
一般口演

一般口演13 データ分析システム

2017年11月21日(火) 16:15 ～ 18:00 G会場 (10F 会議室1006-1007)

[2-G-3-OP13-4] レセプト情報・特定健診等情報データベース（NDB）利用促進に向けた取り組み ―1入院1データ化―

明神 大也¹, 野田 龍也¹, 久保 慎一郎¹, 西岡 祐一¹, 東野 恒之², 今村 知明¹ (1.奈良県立医科大学 公衆衛生学講座, 2.(株)三菱総合研究所 ICTイノベーション事業本部)

【背景】レセプト情報・特定健診等情報データベース（NDB）の生レセプトは月単位で記票されており患者毎の入院期間が把握しづらいため、入院情報を解析するためには、1入院1データ化が必要となる。従来の方法は、DPCの生データと医科入院データをEF化した上で、DPC分析ツール上で処理していた。しかし全国1年分のNDBデータを対象とするとデータ容量が440GBから12.3TBと大きく増加する上、処理に約7ヶ月間を要するため、継続して実施することが現実的に困難であった。そこで本研究では、データ容量の増大を抑えて処理時間を大幅に短縮した1入院1データ化手法を開発した。

【方法】平成25年4月から平成26年3月のNDBデータ（医科入院・DPC）を用いた。医科入院とDPC出来高については、入院基本料の算定日を入院日とみなし、ID0・医療機関コード・病棟区分をキーにして、入院日が途切れるまでを1入院として再構成した。DPC包括についてはすべて入院日とみなした。摘要情報（診療行為/医薬品/特定器材レコード）についても上記の1入院化したテーブルとキーが繋がるように再構成を行った。全国1年分について処理したところ約3日間で完了し、処理後のデータ容量は元のレセプトの4倍程度に収まった。なお、ID0は当教室で開発し、ID1/ID2での名寄せを精緻化したものである。

【結果】平均在院日数を求めたところ、一般病棟では男性30.4日、女性35.6日、特定機能病院では19.4日だった。厚生労働省発表の平成27年度推計平均在院日数は、一般病棟で23.2日、特定機能病院で16.7日となっており、2年間の差を考慮すると、ほぼ信頼に足る結果と考えられた。

【結論】データ容量の増大を抑えて、1入院1データ化を高速に処理することが可能となった。将来的に、入院外レセプトにも適用し、1患者1データ化の実現につなげる予定である。

レセプト情報・特定健診等情報データベース(NDB)利用促進に向けた取り組み

ー1 入院1 データ化ー

明神 大也^{*1}、野田 龍也^{*1}、久保 慎一郎^{*1}、西岡 祐一^{*1}、東野 恒之^{*2}、今村 知明^{*1}

^{*1} 奈良県立医科大学 公衆衛生学講座、

^{*2} (株)三菱総合研究所 ICT イノベーション事業本部

Promoting More Effective Use of the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB)

- retrieving accurate hospitalization data -

Myojin Tomoya^{*1}, Noda Tatsuya^{*1}, Kubo Shinichiro^{*1}, Nishioka Yuichi^{*1}, Higashino Tsuneyuki^{*2}, Imamura Tomoyaki^{*1}

^{*1} Department of Public Health, Health Management and Policy, Nara Medical University,

^{*2} ICT Innovation Division, Mitsubishi Research Institute, Inc.

Background: The National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) is a comprehensive database of health insurance claims data under Japan's National Health Insurance system. The claims are classified by month, medical institution and patient. One of the NDB's problems is the difficulty of retrieving the hospitalization data when admission and discharge are in different calendar months and when more than one hospitalization occurs in a single calendar month. Conventionally, we converted the claims data to "EF files" and controlled by the specific software. However, the amount of data increased by about 40 times and it took about 7 months to convert one year of original data from all over Japan to EF files. Therefore, we developed a new method of retrieving the relevant hospitalization data which reduces the processing time and doesn't increase the amount of data.

Methods: Our study targeted data from health insurance claims filed between April 2013 and March 2014 for medical hospitalization and diagnosis procedure combinations (DPC.) We regarded the day charging a hospitalization fee in DPC comprehension and medical hospitalization and all days in DPC performance as the hospitalization days. We converted the consecutive hospitalization days to a hospitalization sequence. The sequence was by ID0, medical institution code and hospital ward classification. It spent about only 3 days on processing these data and the data amount increased only 4 times.

Results: In order to confirm the reliability of the converted data, we calculated the average number of hospitalization days; male in all hospitals except psychiatric wards 30.4days, female in all hospitals except psychiatric wards 35.6days, male and female in advanced treatment hospital 16.7days. Hospital report by Ministry of Health, Labour and Welfare said the estimated average number of hospitalization days was 16.5 days in general wards, 274.7 days in psychiatric wards and 29.1 days in all wards. The difference between our study and hospital report is thought to be due to the calculation method and the target period.

Conclusion: Our study succeeded a new method of retrieving the accurate hospitalization without increasing the amount of data for much shorter days.

Keywords: NDB, rezept, database, health insurance claims in Japan, patient identification

1. 緒論

わが国には DPC データやレセプト情報・特定健診等情報データベース(以下、NDB)等で、診療に関わる情報を収集する仕組みがあり、貴重なデータとして蓄積されている。それにも関わらず、データベースが各種の問題を抱えているため、十分に利活用が進んでいないと言いがたい状況である。この問題の解消を目指し、DPC・NDB の利活用の円滑化と高速化に向けた手法を開発するとともに、臨床研究への応用を図ることが昨今の喫緊の課題となっている。

NDB の問題点の 1 つに、NDB の DPC レセプトが特殊な形であること¹⁾が挙げられる。本来 DPC レセプトはレセプト共通(RE)レコードや保険者(HO)レコードなどの共通レコードと、包括行為レコードおよび出来高行為レコードに分かれる。診療行為(SI)レコード、医薬品(IY)レコード、特定器材(TO)レコードなどに含まれる出来高行為は一般レセプトと同様のレコードに記録されるが、包括行為の内容についてはコーディ

ング(CD)レコードの中に出来高で算定した場合の内容を記録することになっている。しかし NDB では、包括期間と出来高期間が混在する月のレセプトデータにおいて、包括行為と出来高行為をまとめて包括行為として包括行為レコードに取り込んでいる。その結果、入院期間を把握できなかったり、包括対象の診療行為と出来高の診療行為(処置・投薬など)を分離することができなかったりした。これを解決する手法を 1 入院 1 データ化と称した。

この課題に対して、藤森ら²⁾は DPC の生データと医科入院データを EF 化した上で、DPC 分析ツール上で処理する方法を開発・運用していた。同様の方法を全国 1 年分の NDB データで実施すると、データ容量が 440GB から 12.3TB と大きく増加する上、処理に約 7 ヶ月間を要した。そのため、継続して実施することが現実的に困難であった。

本研究では上記のうち、まず入院期間を把握できない点に着目した。レセプトは月単位・医療機関単位・患者単位で発行されている。入院基本料などが含まれる SI レコードには上

記問題が存在する。加えて、再入院を含め、入退院日の情報はレセプトの摘要欄にコメントとして記載されているが、NDBでは摘要欄のコメント情報は取り込まない。そのため、同一月に2回以上の入院や月またぎの入院では入院期間が把握できなかった。

そこで本研究では、データ容量の増大を抑えて処理時間を大幅に短縮することを前提にし、1 入院 1 データ化に至る 1 ステップである 1 入院期間の把握手法を開発した。

2. 方法

本研究では、医療機関および病棟区分が同一である入院を、1 入院期間と定義した。そのため、同一の医療機関内でも精神病棟から一般病棟に移った場合は入院期間が途切れることとなる。また病棟区分上、ICU・NICU・CCU・SCU など一般病棟に含まれる。

上記の定義を満たすために、平成 25 年 4 月から平成 26 年 3 月の NDB データの DPC レセプトおよび医科入院レセプトを用いた。DPC のうち包括期間の入院は、DPC レセプトの診断群分類 (BU) レコードの入院開始日と退院日 (または入院期間Ⅲの終了日) の情報を取得し、入院開始日から退院日までを入院算定日とした。なお、BU レコードは前述した包括行為レコードや出来高行為レコードではなく、共通レコードに含まれる。DPC のうち出来高期間 (入院期間Ⅲの終了日より後) は総括対象医科入院レセプトに記録された入院基本料または特定入院料が算定された日を入院算定日とした。DPC 指定病院以外・一般病棟以外の入院は、医科入院レセプトに記録された入院基本料または特定入院料が算定された日を入院算定日とした。

DPC の包括期間の有無にかかわらず、同一医療機関・同一病棟区分で、入院算定日が連続したものを 1 入院とした。また同一のレセプト内でも入院基本料または特定入院料が算定されていない期間があれば、別入院扱いとした。

なお患者の名寄せについては、当教室で開発し ID1/ID2 での名寄せを精緻化した ID0³⁾を用いた。

上記の条件を満たすために、SQL server 上で以下の処理を行った。例として 2013 年 4 月 3 日～6 日、4 月 29 日～5 月 1 日、5 月 22 日に入院した場合の処理を示した。

1. レセプト通番・医療機関コード・ID0・病棟区分・入院算定日を列にして、レセプト通番・医療機関コード・ID0・病棟区分ごとに、入院算定日順に並べた。また、レセプト通番・医療機関コード・ID0・病棟区分ごとに、入院算定日の上段の値を lag(入院算定日)に、入院算定日の下段の値を lead(入院算定日)に挿入した。上段・下段がない場合は NULL とした。

レセプト通番	医療機関コード	ID0	病棟区分	入院算定日	lag(入院算定日)	lead(入院算定日)
ABC	1qaz2wsx	hogehoge	1	20130403	NULL	20130404
ABC	1qaz2wsx	hogehoge	1	20130404	20130403	20130405
ABC	1qaz2wsx	hogehoge	1	20130405	20130404	20130406
ABC	1qaz2wsx	hogehoge	1	20130406	20130405	20130429
ABC	1qaz2wsx	hogehoge	1	20130429	20130406	20130430
ABC	1qaz2wsx	hogehoge	1	20130430	20130429	NULL
zyx	1qaz2wsx	hogehoge	1	20130501	NULL	20130522
zyx	1qaz2wsx	hogehoge	1	20130522	20130501	NULL

2. 「入院算定日 - lag(入院算定日)」および「lead(入院算定日) - 入院算定日」がいずれも 1 となった、同一のレセプト通番・医療機関コード・ID0・病棟区分内の入退院日以外のレコードを削除した。その上で、再びレセプト通番・医療機関コード・ID0・病棟区分ごとに、入院算定日の下段の値を lead(入院算定日2)に挿入した。下段がない場合は NULL とした。

レセプト通番	医療機関コード	ID0	病棟区分	入院算定日	lag(入院算定日)	lead(入院算定日)	lead(入院算定日2)
ABC	1qaz2wsx	hogehoge	1	20130403	NULL	20130404	20130406
ABC	1qaz2wsx	hogehoge	1	20130404	20130403	20130405	
ABC	1qaz2wsx	hogehoge	1	20130405	20130404	20130406	
ABC	1qaz2wsx	hogehoge	1	20130406	20130405	20130429	20130429
ABC	1qaz2wsx	hogehoge	1	20130429	20130406	20130430	20130430
ABC	1qaz2wsx	hogehoge	1	20130430	20130429	NULL	NULL
zyx	1qaz2wsx	hogehoge	1	20130501	NULL	20130522	20130522
zyx	1qaz2wsx	hogehoge	1	20130522	20130501	NULL	NULL

3. 同一のレセプト通番・医療機関コード・ID0・病棟区分内の退院日のレコードを削除するため、「入院算定日 - lag(入院算定日)」が 1 で、「lead(入院算定日) - 入院算定日」が 1 以外のレコードを削除した。その上で、入院算定日を入院日に変更し、lead(入院算定日2)を退院日に変更した。加えて、1 日入院である「入院算定日 - lag(入院算定日)」および「lead(入院算定日) - 入院算定日」がいずれも 1 以外のレコードに対し、退院日の列に入院日の値をセットした。

レセプト通番	医療機関コード	ID0	病棟区分	入院日	lag(入院算定日)	lead(入院算定日)	退院日
ABC	1qaz2wsx	hogehoge	1	20130403	NULL	20130404	20130406
ABC	1qaz2wsx	hogehoge	1	20130404	20130403	20130405	20130405
ABC	1qaz2wsx	hogehoge	1	20130405	20130404	20130406	
ABC	1qaz2wsx	hogehoge	1	20130406	20130405	20130429	20130429
ABC	1qaz2wsx	hogehoge	1	20130429	20130406	20130430	20130430
ABC	1qaz2wsx	hogehoge	1	20130430	20130429	NULL	NULL
zyx	1qaz2wsx	hogehoge	1	20130501	NULL	20130522	20130501
zyx	1qaz2wsx	hogehoge	1	20130522	20130501	NULL	20130522

4. lag(入院算定日)およびlead(入院算定日)を削除した。入院日数として退院日 - 入院日 + 1 を挿入し、入院通番としてレセプト通番 & 入院順序をセットした。

レセプト通番	医療機関コード	ID0	病棟区分	入院日	退院日	入院日数	入院通番
ABC	1qaz2wsx	hogehoge	1	20130403	20130406	4	A01
ABC	1qaz2wsx	hogehoge	1	20130404			
ABC	1qaz2wsx	hogehoge	1	20130405			
ABC	1qaz2wsx	hogehoge	1	20130406	20130429		
ABC	1qaz2wsx	hogehoge	1	20130429	20130430	2	A02
ABC	1qaz2wsx	hogehoge	1	20130430	NULL		
zyx	1qaz2wsx	hogehoge	1	20130501	20130501	1	B01
zyx	1qaz2wsx	hogehoge	1	20130522	20130522	1	B02

5. 上記でセットされたテーブルに対し、医療機関コード・ID0・病棟区分・入院日・退院日で自己結合し、入院通番の親子関係テーブルを作成した。

子入院通番	親入院通番
A01	A01
A02	A02
B01	A02
B02	B02

6. 親子関係テーブルを使用して、元テーブルの親入院通番を再帰的に更新、更新データがなくなるまで繰り返した。

レセプト通番	医療機関コード	ID0	病棟区分	入院日	退院日	入院日数	入院通番	親入院通番
ABC	1qaz2wsx	hogehoge	1	20130403	20130406	4	A01	A01
ABC	1qaz2wsx	hogehoge	1	20130429	20130430	2	A02	A02
zyx	1qaz2wsx	hogehoge	1	20130501	20130501	1	B01	A02
zyx	1qaz2wsx	hogehoge	1	20130522	20130522	1	B02	B02

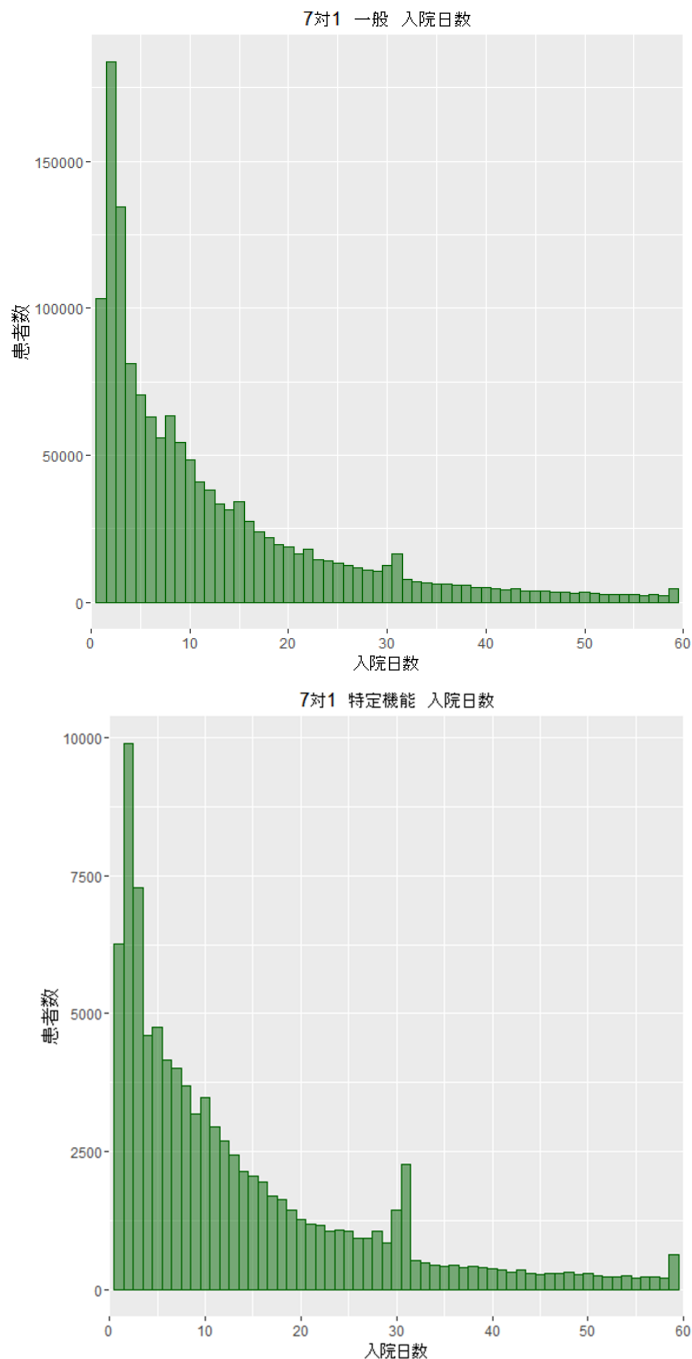
上記処理を全国 1 年分について実行したところ、約 3 日間で完了し、処理後のデータ容量は元のレセプトの 4 倍程度に収まった。

なお、本研究の遂行にあたっては、「人を対象とする医学系研究に関する倫理指針」(平成 26 年 文部科学省、厚生労働省告示)を遵守し、奈良県立医科大学医の倫理審査委員会の承認(平成 27 年 10 月 8 日 受付番号 1123)を得た。

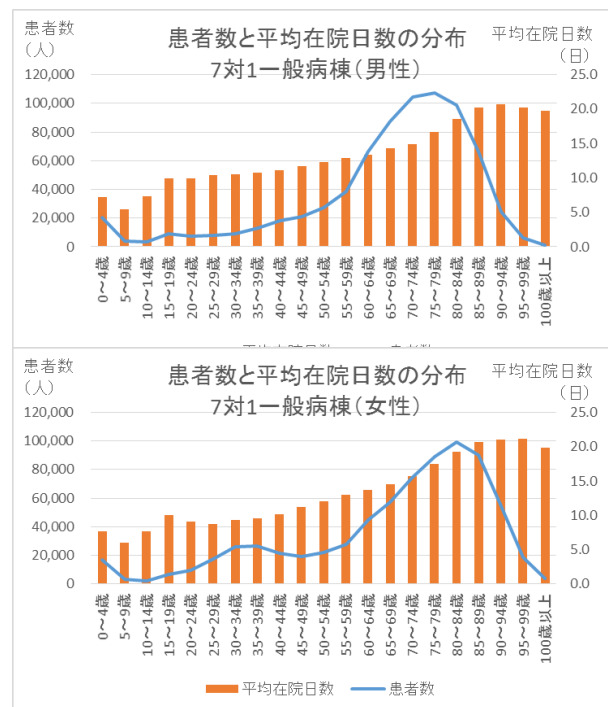
3. 結果

上記方法で行った 1 入院 1 データ化に対し、データの信頼性を調べるため、平均在院日数を求めた。平均在院日数の算出式は、1 入院期間の総和を全入院患者数で割ったもの

である。その結果、男性の精神病棟を除く病棟で 30.4 日、精神病棟で 150.5 日、女性の精神病棟を除く病棟で 35.6 日、精神病棟で 138.9 日となった。7 対 1 入院基本料を算定している病棟の平均在院日数は 15.5 日 (中央値 8 日)、特定機能病棟の平均在院日数は 19.4 日 (中央値 10 日) であった。図 1 にそれぞれの入院日数の分布を示した。



また、性・年齢階級別の患者数及び平均在院日数の分布を図 2 に示した。患者数では、男性は 75 歳～79 歳、女性は 80～84 歳で最大を迎えた。一方平均在院日数では、男性は 90～94 歳に 20.72 日、女性は 95～99 歳に 20.99 日の最大を迎えた。



4. 考察

本研究では従来と異なる方法で 1 入院 1 データ化を実施し、その信頼性評価のために精神病棟などの平均在院日数を算出した。

厚生労働省が実施した平成 27 年度病院報告⁴⁾によると、平均在院日数(平成 27 年度)では、男女合わせて一般病棟で 16.5 日、精神病棟で 274.7 日、全病棟で 29.1 日であった。

値だけで見ると、特に精神病棟で本研究と病院報告に差がある。これは平均在院日数の算出式の違いと本研究の対象期間が原因である。まず平均在院日数の算出式について、本研究では 1 年間の 1 入院期間の日数の総和 ÷ 全入院数であるのに対し、病院報告では 2 × 年間在院患者延べ数 ÷ (年間新入院患者数 + 年間退院患者数) という近似式を用いている。また、本研究の対象期間は 1 年であるため、1 年以上入院していてもその患者の入院期間は 1 年とカウントされた。患者調査によると平成 26 年の精神病棟における 1 年以上の入院数は 18.5 万人と全体の 64.0% であり、うち 5 年以上は 10 万人と全体の 34.6% を占めていた。したがって、精神病棟における平均在院日数が過小評価される結果となった。

次に、本研究では精神病棟を除く全病棟の平均在院日数は 30.4 日であった一方、病院報告では精神病棟を含めた全病棟の平均在院日数は 29.1 日とそれほど差はなかった。加えて、現在の日本の全病床数は約 157 万床、一般病床数は約 90 万床、精神病床数は約 33 万床であり、精神病床は全病床の約 20% を占める。しかしながら、精神病棟のベッド回転率は他病棟に比べて著しく低いため、病院報告の平均在院日数の算出式上、精神病棟を除く全病棟の平均在院日数は本研究の平均在院日数と大きく変わらないと考えられる。

患者調査によると、一般病棟(7 対 1 だけでなく 10 対 1 と 15 対 1 も含む)の入院患者数の最大は男性で 75～79 歳、女性で 80～84 歳であった。また、5～9 歳と 10～14 歳でほぼ同値による最小となっており、いずれも本研究結果と同様の傾向であった。また、女性では 30～39 歳で小さいピークを迎えている。この理由として、正常分娩であっても最後に吸引や鉗子を使った場合は保険診療となり、レセプト上 1 日入院となることが考えられる。

以上のことから、本研究で開発した新たな 1 入院 1 データ化手法は十分信頼に値すると考えられる。今後 3 年分の NDB データを用いて同様の処理を行うが、その際は精神病棟の平均在院日数はより長くなると推定される。

本研究の限界として、上記に示した対象期間が短いことに加えて、本研究の 1 入院の定義上、退院したが同日中または翌日に再入院した場合でも 1 入院として処理される。また、連続して同一医療機関に 1 日入院した場合も 1 入院として処理される問題がある。

実務的な問題点として、入院基本料の取得に当初は診療行為マスタの No.65「入院基本料区分」を参照したが、資料によっては「予備」となっており、内容も「入院基本料加算区分」と矛盾しているため使用は不適切と判断した。代わりに診療行為マスタの No.27「実日数」を参照し、そこから「入院基本料、特定入院料」と「外泊」を参照した。加えて、精神病棟の認知症患者の外泊において、医科診療行為コード 190107590 (外泊 (特定入院料の減額) (精神障害等の患者)) が使用されていたため、このコードも外して集計を行った。

今後、緒論で述べた包括行為と出来高行為をまとめて包括行為レコードに取り込んでいる SI/IY/TO レコードなどを、それぞれ実施日別に分ける処理を行う予定である。それにより各入院期間中の処置、例えば入院中に撮影した CT の回数や、入院中に放射線治療を行った期間、DPC に丸められない製剤を使った期間や回数などを把握できるようになると見込まれる。

5. 結論

データ容量の増大を抑えて、1 入院 1 データ化を高速に処理することが可能となった。今後は包括行為にまとめられている処置などを分割し、入院期間中の医療行為を詳細に把握できるようにする。そして将来的に、入院外レセプトにも適用し、1 患者 1 データ化の実現につなげる予定である。

謝辞

本研究は平成 28、29 年度厚生労働科学研究費補助金 (地域医療基盤開発推進研究事業) 「病床機能の分化・連携や病床の効率的利用等のために必要となる実施可能な施策に関する研究」、平成 28 年度国立研究開発法人日本医療研究開発機構 (AMED) 地域横断的 ICT 活用医療推進研究事業「レセプト等の大規模電子診療情報を活用した薬剤疫学研究を含む医療パフォーマンス評価に関する研究」の一環として実施したものである。

参考文献

- 1) 今村 知明. 地域横断的な医療介護情報の ICT 化により、世界最先端の臨床研究基盤等の構築を加速するための研究. AMED [http://www.amed.go.jp/content/files/jp/houkoku_h27/0501053/2015lk_chiikiICT_001.pdf (cited 2017-Aug-31)].
- 2) 特許公報. 電子レセプトデータ変換プログラムおよび電子レセプトデータ変換システム. 特開 2011-118538.
- 3) 野田龍也、久保慎一郎、明神大也、西岡祐一、東野恒之、松居宏樹、他. レセプト情報・特定健診等情報データベース (NDB) における患者突合 (名寄せ) 手法の改良と検証. 厚生 の 指 標 (in press)
- 4) 厚生労働省. 平成 27 年 (2015) 医療施設 (動態) 調査・病院報告の概況. 2015. [http://www.mhlw.go.jp/toukei/saikin/hw/iryosd/15/dl/gaikyo.pdf (cited 2017-Aug-31)].
- 5) 厚生労働省. 平成 26 年患者調査. [http://www.mhlw.go.jp/file/05-Shingikai-12201000-Shakaiengokyokushougaihoukenfukushibu-Kikakuka/0000108755_12.pdf

(cited 2017-Sep-3)].