚生労働大臣に提供する情報の利用及び提供に関する指針」（平成 22 年厚生労働省告示第 424 号）

限しか含まれていない。つまり、NDB のデータだけでは、患者のその後の病状（治癒したのか死亡したのかなど）がわからない。また、識別番号が割り振られていない一部の病気については、データ中に病名が保存されず、患者の病名がわからない¹⁷。さらに、NDB の二次利用やデータの公開には、厳しい制限が課せられている。前述したとおり、研究者等へのデータの提供は認められているものの、データの機微性のゆえに、利用に当たっては高いセキュリティ要件が課され、さらに利用目的や利用範囲等に関する厚生労働大臣の審査・承認¹⁸を経なければならない。このような手続が、二次利用が進まない一因となっているとの指摘もある¹⁹。

このほかにも、NDB が保有するデータを分析する上で問題点が指摘されている。例えば、一部のレセプトデータと特定健診等データは突合ができないことが問題であるとされている²⁰。また、NDB で個人を識別するために使用する識別子（ID）は、氏名や被保険者証の番号などから生成されるが、結婚して姓が変わる場合や、転居や転職によって被保険者番号が変わる場合には、ID が変化してしまう。このため、個人の継続的な識別が困難となり、長期間にわたる個人の分析ができない²¹。また、特定健診・特定保健指導データについても、同様の課題があるほか、健診を受診した者のデータしか収集しないため、未受診者の健康に関する分析ができないなどの欠点がある²²。

このように、分析に利用する上でいくつか問題点はあるものの、NDB の二次利用や、そのためのデータの公開については、正確な根拠に基づく施策を推進するに当たって有益であるとの期待から、厚生労働省において、NDB の二次利用促進に向けた検討がなされている。現在、より多くの方が NDB データを利用することができるよう、汎用性の高いデータとして、都道府県別や性・年齢階級別に集計した基礎的な集計表が公表されている（NDB オープンデータ）²³。さらに、一定の申請・承認手続により NDB を利用することができる「レセプト情報等オンサイトリサーチセンター」を東京大学等に開設し、試行的な運用を開始している²⁴。

2 DPC データ

(1) DPC データの概要

診断群分類（DPC）は、患者を、病名と行われた診療行為などから分類するものであり、医

¹⁷ 藤森研司「レセプトデータベース（NDB）の現状とその活用に対する課題」『医療と社会』26(1), 2016.4, p.16. <https://www.jstage.jst.go.jp/article/iken/26/1/26_15/_pdf>

¹⁸ 厚生労働省に設置された「レセプト情報等の提供に関する有識者会議」は、NDB のデータ利用の公益性等について検討・意見交換を行い、厚生労働大臣が申請者に対するデータ提供の可否を決定するに当たり、助言することとされている（「レセプト情報等の提供に関する有識者会議」開催要綱（第1回レセプト情報等の提供に関する有識者会議資料1）2010.10.5. 厚生労働省ウェブサイト <<http://www.mhlw.go.jp/stf/shingi/2r9852000000thao-att/2r9852000000the0.pdf>>）。

¹⁹ NDB の二次利用が進まない原因については、満武巨裕「日本のレセプト情報・特定健診等データベース（NDB）の有効活用」『情報処理』56(2), 2015.2, pp.140-144 に詳しい。

²⁰ 満武巨裕「レセプトビッグデータ解析の現状と将来」『実験医学』34(5)（増刊）, 2016.3, p.60.

²¹ 岡本 前掲注(12)

²² 満武 前掲注(20)

²³ 厚生労働省「NDB オープンデータ」<<http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000177182.html>> において、第1回 NDB オープンデータ（平成26年度のレセプト情報と平成25年度の特定健診情報を集計）及び第2回 NDB オープンデータ（平成27年度のレセプト情報と平成26年度の特定健診情報を集計）が公表されている。

²⁴ 同リサーチセンターの概要、活用に向けた課題等については、「みんなが使える「NDB」を、京大病院の黒田氏 オンサイトリサーチセンターは「今のままでは運用に乗らない」」『日経デジタルヘルス』2017.11.27. <<http://tech.on.nikkeibp.co.jp/atcl/feature/15/327442/112600150/?ST=health&P=3>> に詳しい。

療の質を相対的に評価するために開発された手法である。厚生労働省は、平成 15 (2003) 年に、この DPC を用いた診療報酬の包括評価制度²⁵を導入している。同制度の対象となっている病院（以下、厚生労働省から同制度の対象として認定された病院を「DPC 対象病院」という。）は、対象となる入院患者に関する診療報酬を請求する際に、厚生労働省に対して、診療録情報やレセプト情報を含むデータを提出することとされている²⁶。これに基づいて提出されているデータが DPC データである。²⁷

DPC データには、主に、診療録情報とレセプト情報が含まれている。診療録情報として、患者基本情報（生年月日、性別）、入退院情報（入退院日、入退院経路、予定入院又は緊急入院の別等）、診断情報（診断名、入院時併存症、入院後合併症）、手術情報（手術日、術式、麻酔法）が記録されている。また、レセプト情報として、出来高払により支払われる診療報酬の請求情報（個別の診療行為ごとの診療報酬点数が手技料、薬剤費、材料費に区分されて、日別に記録されている。）と、包括的に支払われる診療報酬の請求情報が記録されている²⁸。

DPC データは、政府の成長戦略である「「日本再興戦略」2016」の中短期工程表²⁹に従い、平成 29 (2017) 年にデータベースが構築された。現在、DPC データは、調査年度ごとに DPC データベースに蓄積されるようになっている³⁰。

(2) 二次利用の現状

DPC データは、基本的な診療録情報とレセプト情報が長期間にわたって時系列で記録されており、NDB に比べて情報のバランスに優れているという利点を持つ³¹。また、データ形式が簡易であり、単純な集計から複雑な分析まで、幅広く利用することができると言われる³²。こうした長所を生かして、DPC データを用いた医療の質の評価手法等が開発されている³³ほか、治療法の適切な選択など臨床研究でも成果が上がっている³⁴。

一方で、DPC データを二次利用する上での問題点もある。DPC データは、DPC 対象病院等からのみ提出されるため、NDB に比べて悉皆性が劣る³⁵。また、NDB と同様に、DPC データ

²⁵ 入院患者を疾病や入院日数ごとに分類し、入院 1 日当たりや 1 入院当たりの診療報酬を算定する制度。我が国では、入院患者の入院基本料、検査、画像診断、投薬等に関する費用が包括評価の対象とされ、実際に投入された医療行為の多寡にかかわらず、DPC に基づく点数によって診療報酬が包括的に支払われる。

²⁶ 「厚生労働大臣が指定する病院の病棟における療養に要する費用の額の算定方法」（平成 20 年厚生労働省告示第 93 号）第 5 項第 3 号

²⁷ DPC データについては、松田晋哉『基礎から読み解く DPC—実践的に活用するために— 第 3 版』医学書院, 2011 に詳しい。

²⁸ 同上, p.16.

²⁹ 「「日本再興戦略」2016—第 4 次産業革命に向けて—」（平成 28 年 6 月 2 日閣議決定）首相官邸ウェブサイト <https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/zentaihombun_160602.pdf> の「中短期工程表」において、平成 29 年度より DPC データの一元管理及び利活用を可能とするデータベースのシステム運用を開始するとされた（「日本再興戦略 2016 中短期工程表」2016.6.2, p.26. 同 <https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/kouteihyo_160602.pdf>）。

³⁰ 平成 29 年度 DPC データ提供事業者「DPC データの提供について」2017.12.12, pp.4-5, 9-11. 厚生労働省ウェブサイト <<http://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000188034.pdf>>

³¹ 石川ベンジャミン光一「DPC などの医療ビッグデータ分析の現状と課題」『臨床検査』60(9), 2016.9, p.1005.

³² 石川 前掲注(14)

³³ DPC データの最大の利用事例として、各都道府県が策定する地域医療構想における医療需要予測が挙げられる（畠井正深「DPC が果たす医療政策の展開—これまでの足跡と医療政策へのインパクト—」『病院』76(2), 2017.2, p.103.）。

³⁴ 伏見清秀「DPC データの解析」『医学のあゆみ』259(7), 2016.11.12, p.761.

³⁵ ただし、近年は、DPC 対象病院が増加しており、この問題については改善傾向が見られる。

は、患者の詳細な症状等の情報を含まないため、臨床研究に利用する際には、何らかの方法で追加情報を入手する必要がある³⁶。

DPCデータの二次利用については、厚生労働省において、医療機関ごとの集計データが一部を除いて既に公表されており、一般の利用が可能である³⁷。また、利用申請者の指定した条件や集計条件により抽出されたデータは、平成29年度より、「DPCデータの提供に関するガイドライン」³⁸に基づき、公益性の高い学術研究に対して、一定の処理を加えた集計データの形で提供を行うこととされた。この申請・審査手続はNDBのもの³⁹を準用する形となっている。

表2 NDB及びDPCデータの特徴

	NDB	DPCデータ
データの種類	レセプトデータ、特定健診等データ	診療録データ、レセプトデータ など
データ数	レセプト約128億件（平成21年4月～28年12月）、特定健診等約1.9億件（平成20～27年度）	年間約1000万件
管理者	厚生労働省	厚生労働省
データの提供元	審査支払機関（社会保険診療報酬支払基金、国民健康保険団体連合会）	DPC対象病院、DPC準備病院 ^(注) など
データの提供先	国、都道府県、研究機関への提供、NDBオープンデータの公開	中央社会保険医療協議会、研究機関への提供、医療機関ごとの集計表の公開
主な利用目的	医療費適正化計画、地域医療構想の作成 など	DPCに基づく分類の更新・精緻化、包括評価制度の見直し など
根拠	「高齢者の医療の確保に関する法律」（昭和57年法律第80号）	「厚生労働大臣が指定する病院の病棟における療養に要する費用の額の算定方法」（平成20年厚生労働省告示第93号）第5項第3号
二次利用における主な利点	データの悉皆性、容易な蓄積が可能 など	入院患者の診療録情報が含まれる、病院名付きで結果が公表されている、データの多彩な活用が行われている など
二次利用における主な問題点	臨床情報が最低限しか含まれていない、データの利用・公開が制限されている など	悉皆性が劣る、詳細な臨床情報は含まれていない など

(注) DPC制度に参加することを希望する病院のうち、要件となる基準を満たす病院のこと。「DPC準備病院」の基準を2年間満たした上で、厚生労働大臣が参加を認めると「DPC対象病院」となる。

(出典) 石川ベンジャミン光一「医療ビッグデータの価値を高めるには—生成・蓄積・利用の3局面から考える—」『Monthly IHEP』266号, 2017.8, p.19を基に筆者作成。

3 その他のデータベース

「国保データベース（KDB）」は、各都道府県に設立された国民健康保険団体連合会が管理・運営する健診、医療、介護の各種データを蓄積したデータベースである。医療情報は、主に国

³⁶ 飯原弘二「脳卒中医療の質向上に資するDPCデータ活用」『病院』76(2), 2017.2, p.117.

³⁷ 「DPC導入の影響評価に関する調査：集計結果」厚生労働省ウェブサイト <<http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000049343.html>> において、集計データが公表されている。

³⁸ 厚生労働省「DPCデータの提供に関するガイドライン」2017.8. <<http://www.mhlw.go.jp/file/06-Seisakujouhou-12400-000-Hokenkyoku/0000186364.pdf>>

³⁹ 「レセプト情報等の提供に関する有識者会議」（前掲注(18)参照）による審査。

民健康保険に関するレセプトに限定されているものの、健診や介護データなどとの連結が可能である。集計・分析結果は、国民健康保険の保険者である市町村等に提供され、住民に対する保健指導に用いられるなど、保健事業のための利活用が進んでいる。⁴⁰

がんの診断や治療を受けた患者の情報を登録する「全国がん登録」⁴¹や、全国の外科手術・治療情報を登録する「臨床データベース (National Clinical Database: NCD)」⁴²は、臨床情報を蓄積したデータベースの代表例である。このほかにも、臨床情報を蓄積したデータベースがいくつか構築されている⁴³。

近年、遺伝子情報 (ゲノム情報) を収集、解析して、患者一人ひとりの体質等に合わせた治療を行う個別化医療を推進するため、ゲノムコホート (ゲノム情報と生活習慣や環境がどのように病気と関連するかを調べる研究) データベースもいくつか構築されている⁴⁴。

III 利活用の促進に向けた政府の取組

医療ビッグデータの利活用に向けた政府の取組のうち、①現在、別々に管理・運用されている既存データベースを連結し、一元化することにより、データ分析等を促進させるもの、②これまで二次利用されてこなかった病院等の医療機関が保有する様々な医療情報の利用を活性化させるものの二つの取組を取り上げ、その概要、課題等について述べる。

1 既存データベースの一元化への取組

我が国においては、国民皆保険制度の下、医療・健康分野に関する豊富なデータが存在しており、前述したように、既にNDBやDPCデータなどの大規模なデータベースが運用されている。しかし、これらのデータベースは連結されておらず、データにアクセスできる利用者の範囲も限定されているなど、二次利用による効果は必ずしも十分なものとなっていない。

そのため、厚生労働省は、「データヘルス改革」の一環として、健康・医療・介護分野のデータを一元化した「保健医療データプラットフォーム」を平成32(2020)年から本格稼働させることを目指しており、平成30(2018)年度以降に、詳細な設計に着手することとしている⁴⁵。

⁴⁰ 鎌形喜代実「保健事業推進のための国保データベース (KDB) システムの活用」『保健医療科学』63(5), 2014.10, pp.463-466.

⁴¹ 全国がん登録は、従来、都道府県ごとの登録であったが、「がん登録等の推進に関する法律」(平成25年法律第111号)により、全国で統一した登録になった。都道府県をまたがって診断・治療を受けた患者の情報も正しく把握することができるようになるなど、より正確ながんのデータを集めることを目指している(「全国がん登録とは」国立がん研究センターウェブサイト <https://ganjoho.jp/reg_stat/can_reg/national/public/about.html>)。

⁴² NCDは、手術症例データベースとして、平成22年に、日本外科学会等の外科系臨床学会が連携して設立したデータベースである。平成28年までに、900万件を超える症例が登録されており、臨床研究のための情報基盤として活用されているほか、専門医の申請・資格更新の際に診療実績を証明するための基盤としても活用されている(「NCDについて 沿革」NCDウェブサイト <<http://www.ncd.or.jp/about/history.html>>)。

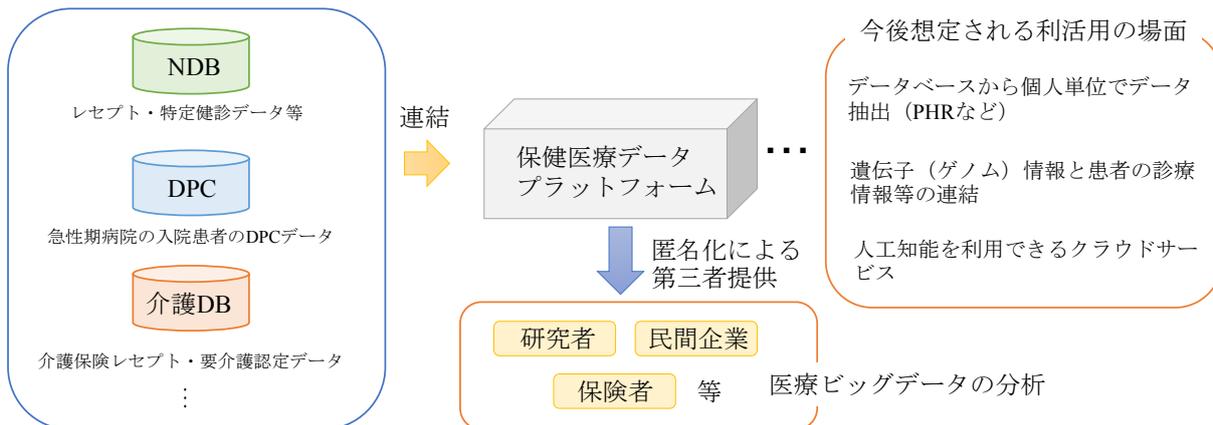
⁴³ 医薬品の副作用を直接把握・評価することを目指して電子カルテやレセプト等を収集する「MID-NET (Medical Information Database Network)」や、脳卒中症例の登録データベースである「J-ASPECT Study」などがある。

⁴⁴ ゲノムコホートデータベースの構築を目指す事業として、「東北メディカル・メガバンク計画」などがある。同計画は、被災地の地域医療再建・健康支援とゲノム医療の実現を目的として、宮城・岩手両県の被災地住民を対象に、15万人規模で全遺伝子情報と生活習慣・病歴情報のデータベース化を行うものである(日本医療研究開発機構バイオバンク事業部ほか「東北メディカル・メガバンク計画全体計画 改定版」2017.4.1. 日本医療研究開発機構ウェブサイト <<https://www.amed.go.jp/content/000006771.pdf>>)。

⁴⁵ 厚生労働省等は、平成29年7月に、医療ビッグデータの活用推進等に係る計画・工程表を公表した(厚生労働省

実現すれば、NDB や DPC データなどの既存データベースが連結されることとなる。また、平成 32 (2020) 年の本格稼働後には、管理しているデータを製薬会社等の民間企業や大学等の研究機関に提供 (以下「第三者提供」という。) する仕組みも計画されている。これには、健康・医療・介護分野のそれぞれのデータベースの情報を一元管理し、健康な時から病気にかかり治療や介護を受けるまでの一連の状況を分析し、医療や介護の質を向上させるなどの狙いがある (図 1 参照)。

図 1 「保健医療データプラットフォーム」の概念図



(出典) 「データヘルス改革—ICT・AI等を活用した健康・医療・介護のパラダイムシフトの実現—」(第7回未来投資会議資料 5) 2017.4.14. 首相官邸ウェブサイト <<http://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/dai7/siryoushou5.pdf>> を基に筆者作成。

個人単位でデータを一元化するため、平成 32 (2020) 年の「保健医療データプラットフォーム」の稼働開始と同時に、個人一人ひとりを識別する「医療等 ID」の運用を開始することが計画されている。現在、その前段階として、基本的に世帯単位となっていた被保険者番号を個人一人ひとりに割り当て、同番号を「医療等 ID」に活用することが決まっている。「医療等 ID」を「個人番号カード (マイナンバーカード)」に記録し、健康保険証の機能を持たせ、オンラインによる健康保険の資格確認が行える仕組みの構築も計画されている⁴⁶。

「保健医療データプラットフォーム」の本格稼働に向けた課題としては、セキュリティ対策の徹底、データの分析や管理のための人材の養成、データベースの開発や運用に要する費用負担への対応等が挙げられる⁴⁷。同プラットフォームの管理・運用を行うことが予定されている社会保険診療報酬支払基金と国民健康保険中央会においては、新たにデータ分析、セキュリティ対策等に従事する人員が必要とされ、厚生労働省と連携して人材の育成に取り組む方針である。データ分析の能力を持つデータサイエンティストのうち、医療・健康分野の知識を有する者は

ほか「国民の健康確保のためのビッグデータ活用推進に関するデータヘルス改革推進計画」2017.7.4. 厚生労働省ウェブサイト <<http://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000170006.pdf>>。
⁴⁶ 患者が病院等を受診する際、健康保険の資格を証明するため、「医療等 ID」が記録されたマイナンバーカードを提示することが想定されている (「病院でもマイナンバーカードが保険証代わり」『読売新聞』2017.1.3)。
⁴⁷ 「データヘルス改革」の課題については、鈴木康裕ほか「データヘルス改革で医療と介護はこう変わる (下)」『社会保険旬報』No.2699, 2018.1.11, pp.45-48; 西村周三「データヘルス改革の現状と課題 (上)」『社会保険旬報』No.2698, 2018.1.11 などに詳しい。

不足しているとされ、将来に向けてそうした人材の育成が大きな課題となる⁴⁸。

2 様々な医療情報の利用の活性化への取組

医療情報には、病院、診療所等で生成される電子カルテデータや、薬局で作成される調剤データなど多種多様なものが存在する。これらのデータは、NDBに記録されているレセプトデータなどとは異なり、これまで余り二次利用が進んでこなかった。しかし、近年、医療・健康分野における研究開発を加速させるため、広範な医療ビッグデータを組み合わせて分析する必要性が指摘されている。

医療情報の利用は、従前から「個人情報の保護に関する法律」（平成15年法律第57号。以下「個人情報保護法」という。）等⁴⁹の制約を受けていたが、平成27（2015）年9月に同法が改正され（以下、改正後のものを「改正個人情報保護法」という。）⁵⁰、病歴や診療結果といった医療情報は、「要配慮個人情報」⁵¹とされ、「オプトアウト方式」⁵²による第三者提供が禁止された。すなわち、医療情報を第三者提供する場合は、原則として、事前に患者一人ひとりに対して同意を得なければならないこととされた。

このため、改正個人情報保護法に対応しつつ、医療情報を円滑に第三者提供することができるよう法制度を整備することが求められ、平成29（2017）年4月に、医療情報の利用により医療等に関する先端的研究開発等を促進することなどを目的とする「医療分野の研究開発に資するための匿名加工医療情報に関する法律」（平成29年法律第28号。以下「次世代医療基盤法」という。）が成立した。これにより、国の認定を受けた認定機関（以下「認定事業者」という。）が、病院、薬局等の医療機関が保有する医療情報を収集し、匿名化して研究機関等に第三者提供することができる仕組みが制度化された。改正個人情報保護法によって制限されることとなった医療機関による医療情報の「オプトアウト方式」による第三者提供は、認定事業者向けに限り、例外的に可能となり、NDBやDPCデータ以外の医療情報の利用が活発になることが期待されている。

図2のとおり、認定事業者は、医療機関から個人情報を含む医療情報の提供を受ける。その際、患者本人は、あらかじめ自らの医療情報を第三者提供することについて通知されており、認定事業者への提供を拒否することができる。患者が拒否した場合は、提供は行われず。医療機関から提供を受けた医療情報は、認定事業者によって個人が識別されないように匿名加工され、匿名加工医療情報⁵³として第三者提供される。

⁴⁸ 伏見 前掲注(34), p.765.

⁴⁹ 民間分野については、個人情報保護法が適用される。公的分野については、国の機関には、「行政機関の保有する個人情報の保護に関する法律」（平成15年法律第58号）が、独立行政法人や国立大学法人等には、「独立行政法人等の保有する個人情報の保護に関する法律」（平成15年法律第59号）が、地方公共団体には、各地方公共団体の個人情報保護条例が、それぞれ適用される（日置巴美「健康・医療情報の活用と個人情報保護法制その他の関係法令（1）」『NBL』No.1098, 2017.5.15, pp.5-6.）。

⁵⁰ 「個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律」（平成27年法律第65号）による改正。

⁵¹ 本人に対する不当な差別、偏見その他の不利益が生じないようにその取扱いに特に配慮を要する情報（改正個人情報保護法第2条第3項）。

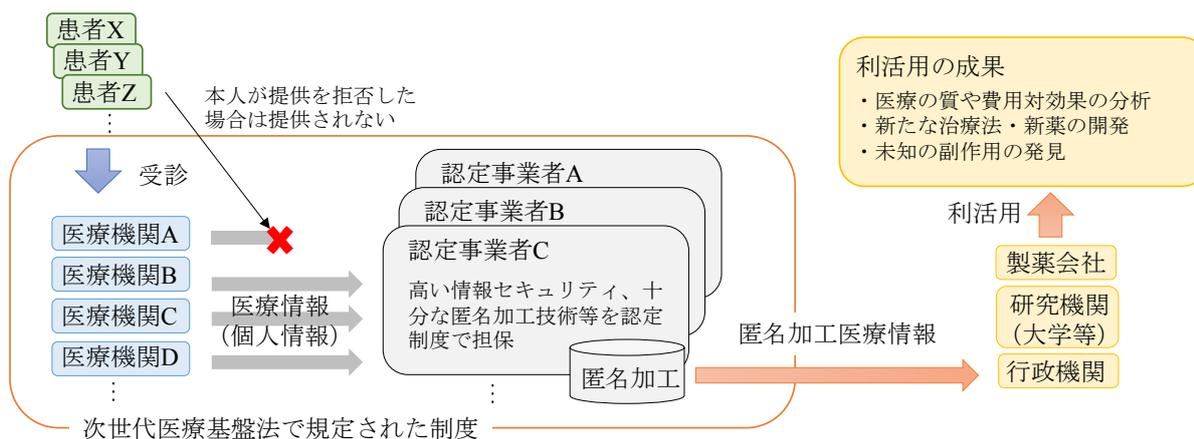
⁵² 「オプトアウト方式」とは、あらかじめ本人に対して個人データを第三者提供することについて通知又は認識し得る状態にしておき、本人がこれに反対しない限り、同意したものとみなし、第三者提供をすることを認めること。

⁵³ 「匿名加工医療情報」とは、特定の個人を識別することができないように医療情報（個人情報）を加工したもので、当該情報を復元することができないようにしたもの（次世代医療基盤法第2条第3項）。改正個人情報保護法にお

次世代医療基盤法の規定により、認定事業者は、高い情報セキュリティを確保し、十分な匿名加工技術を有することなどが求められている。政府は、平成 30 (2018) 年 5 月 11 日の次世代医療基盤法の施行に向けて、適切な事業運営に向けた基本方針や認定事業者を決定するための基準の策定などを進めている。同基本方針では、医療機関に対して、患者への情報提供の意思確認を最初の受診時に書面で行うことを求めるとともに、提供後も、個人を識別できる情報は可能な限り削除することを求めることができるなど、患者に不利益が生じないための措置を盛り込む方針である。⁵⁴

次世代医療基盤法による医療情報の利用については、いくつかの論点が指摘されている。例えば、認定事業者の倒産等による個人情報の散逸を防ぐために、認定事業者の事業の安定性を担保する必要があるが、基準を厳しくしすぎれば参加できる事業者が限定されてしまうため、規制の在り方が課題とされる⁵⁵。また、認定事業者が行う匿名加工に関して、どこまで加工を求めるかという点も検討を要する。加工に伴い、情報が削除されればされるほど、二次利用した際の有用性が失われるため、匿名性と有用性の両立が課題とされる⁵⁶。さらに、医療機関が認定事業者に医療情報を提供することは義務ではなく、十分な量の医療情報が蓄積されないことが懸念されるため、医療情報提供のインセンティブ設計の検討が提案されている⁵⁷。

図 2 次世代医療基盤法の概念図



(出典) 「医療分野の研究開発に資するための匿名加工医療情報に関する法律案の概要(次世代医療基盤法案)」内閣官房ウェブサイト <<http://www.cas.go.jp/houan/170310/siryoul.pdf>> を基に筆者作成。

ける「匿名加工情報」に相当する(改正個人情報保護法第2条第9項)。
⁵⁴ 「医療ビッグデータ法：患者の意思、初診で確認 情報提供の方針」『毎日新聞』(大阪) 2018.4.11; 「5000万人診療情報 集約」『読売新聞』2018.4.24。
⁵⁵ 「情報を極める 個人情報保護法⑨ 医療情報、新法で流通促進」『日本経済新聞』2017.7.3。
⁵⁶ 同上; 大井哲也「個人情報、扱いやすい形に 匿名加工の法整備で活用の幅広がる」『日経産業新聞』2016.11.25。
⁵⁷ 医療情報の提供に対するインセンティブとして、情報提供への診療報酬による評価や、事業者に対する国のコスト負担補助等が提案されている(日諸恵利「医療・介護・ヘルスケア分野における ICT化と情報利活用—個人情報のは、「守る」時代から「使う」時代へ—」『みずほ情報総研レポート』Vol.13, 2017, p.6. <https://www.mizuho-ir.co.jp/publication/report/2017/pdf/mhir13_ict.pdf>)。

おわりに

政府は、現在、医療・健康分野を我が国の経済成長のための重点分野の一つと位置付け、人工知能やビッグデータといった先端技術の活用を推進している。具体的には、これまで別々に管理・運用されていたNDB、DPCデータなどの既存データベースを一元化したり、これまで二次利用されてこなかった電子カルテデータなどの医療情報の利用を活性化させたりする取組を行っている。

従前から指摘されているように、医療ビッグデータは、個人の健康状態、病歴等の機微に関わる情報を多く含み、利活用に際しては、個人情報保護に細心の注意を払うことが大前提となる。二次利用の更なる推進に当たっては、セキュリティ対策の徹底、医療ビッグデータを扱うことのできる人材の育成、医療情報が流通する仕組みの確立等が課題とされる。

政府においては、個人情報保護と二次利用促進とを両立させるためのバランスのよい制度設計・運営が求められるとともに、国民が抱いている懸念の一つひとつに対応し、丁寧に説明していくことが必要である。今後に予定される「保健医療データプラットフォーム」の本格稼働や、次世代医療基盤法の施行により、医療ビッグデータの二次利用が促進され、我が国の医療の質の向上につながることを期待したい。