

뇌졸중 환자 분석 2단계

• CT, MRI 촬영

- ① 진료내역 테이블에서 CT, MRI 촬영 건을 추출
 - 항목코드(CLAUSE_CD)=특수장비('S') and
 - 분류유형 코드(DIV_TYPE_CD)=수가('1') and
 - 분류코드(DIV_CD)= 앞 두 자리가 CT, MRI('HA','HE','HF')

- ② 명세서 일반테이블의 뇌졸중 전체 건에 대해 ①의 결과를 merge

- ③ CT, MRI를 촬영하지 않은 건은 삭제하고 환자기준으로 중복삭제

※ ②에서 환자중복삭제 후 merge는 안됨

뇌졸중 환자 분석 2단계

- 표본환자 중 CT, MRI 검사를 받은 환자

성별	구분	표본환자수	CT, MRI	백분율
남성	0-9세	46	40	87.0%
	10-19세	49	40	81.6%
	20-29세	126	85	67.5%
	30-39세	387	279	72.1%
	40-49세	1,368	924	67.5%
	50-59세	2,681	1,765	65.8%
	60-69세	3,496	2,264	64.8%
	70-79세	3,781	2,359	62.4%
	80-89세	1,558	881	56.5%
	90세이상	120	62	51.7%
여성	0-9세	43	38	88.4%
	10-19세	52	39	75.0%
	20-29세	90	63	70.0%
	30-39세	207	134	64.7%
	40-49세	871	610	70.0%
	50-59세	1,686	1,111	65.9%
	60-69세	2,885	1,820	63.1%
	70-79세	5,292	3,000	56.7%
	80-89세	3,275	1,601	48.9%
	90세이상	393	181	46.1%
Total		28,406	17,296	60.9%

전자간증 위험인자 분석

- ◆ 연구제목 : 환자표본자료를 활용한 산모의 전자간증 위험인자 분석
- ◆ 연구대상 : 2009년도 전체 산모 438,124명 중 약 13%에 해당하는 분만 산모 56,474명
- ◆ 전자간증의 위험인자 : 나이, 분만력, 다태아임신 및 분망 방법을 확인 분석

- ◆ 상병정의: 임신중독 ('O11','O13','O14','O15')

- ◆ 데이터 핸들링
 - ✓ nps_40 테이블에서 상병코드(SICK_SYM)가 ('O11','O13','O14','O15')인 경우 추출
 - ✓ nps_30 테이블에서 처치코드(div_cd)가 분만 ('R4351','R4353','R4356','R4358','R3131','R3133','R3136','R3138','R4361','R4362','R4517','R4519','R4518','R4520','R4514','R4516','R3141','R3143','R3146','R3148','R4522','R4380','R4507','R4508','R4509','R4510','R5001','R5002')인 경우 추출

 - ✓ 추출한 30테이블, 40테이블과 20테이블을 각각 keycode로 병합(기준은 20테이블)

임신중독증 위험인자 분석

20_30병합 테이블을 분만 종류별로 더미코딩(1 or 0)

자연분만	초산부(R4351)
	경산부(R4356)
	초산 다산부(R4353)
	경산 다산부(R4358)
제왕절개	초산부(R4517)
	초산 다산부(R4519)
	경산부(R4518)
	반복 1태아(R4514)
	경산 다산부(R4520)
	반복 다산부(R4516)
둔위분만	초산부(R4361)
	경산부(R4362)
유도분만	초산부(R3131)
	초산 다산부(R3133)
	경산부(R3136)
	경산 다산부(R3138)
흡입분만	초산부(R3141)
	초산 다산부(R3143)
	경산부(R3146)
	경산 다산부(R3148)
질식분만	(R4380)
질식시도후 제왕절개	(R4522)

임신중독증 위험인자 분석

- 20_40병합테이블을 다시 임신중독파일과 임신성당뇨파일로 나눈 후 더미처리 후 환자기준으로 중복삭제
- 20_30병합파일에 임신중독파일과 임신성당뇨파일을 각각 덮어 씌움(기준은 20_30병합파일)>>20_30_40병합파일 생성
- 20_30_40병합파일을 환자기준 중복삭제할 필요 없음(출산처치 코드이므로)
- 종속변수 : 임신중독증 여부
- 독립변수 : 35세 구분, 다태아, 제왕절개, 둘째아이분만

전자간증 위험인자 분석

◆모집단 추정 결과

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald	
		Confidence Limits	
제왕절개	2.744	2.647	2.845
다태아여부	1.986	1.824	2.162
둘째출산	0.494	0.476	0.514
35세 이상	1.678	1.607	1.751

◆표본자료 추정결과

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald	
		Confidence Limits	
제왕절개	2.261	2.166	2.359
다태아여부	1.899	1.71	2.11
둘째출산	0.529	0.506	0.553
35세 이상	1.754	1.667	1.845

표본자료 는 SamplingWeight 샘플링 가중치 부여 후 모집단 추정

전자간증 위험인자 분석

◆ 결과 :

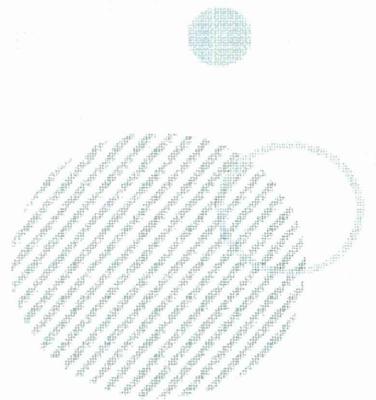
- ✓ 표본 자료를 바탕으로 한 분석에서는 다태아 임신 (OR 1.90, 95% CI 1.71, 2.11) 및 35세 이상고령임신 (OR 1.75, 95% CI 1.67, 1.85) 인 경우 전자간증의 위험이 증가하였으나 경산부인 경우 전간간증 위험이 감소 (OR 0.53, 95% CI 0.51, 0.55)
- ✓ 모집단을 바탕으로 한 분석 결과에서도 이와 일치하는 소견으로 (다태아 임신, OR 1.99, 95% CI 1.82, 2.16),(35세 이상 고령임신, OR 1.68 95%, CI 1.61, 1.75),(경산부, OR 0.49, 95% CI 0.48, 0.51).

◆ 결론 :

- ✓ 본 연구에서는 표본 자료 및 모집단 자료를 통해 전자간증의 위험인자를 분석하였으며 이를 통해 고령임신 및 다태아임신인 경우 그 위험이 증가하는 것을 확인
- ✓ 두 자료간의 결과가 일치하는 소견을 보여 표본집단의 모집단에 대한 예측 자료로서 의미가 있는 것을 확인

09

결론 및 제언



건강보험자료의 유용성

- 대표성

- 전국민을 대표하는 자료로서 일반화 용이
- 결과는 주의 깊게 해석

- 자료의 완결성

- 행위별 수가제 하의 청구자료 구축으로 세부 의료이용 내역 확인 가능
- 약물사용에 대한 자세한 정보를 가지고 있음
- 드물게 발생하는 사건(희귀질병, 합병증, 약물 부작용, 희소하게 사용되는 진료행위)에 대한 연구 가능

- Real world

- 제한적 실험적 환경이 아닌 실제 보건의료환경 반영
- 현황 및 추세를 보여줌
- 이미 구축된 자료를 활용하므로 비용과 시간 단축

제한점 -1

- 정보 오분류

- 급여가 인정된 의료이용만 포함됨
 - 비급여 진료, 처방없이 약구입(뇌졸중/아스피린), 미용 등 미포함
- 의료급여환자, 보건소진료환자, DRG대상환자의 경우 명세서일반사항은 있으나 구체적 진료내역이 없음

- 진단명의 정확성 문제

- 청구자료 주상병과 의무기록의 일치도: 약 70% (2003년)
 - 입원 > 외래, 위중한 환자 > 다빈도 경증 질환
 - 종합병원 > 병원 > 의원
- 진단명, 시술에서 개인차, 관습적 요인을 완전히 배제하기 어려움
- 자료의 특성, 환자의 의료이용행태와 질병의 고유특성, 의사의 진료과정과 임상환경, 병원의 전산망과 청구과정, 건강보험급여제도 등을 충분히 파악해야 올바른 해석 가능

제한점 -2

- 표본자료는 적정수준이상의 모집단의 빈도를 확보해야 유의성을 보장하므로 표본자료의 특성상 발생빈도가 높은 상병 일수록 설명력이 커지며 희귀빈도의 상병일수록 설명력은 감소

- 환자 상태에 대한 정보 부족

- 임상검사 및 신체기능에 대한 측정값, 삶의 질, 만족도
- 몸무게, 키, 거주지, 생활습관, 질병력, 가족력
- 각종 lab data

- 주요 자료

- 병원 의무기록
- 국민건강영양조사자료
- 건강검진자료

건강보험자료의 정확한 사용

▶ 개인정보 보호

- 본래의 연구목적에 맞게 사용
- 자료 유출에 주의

▶ 신중한 가설 설정과 보건 의료환경 검토

- 환자의 의료이용행태와 질병고유의 특성
- 의사의 진료과정과 임상환경
- 병원의 전산망과 청구과정
- 심사평가 관련 제도 및 건강보험 급여제도

▶ 기술적인 문제

- 자료처리 기법 및 통계적 분석방법론
- 하드웨어, 용량 문제 고려

II. 研究成果の刊行に関する一覧表

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ

雑誌等

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
猪飼宏, 大坪徹也, 今中雄一	全国網羅的なレセプト情報の有効活用に資するデータベース基盤の課題と設計	日本医療・病院管理学会誌	第50巻 Supplement	235	2013
加藤源太, 平野景子, 赤羽根直樹.	レセプト情報・特定健診情報等情報データベースの利活用について—これまでの経緯を踏まえて—	統計	2014年10月号.	pp.8-13	2014
今中雄一	脳梗塞t-PA治療の経済評価と均てん化施策	脳と循環	Sep Vol.19, No.3.	pp.49-52.	2014
今中雄一	混合診療と医療改革—新制度 広くデータ評価を.	日本経済新聞	2014年7月25日	経済教室	2014

Ⅲ. 研究成果の刊行物・別刷

2014年10月1日発行(毎月1回1日発行)統計 第65巻 第10号

統計

10

2014

October

特集：データヘルス事業とレセプトNDBの有効利用



一般財団法人 日本統計協会

特集 データヘルス事業とレセプトNDBの有効利用

- データヘルスの推進について 安藤 公一 **2**
- レセプト情報・特定健診等情報データベースの利活用について
—これまでの経緯を踏まえて— 加藤 源太・平野 景子・赤羽根直樹 **8**
- レセプト情報・特定健診等情報サンプリング
データセット分析とその意義 伊藤 弘人・奥村 泰之 **14**
- DPCとNDB及び患者調査データを活用した
地域医療の計画化に関する一考察 松田 晋哉 **21**
- 薬剤疫学とレセプト情報 関 真美・津本 周作・岡本 基・椿 広計 **28**

論考

- NDBの活用の現状、課題、提言 野口 晴子 **34**

統計を深く知る

- 厚生労働省縦断調査開始の経緯と現状、今後の課題
..... 厚生労働省大臣官房統計情報部人口動態・保健社会統計課世帯統計室 **41**

統計ウォッチング

- 多面的尺度で考える空洞化 冨浦 英一 **46**
- マクロで見たリソースの世代間移転の変化と高齢者での需要増加 小川 直宏 **50**

統計をとりまく動き

- 岐阜県が取り組む統計教育について 岐阜県環境生活部統計課 **54**

統計討評

- 誤差率について 島村 史郎 **56**

統計でみる都道府県

- おもてなし、やまなし 山梨県企画県民部統計調査課 **58**

連載

- 歴史の中の統計学(7) —統計学以前の統計学(1) 政治算術— 竹内 啓 **64**

- つれづれなる記 永山 貞則 **71**

英国王立統計学会 Significance から

- チャンピオンズリーグ組合せ抽選の偶然 横山 暁 **72**

書評

- 松田 茂樹 著 少子化論—なぜまだ結婚、出産しやすい国にならないのか 阿藤 誠 **78**

- スタチスチック **80**

レセプト情報・特定健診等 情報データベースの利活用について

— これまでの経緯を踏まえて —

加藤 源太・平野 景子・赤羽根 直樹

はじめに

平成18(2006)年の医療制度改革において、医療費を増加させるような構造的要因に着目し、中長期的に医療費適正化を進める仕組みを構築する目的で、医療費適正化計画という枠組みが導入されることとなった。その際、医療費適正化計画策定の際の調査分析に資するデータベースとして整備されることとなったのが、レセプト情報・特定健診等情報データベースである。このデータベースは、高齢者の医療の確保に関する法律を根拠として、保険者の特定健診および特定保健指導情報、ならびにレセプト情報を厚生労働省保険局において格納・管理するデータベースである。

このデータベースであるが、医療費適正化計画策定に資する目的以外にも、医療サービスの質の向上等を目指した正確なエビデンスに基づく施策の推進や、これら施策に有益な分析・研究、また学術研究の発展に資する目的で行う分析・研究への利活用について、「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」にて、一定の審査を行ったうえで認めるべきではないかという提言がなされた。加えて、平成22(2010)年6月22日閣議決定「新たな情報通信技術戦略工程表」においても、当データベースの活用に関する方針が示されることとなった。こうした背景を踏まえ、同年に「レセプト情報等の提供に関する有識者会議」(以下「有識者会議」とする)が設立され、本来目的以外の目的に対しても、有

識者会議の審査を経た上で平成23(2011)年度よりデータの利用が認められることとなり、現在に至っている。

本稿では、当データベースの利活用について、これまでの経緯、ならびに今後の方向性に関する事項を中心に概要を紹介する。また、実際に利活用する場合に留意すべき点などについても、言及することとする。

データベースに格納されている情報について

このデータベースは、大きく分けてレセプト情報、ならびに特定健診・特定保健指導情報によって構成されており、両者において、データの性質や構造が大きく異なっている。

まずレセプト情報についてであるが、レセプトとは、保険診療を行った医療機関が、診療報酬点数表に基づいて診療報酬(医療費)を保険者に請求するために、患者一人について毎月月末に作成する診療報酬明細書のことである。平成18(2006)年以降、このレセプトについてオンライン・電子媒体での請求を推進しているところであり、平成21(2009)年度診療分より、電子化されたレセプト情報がこのデータベースに格納されている。したがって、全ての保険診療に関するレセプト情報がデータベースに格納されているわけではなく、医療機関でのレセプト請求の電子化の状況によって、データベースに格納されるレセプト情報の割合は影響を受けることとなる。平成26(2014)年

8月現在、各医療機関におけるレセプト請求の電子化率はレセプトの種類によって差がみられるものの、社会保険診療報酬支払基金の調査によれば、電子レセプト請求の普及状況は全体で90%以上となっている（図1・図2）。レセプトは保険診療に関する診療報酬明細書であり、主な記載項目として、たとえば氏名や生年月日・性別などといった患者情報、医療機関や保険者、被保険者等の情

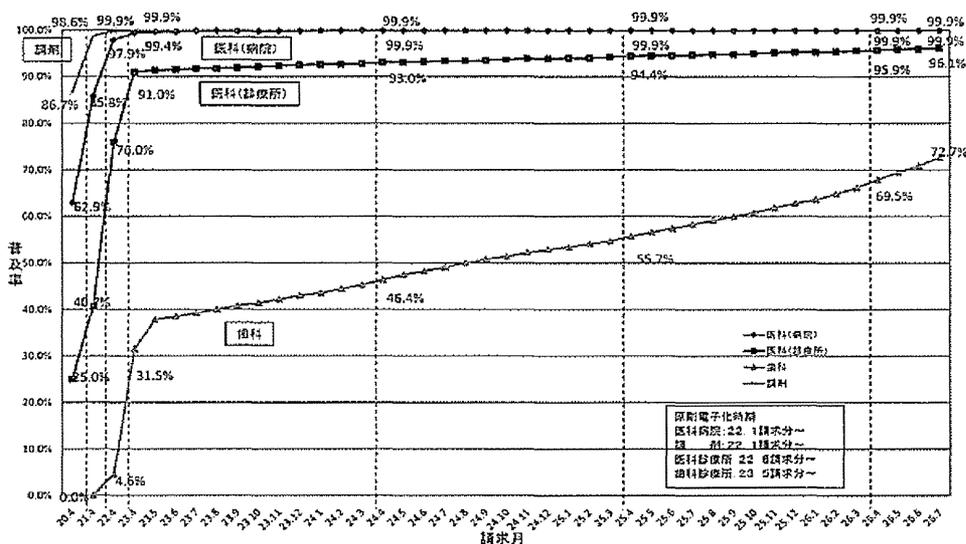
報、そして保険診療に関連する傷病名や治療内容、投薬等の情報が含まれている。一方、これらのレセプト情報をデータベースに格納する際には、患者氏名や生年月日の「日」、保険医療機関の所在地及び名称、カルテ番号等、被保険者証（手帳）等の記号・番号等は削除されている。その際、生年月日および性別情報に加えて保険者番号や被保険者証（手帳）等の記号・番号といった情報、あ

図1 電子レセプト請求書普及状況（件数ベース）（平成26年7月請求分）

		普及率	
		50%	100%
総計	電子レセプト	94.6%	
	紙レセプト		5.4%
		オンライン	電子媒体
		70.9%	23.7%
医科	40床以上		99.9%
	40床未満	オンライン	99.9%
	病院計 1,029万件		99.9%
	診療所 3,276万件		96.1%
	医師計 4,305万件		97.0%
歯科	1,070万件	72.7%	紙
調剤	2,514万件		99.9%

資料：社会保険診療報酬支払基金調べ

図2 医療機関のレセプト電子化の推移（レセプト件数ベース）



資料：社会保険診療報酬支払基金調べ

るいは氏名情報から二種類のハッシュ値を作成し、そのハッシュ値が個人IDの代わりとして個々のレセプト情報に付与され、データベースに格納されている。このハッシュ値を用いることで、個々のレセプト情報を同一人物の情報として特定することができ、このハッシュ値で紐付けることにより、複数の医療機関を受診する患者の受診動向や、継続的に治療を受ける患者の受診動向を評価することが可能となる。後述する特定健診・特定保健指導情報についても同様の処理が施されている。

特定健診情報は、平成20（2008）年度より、医療保険者（国民健康保険・被用者保険）に対して40歳以上75歳未満の被保険者・被扶養者を対象とする内臓脂肪型肥満に着目した特定健診の実施が義務付けられることとなったが、その際の受診情報が元となっている。同じく、特定保健指導情報は、特定健診の受診者のうちで一定の基準に該当する者に対して特定保健指導が実施されることとなっているが、その際の保健指導に関する情報が元となっている。

特定健診・特定保健指導情報は、個々の受診者ごとに実施された特定健診・特定保健指導に関する情報が含まれている。主な項目として、たとえば氏名や生年月日・性別などといった受診者情報、特定健診・特定保健指導の実施機関や保険者、被保険者等の情報、そして特定健診については問診

結果や生活習慣病に関連した測定項目の結果、あるいは特定保健指導については保健指導レベルや支援形態、などといった情報が含まれている。一方、これらの情報をデータベースに格納する際には、受診者氏名、特定健診・特定保健指導の実施機関の郵便番号や所在地、名称、電話番号、医師の氏名、被保険者証（手帳）等記号・番号等は削除されている。そして、レセプト情報と同様、個人IDの代わりとして個々のレセプト情報に二種類のハッシュ値が付与され、データベースに格納されている。

特定健診・特定保健指導情報については、レセプト情報と異なり、全ての受診者情報がデータベースに格納されている。年度により受診率に変動がみられるが、平成24（2012）年度受診分の特定健診情報件数は約2,500万件、特定保健指導情報件数は約84万件となっている（表1）。

施策や研究にあたっての利用一本来目的以外での利用について

「はじめに」でも述べたとおり、医療費適正化計画策定に資する目的以外に、正確なエビデンスに基づく施策の推進や、これら施策に有益な分析・研究、また学術研究の発展に資する目的で行う研究に対しても、平成23（2011）年よりデータ提供が行われることとなった。

データ提供にあたっては、平成23（2010）年10月に有識者会議が設置され、データ提供に関する各種手続き等が規定された「レセプト情報・特定健診等情報の提供に関するガイドライン」（以下「ガイドライン」とする）が定められた。このガイドラインに則

表1 平成26（2014）年8月時点でのデータベースに格納されたデータ件数

	レセプト情報件数	特定健診情報件数	特定保健指導情報件数
平成20(2008)年度		約2,000万件	約39万件
平成21(2009)年度	約12億1,700万件	約2,200万件	約58万件
平成22(2010)年度	約15億1,100万件	約2,300万件	約61万件
平成23(2011)年度	約16億1,900万件	約2,400万件	約72万件
平成24(2012)年度	約16億8,100万件	約2,500万件	約84万件
平成25(2013)年度	約17億2,800万件		
平成26(2014)年度	約2億9,600万件		
合計	約80億5,200万件	約1億1,400万件	約315万件