

特集：臨床研究実施の現状と課題

第2部 医療データベースの疫学研究への活用：医療データベースの研究利用と社会還元

レセプトデータベースの研究利用

国立保健医療科学院 医療・福祉サービス研究部地域医療システム研究分野

岡本悦司

2006年1月のIT新改革戦略が初めて示した「レセプトデータの学術的(疫学的)利用のため、ナショナルデータベースの整備及び制度的対応を2010年度までに実施する」という提言が、5年を経ていよいよ実現に達した。ナショナルデータベース(NDB)の研究利用をめぐる現状と動向について概観する。

1 法的根拠

NDBは正式には「レセプト情報・特定健診等情報データベース」と呼ばれ、高齢者医療確保法第16条「厚生労働大臣は……医療費適正化計画の作成、実施及び評価に資するため、次に掲げる事項に関する情報について調査及び分析を行い、その結果を公表する」を根拠にデータが収集される。この条文からも明らかなように、NDBは医療費適正化計画(メタボ健診、療養病床再編、医療計画等)の作成、実施、評価を目的とするものであって、IT新改革戦略が提言したような「学術的(疫学的)利用」を主目的とするものではない。

では、国以外の民間研究者による学術的利用を認める法的根拠はどこにあるか。

高齢者医療確保法に基づき「情報の利用及び提供に関する指針」(表1)を大臣告示(告示は通達とは異なり、官報に掲載され法律に準じる効力を有する)し、そのなかで例外的扱いを認める、という根拠構成となっている。このため、2007年7月より「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」が5回にわたって開催され、

表1 情報の利用及び提供に関する指針(2010年12月24日大臣告示)抜粋

第2(データの利用目的)に規定する場合を除き、データは管理責任者以外の者に提供してはならない。ただし、次のいずれかに該当する場合は、この限りではない。

①厚生労働省、その他の国の行政機関、地方公共団体等が、法令に定める所掌事務の遂行に必要な限度で、かつ、医療サービスの質の向上等を目指した正確な根拠に基づく施策の推進のために利用する場合であって、当該データの利用が公益性が高いものとして厚生労働大臣が承認した場合

②①に規定する以外の場合であって、①に規定する施策の推進に有益な分析・研究又は学術研究の発展に資する目的で行う分析・研究のためにデータを利用しようとする者が、当該分析・研究に必要な限度で当該データを利用する場合であって、当該研究の目的、研究計画、データの分析方法、データの使用・管理方法等について審査した上で、当該データの利用が公益性が高いものとして厚生労働大臣が承認した場合

(2)(1)のいずれかに該当する場合にあっては、利用目的の達成に必要な範囲で、必要に応じて加工を行った上でデータを提供するものとする。

2008年2月に最終報告書がまとめられた。この報告書にそって告示案が作成されパブリックコメントが2009年10月に実施されたが、政権交替等の影響もあって告示が出たのは1年以上遅れた2010年12月になった。表1はその抜粋であり、研究利用が例外的利用であること、そして利用の可否は厚生労働大臣が、利用目的、必要性、緊急性、人的体制、管理方法等について有識者会議の意見を聞いて判断すること、が示されている。

有識者会議は2010年10月にスタートし、情報提供の手続きや審査基準を定めたガイドラインが、パブリックコメントを経て2011年3月31日に公表された。

2 ガイドラインの内容

このほど公表されたガイドラインの内容を要約すると、次のようになる。

申出者の範囲は、国、都道府県等の行政機関、大学・大学院の研究者に加えて、医療保険者の中央団体（健保連等）、研究開発型独立行政法人（国立健康・栄養研究所等）、医療サービスの質の向上等をその設立目的の趣旨に含む国所管の公益法人（日本医師会や医療経済研究機構等）に所属する研究者ならびに研究の実施に要する費用の全部または一部を国の行政機関から補助されている者、とされる。重要な点は、申出をするのは研究者個人であって所属機関ではない、という点（所属機関の承認は必要だが）。製薬企業等の営利企業は、国からの補助事業である場合を除いて除外されている。

NDB に収録されるデータは匿名化されてはいるが、個人情報に準じた扱いとされ、申出者と、実際にデータに触れる利用者全員の誓約書を提出しなければならない（研究の一部を外部委託する場合も委託先職員の利用者の氏名を申出書に記載しなければならない）。

データ利用目的では、NDB では医療機関コード、保険者番号等は匿名化せずそのまま収集しているが、それらの研究利用は認められない（地域性の分析等、有識者会議が認める場合は例外とされるが、それでも医療機関は郡市区、保険者番号は都道府県まで）。違反した場合は、利用者と所属機関名の公表が行われる。また、他の個体識別が可能となる可能性があるデータ（たとえばがん登録）とのリンケージ（照合）も禁じられる。

審査基準では、必要性、緊急性、申出者の過去の実績、データ分析に係る人的体制に加えて「学術論文等の形で成果が公表される」ことが要求される。目的があくまで学術研究に限定する、という考えからである。

NDB は統計法に基づく統計調査ではなく、それゆえ統計法のデータ利用に関する規定も罰則も適用されない。統計法ではたとえば、提供された匿名データを「自己又は第三者の不正な利益を計る目的で提供し、又は盗用した者」には罰則が規定されている（第 61 条）。しかしながら、高齢者医療確保法には、研究者への個票提供は想定されていないので、この

ような罰則規定はない。そうした不備を補うため、ガイドラインは不適切利用に対する措置についても詳細に規定している。

3 データベースの内容

NDB に含まれるのは、電子化された全レセプト（紙レセプトおよび生活保護レセプトは含まない）ならびに特定健康診査・保健指導データである。後者は 2008 年度分からであるが、レセプトについては 2009 年 4 月診療分以降の収集となった。データ量は、2010 年 8 月末時点でレセプト約 15 億 9800 万件、特定健康診査・保健指導データ約 2065 万件蓄積されている。ファイルサイズは特定健康診査・保健指導データが 247 GB、レセプトは 1 月分で 167 GB、年間にすると約 2 TB という膨大なもの。

含まれる内容をレセプトにしばって述べると、基本的には紙レセプトと同一であるが、傷病や医薬品等が文字ではなくコード化され、一定のフォーマットに従った csv ファイルである。電子レセプトの様式については支払基金サイトに掲載されている「電子レセプト作成の手引き」を参照されたい。

http://www.ssk.or.jp/rezept/iryokikan/download/files/jiki_i01.pdf

氏名等の個人情報はすべて匿名化して収集されるが、重要な点は氏名等をハッシュ関数で暗号化し、同一人を追跡できるように作られている点である。これは特定健康診査・保健指導データとレセプトデータとを突合したり、2 年間のデータを比較して腹囲等への効果を測定したり、と「医療費適正化計画の作成、実施、評価」という本来目的にかなった仕組みであるが、同時に疫学研究にとってもきわめて有利な条件となる。ハッシュ値は図 1 のように 2 つ（各 64 ケタ）生成され、たとえば氏名が変更されても記番号が同一であれば追跡可能である。

なお、これは有識者会議でも議論されたことであるが、NDB には氏名等が匿名化されているほかはレセプトデータは医療機関から提出されたそのままのかたちで収載されており、分析には適していない。図 1 の①を見るとわかるように、各行には通番すらふられておらず、たとえばこのなかから男性だけのレセプト、東京都にある医療機関だけのレセプトを抽出することは不可能であることが理解できよう



(医療機関レコードはあるが、そのデータは各行に入っていない)。そのため、まずデータの各行に通番をふったり、またどの行がどのレセプトに属するかを示すレセプト ID を挿入したりする必要が出てくる。これらは申請した利用者の側で行わなければならない作業であり、そのためのプログラムも開発されてはいるものの、相当なデータ処理能力のある者でなければ、申請が承認されても実際の分析が困難になることが懸念される。

4 今後の見通し

2011～12 年度の 2 年間は試行期間とされ「提供依頼申出書の受付は、年に 4 回程度の一定期間に行う(ガイドライン案)」とされ、事務局や有識者会議の処理能力をふまえ各回 10 件程度が予定されている。実際にスタートするとどれくらいの申出があるかはわからないが、もし多数にのぼれば、処理がいつかなくなる懸念もある。

NDB の研究利用は米国、韓国、台湾の各国でもさかんに行われている。申請を受け付けるだけでなく、研究者による利活用を積極的に支援するため、たとえば米国は ResDAC (Research Data Assistance

Center), 台湾は国家衛生研究院といった支援体制まで構築している。人口 1 億 2000 万人のわが国 NDB は、少なくとも量的には世界最大級であり、それをわが国のみならず世界の「公益」に資する研究資源として活用していくためにも、こうした研究者への支援体制の充実が求められるであろう。

参考文献

- 1) 藤田利治ほか、厚生労働科学研究費補助金医薬品・医療機器等レギュラトリーサイエンス総合研究事業「レセプト等を利用した薬剤疫学データベース作成に関する研究」平成 21 年度総括・分担報告書、統計数理研究所 (2010 年 4 月)
- 2) 本田達郎ほか、厚生労働科学研究費補助金政策科学推進研究事業「医療ナショナルデータベースに関する諸外国の整備状況および日本におけるデータベースのあり方研究」医療経済研究機構 (2009 年 3 月)
- 3) 久保田潔、薬の安全性と医療データベース、ESTRELA 184 号 2～7 頁 (2009 年 7 月)
- 4) 岡本悦司、国における情報データベースの構築とその活用、薬剤疫学 2009 ; 14 (1) : 47-52.
- 5) 岡本悦司、医薬品安全対策とナショナルデータベース、JAPIC NEWS 319 号 2～3 頁 (2010 年 11 月)