

健康保険審査評価院の患者標本資料(HIRA-NPS)紹介

김 록영 (キムロクヤン) 主任研究員
健康保険審査評価院統計管理部

1. 前書き

1960～70年代の韓国は標本資料に関する具体的な体系も確定できない状態であったが、1990年代は統計庁で保管している調査資料が一部提供され、科学的な標本設計が行われるようになった。保健分野においては2000年代でも保健調査資料は10種程度であったが、2010年になって40を超える保健調査資料やパネル調査資料が保有されるようになった。



韓国は現在、全国民の98%が国民健康保険に加入しており、健康保険の請求資料は国の保健医療を代表する資料になっている。しかし、今までの保健分野の資料は主に実体調査を基礎にしたもので、健康保険の請求資料に対する標本資料が提供される状態ではなかった。

台湾は1995年3月1日に韓国と類似の制度で単一の国家医療保険を始め2007年現在、全人口の98.4%が登録されており、国民医療保険DB (National Health Insurance Research Database)には医療償還のための保険者登録資料と健康保険請求資料が含まれている。このDBは健康保険局 (Bureau of National Health Insurance) 傘下の国立保健研究所 (National Health Research Institutes) の管理の下に構築され、研究用にも提供されている。これには請求用の月単位標本資料とともに、患者基準で実施される1年単位標本資料、5年単位のパネル資料まで多様な方式で健康保険請求資料に関する標本資料が提供されている。

月単位の標本資料は主に季節に敏感な流行性疾病に適切に対応するための資料として活用され、1年単位あるいは5年単位のパネル資料は主に研究用として活用されている。

米国の Agency for Healthcare Research and Quality (AHRQ) は、連邦政府に属する研究機関で保健分野に関連する研究を支援している。AHRQは37の州政府及び地域社会、保健医療産業界などからのデータを収集して医療DBを構築している。

AHRQで管理する調査プログラムのひとつである Healthcare Cost and Utilization Project (HCUP) は、米国で最も大きいDBを構築しており、HCUPの提供資料の中で最も包括的な全国入院患者サンプルデータ (National Inpatient Sample) は、リハビリ医療機関を除いたアメリカ病院協会 (American Hospital Association) に属するすべての医療コミュニティを包含している。

NISはコミュニティに加入した37州の約3,900の医療機関から収集したデータが基盤で、加入医療機関から毎年約20% (800～1,100機関) を標本抽出した医療機関の全入院資料 (5百万～8百万の入院件数) を含んでいる。

これに対し韓国の場合、健康保険審査評価院と国民健康保険公団で資料処理室を運営しており、健康保険請求資料を外部研究者が直接加工して結果を算出しているが、接近性と便宜性の面において限界が存在する。

また診療情報資料は年間約10億件以上に達するので、その容量が膨大で使用者の保存容量、処理速度など収容能力を超え、適切な資料確保を不可能にしている。そのため、多様な需要層に対する接近性、便宜性、即時性の確保のための代案の一つとして我が国の健康保険請求資料に対する標本資料を開発した。

米国は退院データの性格上、非給与分を含むが特定の医療機関への患者の入院から退院までを1単位とする資料なので、その患者の再入院や他の医療機関への入院に関するデータを患者ごとに区分することができない。これに対して韓国や台湾では請求資料には非給与分は含まれないが、患者単位の資料なので再入院や他医療機関の受診も含むことが可能である。

国家別比較	韓国 (HNPS)	米国 (NIS)	台湾 (NHRD)
抽出単位	患者	病院	患者
提供単位	患者単位	機関別退院患者単位	患者単位
層化変数	性、年齢区間	病院特徴 地理学的位置	単純無作為抽出
提供対象	すべての研究者	すべての研究者	国家研究機関及び研究者 (一般の学習用資料利用)
標本規模	入院患者13% (約70万名) 外来患者1% (約40万名)	約700万件程度の機関 単位 (入院資料)	健康保険登録者の100万名

表1. 国家別標本資料比較

2. 患者標本資料の開発背景及び目的

健康保険の請求資料は、制限的な実験環境での使用を対象にしたものではなく、実際の保健医療環境を反映するデータである。そのため比較的一般化が容易で構築された資料を活用することで研究に要す時間や費用を縮めることが可能である。韓国は世界的にも例が少なく、国家が運営する単一の健康保険資料を保有しており、この資源は保健医療分野の国家政策の樹立及び国民の健康増進に係わる研究の基礎資料として活用できるものであり、現在も多様な分野で健康保険請求資料を活用した研究需要が急増している状態である。

そのために健康保険審査評価院では健康保険請求資料に対する多様な分野の需要に応じる方案の一環として、代表的な我が国の健康保険資料の標本資料を開発することになった。標本資料の開発及び提供の目的は健康保険資料に対する理解及び活用度を高めて保健医療研究を活性化させるためのものである。

本研究は健康保険請求資料に関する接近性を拡大するための方案として効率的な標本設計方法論に関するものである。標本資料の公開によって国家社会的な便益を向上させ、持続的な還流過程を通じて多様化されて改善された資料が開発されることで期待する。

3. 標本対象及び抽出方法

ガ. 患者標本資料の概要

健康保険審査評価院で開発した患者標本資料の概要を次に示す。

- 2009年の1年間に医療を利用したすべての患者(約4600万名)を性別、5歳単位の年齢区別の患者単位の層化系統抽出
 - 標本抽出患者の1年間のすべての診療内訳と処方内訳を含む。
 - 標本基準変数は 1年間患者当たり総診療金額(最大分散を持つ変数)
 - 入院患者抽出比率13%(母集団の1年間入院患者 1人当り平均診療金額標準偏差の0.5%に相当する金額を標本の許容誤差範囲とした。)
 - 外来患者抽出比率1%(入院と等しい誤差範囲で必要な標本人員は0.05%であるが、それを1%に拡大して提供。)
- 該当年度の1月から翌年6月までの審査決定資料を基準とした該当年度1年分の診療内訳。
- 容量 : 33G (圧縮後3.44G) , 40,989,560行(件)、DVD1枚

ナ. 母集団(健康保険請求資料)の特徴

患者標本資料の対象母集団としての健康保険請求資料とは医療機関における患者の診療費用のうち、国民健康保険が負担する部分に対して支給依頼をするために健康保険審査評価院に請求する資料である。

国民健康保険公団は国民を対象として国民健康保険財政に対する取り立て、管理及び支給業務を担当し、健康保険審査評価院は医療機関が請求する請求資料に対する診療費審査結果を健

康保険管理公団に知らせて支給を要求している。

我が国の1年間の健康保険請求患者数は2009年基準で45,969,893人、年央人口数49,773,145人の92.4%である。

保険請求は診療行為、薬品、治療材料の三種類の項目で構成されている。明細書は入院の場合は入院から退院を単位とし、月単位に分離請求(入院が持続する場合は明細書が分離して請求される)、外来の場合は月間診療内訳統合請求(ただし医院級では訪問日ごとに仕分けして請求)、薬局の場合は日別に仕分けして(2005年以後)請求される。

㉔. 健康保険請求資料の標本抽出単位の設定

健康保険資料を活用した標本抽出方法の理解のために数種の図を紹介する。すでに述べたように医療機関では月別あるいは日別で健康保険審査評価院に患者当たり1件の健康保険療養給与を請求する。

標本資料はすべてのエピソード及び疾病定義に対する基準を満足させるために1年間の健康保険請求患者を1単位として抽出した。患者を1単位として抽出することによって、それぞれの疾病特性及び定義によって疾病のエピソードを構築して研究に活用することができる。

㉕. 標本抽出方法

標本抽出は層の分割を可能とするに十分な情報が存在するので単純無作為方式よりは層化抽出方式を優先考慮した。病院請求資料の場合患者の年齢及び性別区分の階分割を可能とするに十分な情報になるので層化抽出方式を使用した。

層化標本抽出は単純無作為標本抽出に比べ標本数でも全体母集団の特性をよく代表することができる標本抽出方法なので標本資料の効率性を高めるために層化抽出方式を使用した。層化変数は性別、年齢5歳単位の16区間で全32の層を持つ層化抽出で、最も一般的に使われる層化系統比例確率抽出方式を使用した。

4. 結果及び妥当度の検証

㉖. 標本抽出結果

2009年の1年間に健康保険審査評価院で療養給与が請求された全患者数は約4,600万人で入院経験がある患者(全入院患者)は約500万名、入院経験がなく外来だけの来院患者(全外来患者)は約4,000万人であった。全入院患者の13%、全外来患者の1%を標本抽出して、患者標本資料に含まれた標本患者数は約110万名であった。

表2 標本資料の概要

全体入院患者	5,472,670名	標本入院患者	711,457名
全体外来患者	40,497,223名	標本外来患者	404,583名
母集団患者数	45,969,893名	標本資料患者数	1,116,040名

表3 標本資料の容量

区分	容量
明細書一般内訳	27,320,505列(3.82G)
診療内訳	313,011,694列(21.13G)
傷病内訳	68,807,094列(1.56G)
処方せん詳細内容	65,477,122列(3.93G)
療養機関現況	80,418機関(6.84MB)
全体	34Gbyte(DVD_CD 1枚)

4. 妥当性の検証

1) 統計的妥当性の評価結果

標本検定方法は分散分析(一員일원分散分析)を使用して95%信頼水準で二つの仮説をすべて採択して母集団と標本間に差がないことを確認した。抽出された標本の効率性を評価するために母集団と標本集団の分散の相対比率を使用した。

患者単位の母集団に対する分散の相対比率は98.02%である。健康保険請求資料を患者単位にした場合、最大分散を持つ連続変数は1年間の患者1人当り診療費である。

母集団と患者標本資料の患者当たりの診療費を月別で比較した結果、(図2)の標本資料と母集団の月別患者当たり平均診療費推移がほとんど一致し、分散の相対比率は0.98と高い一致率を示した。

入院患者の多頻度疾患を主傷病基準で母集団と標本資料の発病件数を比較した相対比率は、入院患者抽出割合である13%に近似した値を示した。

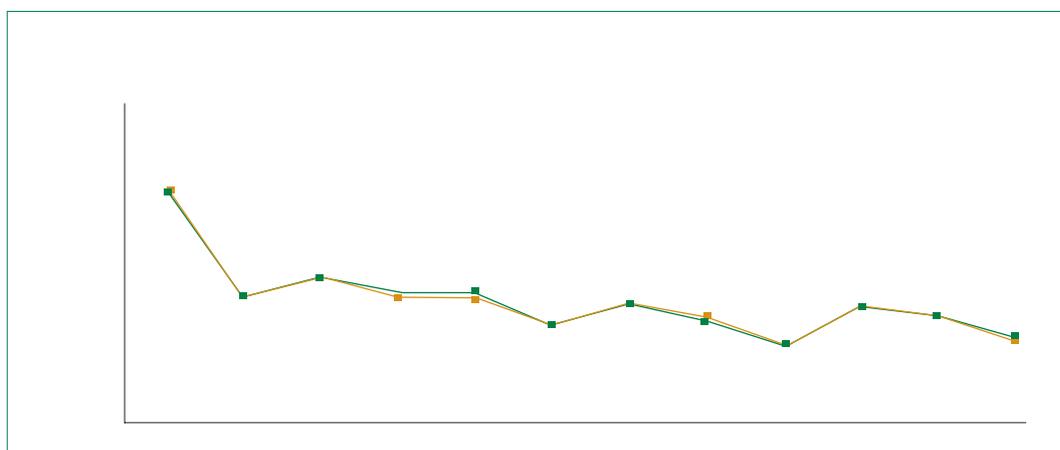


図 2. 月別 1人当たり平均請求費用の母集団と標本集団比較

95%信頼水準で母集団推定誤差区間を算出した結果、慢性腎不全症(N18)の全患者は1,159,366,039人の推定値段を持つ母集団の慢性腎不全症患者は110,354人で全患者と各年齢区間はすべて信頼区間内に真値が存在することが確認できた。

2) 関連学会¹⁾による資料の妥当性検討

健康保険審査評価院は患者標本資料の妥当性評価のために、保健医療関連5学会とMOUを結んでいるが、MOU学会会員による標本資料活用研究の主要結果を見れば「我が国糖尿病有病率推定及びDPP-4抑制剤使用様相評価²⁾」で標本資料を利用した糖尿病有病率推定結果が母集団の分析結果と一致し、血糖降下剤の各薬効群別処方率推定結果が母集団分析結果と一致することが確認できた。

また外来患者の各薬効群別処方率はすべて、推定値の95%の信頼区間内に真値が存在した。

「視力損失と失明による社会的疾病負担費用推計³⁾」の研究では、標本資料と母集団のすべてで女性が男性に比べてすべての主要眼疾患(白内障、緑内障、黄斑変性、糖尿網膜弁中、網膜静脈閉鎖)で医療利用患者の割合が高いことが明らかになり、白内障、緑内障、黄斑変性の年齢別推移も母集団と標本資料が類似の様相であることが明らかにされた。

95%信頼水準で母集団推定誤差区間を算出した結果、慢性腎不全症(N18)の全患者は1,159,366,039人の推定値段を持つ母集団の慢性腎不全症患者は110,354人で全患者と各年齢区間はすべて信頼区間内に真値が存在することが確認できた。95%信頼水準で母集団推定誤差区間を算出した結果、慢性腎不全症(N18)の全患者は1,159,366,039人の推定値段を持つ母集団の慢性腎不全症患者は110,354人で全患者と各年齢区間はすべて

信頼区間内に真値が存在することが確認できた。95%信頼水準で母集団推定誤差区間を算出した結果、慢性腎不全症(N18)の全患者は1,159,366,039人の推定値段を持つ母集団の慢性腎不全症患者は110,354人で全患者と各年齢区間はすべて信頼区間内に真値が存在することが確認できた。

表4 入院患者の30種の多頻度疾病での標本資料と母集団の相対比率

多頻度疾病順位	疾病コード	母集団頻度	標本頻度	百分率
1	O800	188,840	24,484	12.97
2	J189	139,603	18,025	12.91
3	A09	127,301	16,373	12.86
4	I841	126,508	16,510	13.05
5	H259	79,074	10,243	12.95
6	O820	78,332	10,298	13.15
7	K359	74,155	9,581	12.92
8	M511	69,381	9,052	13.05
9	I639	59,495	7,847	13.19
10	I10	53,087	6,860	12.92
11	S335	52,243	6,681	12.79
12	H251	45,861	5,903	12.87
13	O821	43,856	5,739	13.09
14	N10	42,737	5,519	12.91
15	H258	42,503	5,645	13.28
16	J459	41,512	5,431	13.08
17	N393	39,918	5,257	13.17
18	J209	38,260	4,958	12.96
19	J180	36,630	4,675	12.76
20	J039	33,588	4,257	12.67
21	D259	32,834	4,311	13.13
22	M4806	32,383	4,122	12.73
23	I209	30,714	4,018	13.08
24	M512	29,983	3,866	12.89
25	I200	29,760	3,858	12.96
26	C169	29,430	3,819	12.98
27	J157	28,704	3,714	12.94
28	S3200	27,184	3,479	12.80
29	M170	26,832	3,497	13.03
30	E119	26,500	3,469	13.09

表5. 標本資料を利用した慢性腎不全症(N18)患者の母集団推定

年齢区間 (歳)	95%信頼下限	95%信頼上限	母集団患者頻度
1~4歳	11	81	48
5~9歳	16	92	88
10~14歳	78	199	192
15~19歳	127	749	430
20~24歳	123	1,053	625
25~29歳	1,034	2,574	1,635
30~34歳	1,820	3,715	2,802
35~39歳	3,657	6,275	4,918
40~44歳	5,454	8,600	6,696
45~49歳	8,888	12,763	9,897
50~54歳	9,986	13,928	12,298
55~59歳	10,797	14,849	11,796
60~64歳	11,266	15,349	12,808
65~69歳	13,960	18,386	14,940
70~74歳	12,290	16,323	13,858
75歳以上	15,854	20,330	17,323
Total	109,263	121,252	110,354

5. 結論及び提

言

本研究の結果で母集団である2009年健保請求資料に対する代表性と個人情報保護が一定水準を満足した標本資料を算出することができた。

今後の標本資料は持続的な開発及び補完を通じて資料の提供領域を拡大して行く計画である。

患者標本資料活用時の注意事項は、給与が認められた医療利用内訳は含むが、非給与の内訳または処方せんなしに購入できるアスピリンなどの一般医薬品の情報は標本資料で確認できないことである。

診断名の正確性は外来よりは入院患者が、多頻度軽度疾患よりは危重な患者で高くなり、医院級よりも総合病院級の療養機関で高くなる傾向がある。

診断名及び手術においては医師の個人差や慣習的要因の完全排除は困難なので資料の特性、患者の医療利用行動と疾病の固有特性、医師の診療過程と臨床環境、病院のネットワークと請求過程、健康保険給与制度などを十分に把握することで正しい解釈が可能になる。

患者標本資料の制限点はすべての標本資料が持っている共通の限界点として標本資料内の観測値は確率によって抽出された資料なので適正水準以上の標本数を確保することで代表性、有意性の保障を受けることができる。

例えば本患者標本資料で特定年齢帯の稀貴な疾患の発生頻度の場合、標本抽出頻度が非常に少なく代表性と説明力が劣ることになる。そのため標本資料の説明力は多頻度の傷病ほど大きくなり、傷病の頻度が下がれば減少するようになる。

これを補う方法として発生頻度が500件以下の傷病に関しては、全数提供する方案があり資料の提供領域を1年単位でなく一定期間のパネル資料を構築して抽出頻度率を高める方案も考慮している。

また標本資料変数の提供範囲を拡大して歯科、産婦人科などの特定診療科や65歳以上の高齢患者で標本資料の細分化の開発方案も考慮している。

健康保険審査評価院の患者標本資料を通じて、多様な保健医療分野で研究に活用されて公共の便益を向上することができることが期待される。

文末注

- 1) 大韓予防医学会、保健経済政策学会、保健情報統計学会、保健行政学会、韓国疫学会
- 2) 박병주, 성중미 我が国の糖尿病有病率推定及び DPP-4 抑制剤使用様相評価
- 3) 최상은 視力損失と失明による社会的疾病負担費用推計

参考文献

1. 김재용, 임지혜, 김화영 診療エピソード重症度補正及び予測用指標開発 健康保険審査評価院; 2007.
2. 남궁평 標本調査設計と分析 2007
3. 健康保険審査評価院調査研究室 標本調査方法論開発に関する研究 健康保険審査評価院; 2003.
4. 신의철ほか 診療情報の適正活用方案に対する研究 大韓予防医学会健康保険審査評価院; 2008.
5. 韓国保健情報統計学会保健社会の調査研究での統計と情報2009韓国保健情報統計学会秋季学術大会。
6. AHRQ. Design of the Nationwide Inpatient Sample(NIS). Agency for Healthcare Research and Quality; 2005.
7. AHRQ. Intriduction to the HCUP Nationwide Inpatient Sample(NIS). Agency for Healthcare Research and Quality; 2007.
8. Bethel. J. W. An Optiomum Allocation Algorithm for Multivariate Surveys, Proceedings of the Social Statistics Section. American Statistical Association; 1985.
9. Cheng. S. H. Introduction to National Health Insurance database. Chin. J. Public Health 1999; 18: 235-236.
10. Chen. C. J. Lin. L. H. Use of National Health Insurance database in academic research: experiences from analysis of major disease certification profile. Chin. J. Public Health 1997;16: 513-521.
11. Chatterjee. S. Multivariate Stratified Survey. Journal of American Statistical Association 1968.
12. Hidiroglou. M. A. Srinath. K. P. Problems associate with designing sub- annual business surveys. Journal of Business and Economic Statistics 1993.
13. Inho P. Hyunshik L. THE DESIGN EFFECT: DO WE KNOW ALL ABOUT IT?. Proceeding of the Annual Meeting of the American Statistical Association 2001.
14. National Health Research Institute. Researches of National Health Insurance claim data-base.

건강보험심사평가원

환자표본자료 (HIRA-NPS)

HIRA-NPS : National Patients Sample



01 신청 및 제공 방법

신청 접수 2012. 4. 30 (월) 부터

접수 방법 심평원 홈페이지 (www.hira.or.kr)
(홈페이지 → 정보 → 통계자료실 → 통계자료)에서
자료신청 양식을 다운받아 작성 후 직접 혹은
이메일, 팩스로 제출
※ 팩스번호 : 02-6710-5834, 담당 : 02-2182-2548
※ e-mail : songahchae@hiramail.net, 담당 : 02-2182-2544

처리 절차 자료 신청 → 건강보험심사평가원 계좌입금 →
입금확인 → 우편발송 혹은 직접수령

문의처 1644-2000 (건강보험심사평가원 콜센터)

02 교육안내

일시 2012. 5. 21 (월) 14:00 ~ 16:00

장소 서울특별시 서초구 효령로 168(서초동)
건강보험심사평가원 본원 지하 대강당

등록 신청 기간 2012. 4. 30 (월) ~ 2012. 5. 13 (일)

교육 등록 심평원 홈페이지 (www.hira.or.kr)
(홈페이지 → 정보 → HIRA교육 → 심사평가교육 →
교육일정 및 신청 → 환자표본자료(HIRA-NPS) 교육)

교육 내용 건강보험심사평가원 환자표본자료의 소개 및 활용 방법

문의처 02-2182-2515 kimlog2@hiramail.net

찾아오시는 길 지하철 3호선 남부터미널 6번 출구로 나와
300미터 직진

- 定義: 原資料 (raw data) で、個人や法人の情報を削除した二次資料
- 算出基準: 1年間の医療利用をしたすべての患者を対象に診療内容と処方履歴を含めて性別、年齢(5歳単位)の区間に応じて抽出
- -入院患者抽出率13%、外来患者抽出率1%
- -その年の1月~6月までの審査決定された1年分の診療内容で構築
- 提供先: 国内に居住する一般研究者
- 申込方法: 資料請求フォームをダウンロードして作成し、担当者に直接またはEメール、ファックスで提出
- -ファックス番号: 02-6710-5834、担当: 02-2182-2548 - e-mail: songahchae@hiramail.net、担当: 02-2182-2544, 02-6710-5834, 02-2182-2544
- 提供方法: 口座に入金された申請者に郵送または直接提供
- 手数料: 300,000ウォン(送料含む)
- ※口座番号: ハナ銀行100-200900-76804