

국민건강보험 청구자료원을 이용한 임상연구의 입문

류동열

이화여자대학교 의과대학 내과학교실, 이대목동병원 휴먼헬스빅데이터연구소

Introduction to the Medical Research Using National Health Insurance Claims Database

Dong-Ryeol Ryu

Department of Internal Medicine, Ewha Womans University School of Medicine,
Research Institute for Human Health Information, Ewha Womans University Mokdong Hospital, Seoul, Korea

All Korean citizens should join the National Health Security System by law. The National Health Insurance Service (NHIS) and the Health Insurance Review and Assessment Service (HIRA) are one of major components to support this system, and all data about medical expenses for the medical claims are stored and managed in the institutions. Recently, medical research using administrative claims databases has dramatically progressed in Korea and worldwide, and the methods how to use them are briefly reviewed in this article. Research using these databases have several strengths. Researchers can perform the complete enumeration survey in a real world. They can get new valuable findings because the number in the database is usually large enough to detect the minute difference with a big statistical power. They can obtain more detailed and reproducible results. Moreover, they can investigate a very rare disease or infrequent side effects of drugs. However, we must recognize that research using administrative claims database also has several incoherent limitations. These databases have not been constructed originally for research, but for reimbursement. Therefore, there are no important data including medical history and laboratory findings of each patient, which are crucial to adjust baseline characteristics. In addition, it is hard to discover causal relationship and direct association with the included information. In spite of limitations, researchers can easily use these databases for their research now than ever, and the results may be utilized not only to expand the academic knowledge but also to influence the determination of national healthcare policy. (**Ewha Med J 2017;40(2):66-70**)

Received January 13, 2017
Accepted April 10, 2017

Corresponding author

Dong-Ryeol Ryu
Department of Internal Medicine, Ewha
Womans University School of Medicine, 1071
Anyangcheon-ro, Yangcheon-gu, Seoul 07985,
Korea
Tel: 82-2-2650-2507, Fax: 82-2-2650-2501
E-mail: drryu@ewha.ac.kr

Key Words

Administrative claims; Database; Medical
research

서론

최근 다양한 국민건강보험 청구자료원에 접근이 용이해지고 이를 이용한 연구방법론이 발전되어 기존 자료원과 연구방법으로는

분석하기 어려웠던 수많은 연구 결과를 얻을 수 있게 되었다. 이로 인하여 임상 의사들은 질병의 유병률, 발생률, 질병으로 인한 주요 임상 결과의 발생률 등을 전국 단위로 전수조사를 할 수 있게 되었고, 질병의 사회-경제적인 부담, 새로운 치료 방법의 경제성 평

가, 그리고 보기 드문 희귀한 부작용에 대한 실마리를 찾아낼 수 있게 되었다.

이러한 자료원들에는 매우 많은 대상자를 포함하고 있기 때문에 의미 있는 작은 차이도 통계적으로 찾아낼 수 있는 분석도구가 되며, 오차범위가 작기 때문에 좀 더 정교하고 재현성이 높은 결과를 얻을 수도 있다. 또한, 결과 해석에 중대한 영향을 미치는 혼란변수 및 매개변수에 대한 보정이 가능하고, 임상적으로 중요한 결과인 사망, 질병의 재발 및 악화, 합병증 발생 등에 대해 분석할 수 있으며, 드물게 발생하여 적은 규모의 연구에서는 찾아내기 어려운 희귀한 질병이나 예측하지 못한 부작용을 찾아낼 수도 있다. 그 뿐만 아니라, 비용에 대한 분석이 가능하기 때문에 비용-효과 분석 또는 비용-효용 분석에 이용될 수도 있다.

이러한 장점에도 불구하고, 대규모 청구 자료원들은 구축된 원래의 목적이 연구용이 아닌 경우가 대부분이어서 연구를 위해서는 데이터베이스 분석 전문가의 도움을 받아 일차 가공을 거쳐야 하며, 보안이 필요한 자료원인 경우에는 일반 연구자의 접근이 제한적이고, 일반 연구자가 쉽게 파악할 수 없는 구조를 가지고 있기 때문에 분석 경험이 많은 통계 전문가와 협업이 필요하다. 또한, 두 가지 사건 사이의 선후관계를 파악하기 어려운 경우가 흔하게 있고 검사결과나 임상양상에 대한 의무기록 자료가 포함되어 있지 않은 경우가 대부분이며, 결과에 영향을 미치는 중요한 변수에 대한 보정을 시행할 수 없는 경우가 발생하기 때문에 원인-결과적 인과관계를 정확하게 파악하기 어려운 경우가 많다. 특히, 보험청구용 자료원의 경우 상병코드가 부정확하게 입력된 경우가 있을 수 있는데 이런 부정확성을 보완하고 질병을 정의하기 위해서는 여러 가지 다양한 방법을 추가하여 해당 학계에서 합의되는 적절한 조작적 정의(working definition)가 필요한 경우도 있다.

따라서, 연구자들이 연구하고자 하는 청구 자료원의 특성을 잘 파악하고, 경험 있는 통계전문가와 함께 분석 가능한 가설을 설정

한 뒤 과학적인 방법을 이용하여 연구를 수행하여야 가치 있는 결과를 얻을 수 있을 것으로 생각한다.

우리나라의 의료보험 제도는 1977년 처음으로 시행된 후 1989년 전국민으로 대상이 확대되었고[1,2], 2015년 12월 현재 의료보장 적용인구 5,203만 명 중 5,049만 명(97.0%)이 건강보험을 적용 받고 있으며 154.4만 명(3.0%)이 의료급여를 받고 있다[3]. 국내에 거주하는 모든 국민은 국민건강보험 가입자, 의료급여 수급권자, 그리고 유공자 등 의료보호대상자로 분류되어 기본권으로서의 의료보장을 받고 있다. 건강보험 가입자가 의료기관을 이용한 경우 건강보험심사평가원(Health Insurance Review & Assessment Service, HIRA)에서 의료기관이 청구한 진료비에 대한 심사를 진행한 뒤 그 결과를 국민건강보험공단 (National Health Insurance Service, NHIS)에 통보하며 공단은 이 결과에 따라 의료기관이 청구한 금액을 지불하게 된다(Fig. 1). 따라서, 진료와 관련된 모든 자료가 HIRA 및 NHIS 데이터베이스에 축적되게 되는데, 이 자료원을 연구 목적으로 개방하고 있기 때문에 절차에 따라 승인을 받은 후 연구에 활용할 수 있다.

본 론

1. 건강보험심사평가원 자료의 특성 및 이용 방법

건강보험심사평가원에서는 각 요양기관에서 수행한 진료의 적정성을 평가하고 청구된 진료비에 대한 심사를 위하여 데이터베이스를 구축하고 있다. 이 자료에는 상병명 등의 일반정보, 처치 및 수술료 등의 진료내역 정보, 그리고 청구비용 정보가 포함된 청구명세서, 약제 내역, 요양기관 정보, 의약품 정보가 포함되어 있고 이를 연구용으로 제공받을 수 있다. 그러나, 자격자료에 기반한 개인정보가 미비하고 임상자료가 없으며 일반연구자에게는 5년의 기간을 정해 자료를 제공하고 있기 때문에 장기간의 연구가 불가

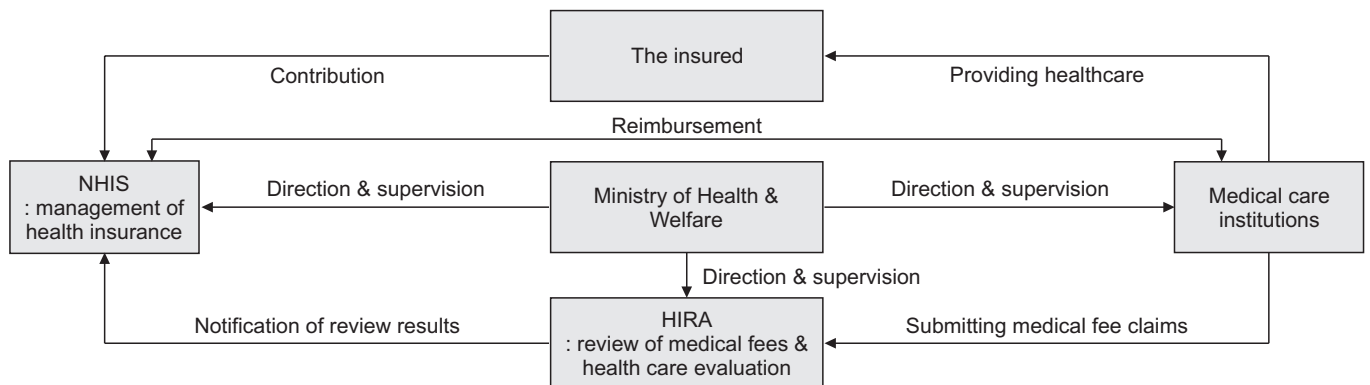


Fig. 1. Structure of the national health insurance program in Korea. NHIS, national health insurance service; HIRA, health insurance review and assessment service.

능하다는 단점이 있다. 심사평가원 자료의 구독 및 신청에 대해서는 보건 의료빅데이터 개방시스템 홈페이지(<http://opendata.hira.or.kr/home.do>)에서 관련 정보를 얻을 수 있다. 필요한 맞춤형 자료에 대하여 이메일로 상담하는 과정을 거친 후 홈페이지에서 이용신청을 하고 요청자료 신청 접수 후 2-4주 정도의 자료 추출 기간을 거친 뒤 연구자는 원격분석시스템을 통하여 분석할 수 있다. 원격분석시스템이 도입된 이후 심사평가원의 분석실에 직접 방문하여 분석하지 않아도 되고 분석실의 한정된 장소 때문에 발생하였던 불편감이 획기적으로 개선되었다. 그러나, 자료 제공 비용이 만만치 않기 때문에 처음 심사평가원 자료를 이용하여 분석을 시행하는 경우에는 '환자데이터셋'을 이용하여 충분히 기초자료 및 분석방법을 확립한 이후 전체 의료빅데이터 이용한 연구를 시행하는 것이 바람직할 것으로 사료된다. 환자데이터셋에 대한 정보와 신청은 역시 보건 의료빅데이터 개방시스템 홈페이지에서 진행할 수 있는데, 여기에는 입원환자 데이터셋, 전체환자 데이터셋, 고령환자 데이터셋, 그리고 소아청소년환자 데이터셋이 있으며, 1년간 의료서비스를 이용한 모든 환자를 대상으로 성별 및 5세 단위 연령구간에 따라 1%-20%의 비율로 층화계통추출한 자료이며 진료내역과 처방내역을 포함하고 있다.

심사평가원 데이터베이스를 이용하여 2005-2008년 사이 국내에서 혈액투석 및 복막투석을 시행한 투석 신환자를 대상으로 처음 선택한 투석방법에 따른 생존율 차이 여부를 분석할 수 있었고 [4], 65세 이상 고령의 투석 신환자의 사망과 관련된 위험인자 분

석도 가능하였다[5].

2. 국민건강보험공단 자료의 특성 및 이용 방법

국민건강보험공단 자료원은 심사평가원 자료와 거의 비슷한 구조를 가지고 있으나, 심사평가원 자료에는 없는 건강보험료 산정에 필요한 보험가입자의 자격자료와 건강보험공단에서 시행중인 건강검진에 대한 자료가 추가되어 있다는 점이 다른 점이다. 따라서, 의료기관 이용자의 개인적 특성인 소득수준, 거주지, 보험자격의 상실 여부 등에 대한 정보가 포함되어 있기 때문에 상기의 개인적인 기초 특성을 보정해야 할 필요성이 있는 연구를 시행할 때 도움이 된다. 또한, 건강보험공단에서는 건강보험가입자들을 대상으로 일반건강검진, 생애전환기 건강검진, 암검진, 그리고 영유아 건강검진을 시행하고 있는데 이에 대한 자료가 수집된 건강검진 데이터베이스를 분석해 볼 수 있는 것도 큰 매력 중에 하나이다. 건강보험공단 자료에 대한 정보는 공단 홈페이지(<https://nhiss.nhis.or.kr/bd/ay/bdaya001iv.do>)에서 확인할 수 있으며, 크게 표본연구 데이터베이스와 맞춤형 데이터베이스로 분류된다. 맞춤형 데이터베이스는 공단이 수집, 보유, 관리하는 건강정보를 정책 및 학술 연구 목적으로 제공하는데, 자료제공을 승인 받는 경우 현재로서는 원주 건강보험공단 본부의 분석실에서만 분석이 가능하지만 조만간 원격접속에 의한 자료 분석이 가능해질 전망이다.

표본연구 데이터베이스는 연구자가 기관생명윤리위원회의 심의를 받은 연구계획서를 지정 서식과 함께 제출하여 공단의 심의를

Table 1. Summary of major results on CKD and ESRD analyzed using healthcare database in Korea

Parameter	Figure	Reference
CKD prevalence in adults	10.6%	[7]
Annual growth rate in number of CKD patients (2006-2010)	8.2%	[8]
Number of treating CKD patients	0.16 million	[9]
Estimated number of treating CKD patients	3.7%	-
CKD as a mortality cause	12/100,000 PY	[6]
Total medical cost of CKD in 2014	KRW1.4 billion	[9]
Annual growth rate in cost of CKD patients (2006-2010)	10.2%	[8]
Ratio of annual medical cost of CKD stage 3, 4, and 5 over that of hypertension	3.0, 4.9, and 21.1	[6]
Crude all-cause mortality rate of ESRD patients	116/1,000 PY	[10]
Overall hazard ratio of PD over HD	1.20	[4]
Hazard ratio of PD survival in 2008 vs. 2005	0.75	[11]
Standard mortality ratio of ESRD patients compared to the general population	10.9	[12]
Total medical cost of HD in 2013	KRW1.6 billion	[13]
Annual growth rate in cost of HD patients (2009-2013)	6.4%	[13]
Ratio of annual medical cost of HD and PD over that of hypertension	86.7 and 65.1	[6]

CKD, chronic kidney disease; ESRD, end-stage renal disease; PY, patient-years; PD, peritoneal dialysis; HD, hemodialysis.

받으면 개인정보 식별이 불가능한 형태의 코호트 자료를 배송 받게 되므로, 연구실에서 편리하게 분석할 수 있는 장점이 있다. 이 중 표본코호트 데이터베이스는 2002년 당시 전체 국민의 약 2%인 100만 명의 건강보험가입자 및 의료급여수급권자(외국인 제외)를 층화표본추출하고 이들을 2013년까지 추적관찰한 자료로, 이들의 자격자료, 진료자료, 건강검진자료 및 영양기관자료가 모두 포함되어 있다. 특히, 통계청에서 수집한 사망에 대한 자료가 포함되어 있기 때문에 사망을 임상결과로 한 연구에서 사용할 수 있는 소중한 자료원이다.

건강검진코호트 데이터베이스는 2002년 자격유지자 중 2002-2003년 일반건강검진을 받은 전체 500여만 명 중의 약 10%를 추출한 뒤 2013년까지 추적관찰한 자료로 여기에도 역시 이들의 자격자료, 진료자료, 건강검진자료 및 영양기관자료가 모두 포함되어 있다. 따라서, 우리나라 성인의 건강행태 및 이와 관련된 임상 결과를 분석하는 연구를 진행할 수 있다. 그 외, 60세 이상의 대상자들로 구성된 노인코호트 데이터베이스, 직장여성코호트 데이터베이스, 그리고, 영유아검진코호트 데이터베이스가 있다.

2016년에는 건강보험공단 맞춤형 자료 및 표본코호트 자료를 분석 결과가 포함된 '국가건강검진 항목 중 신장질환 검진의 타당성 분석 연구'가 발표된 바 있으며[6], 만성콩팥병 단계의 변화에 따른 혈압 조절 정도를 분석한 연구를 진행하고 있다.

3. 기타 자료원

통계청에서는 매년 인구, 보건 등에 대한 통계를 발표하고 이를 홈페이지에 고시하고 있다(<http://kosis.kr/>). 질병관리본부에서는 다양한 표본조사를 시행하고 있는데, 이 중 제7기 사업이 시행중인 국민건강영양조사에는 국민의 광범위한 건강행태 및 검사 자료들이 포함되어 있고 이를 분석한 결과가 매년 국민건강통계로 발표되고 있는데(<https://knhanes.cdc.go.kr/knhanes/index.do>), 간단한 회원가입 만으로 원시자료를 구독하여 연구에 이용할 수 있다. 또한, 한국인유전체역학조사를 수행하고 있는데, 여기에는 안산코호트 등의 전체 6개 코호트로 구성되어 2014년말까지 누적 기반조사 기준 약 24만 5천명 규모의 코호트를 구축하였고 수집된 역학, 임상, 유전 정보자료 및 인체유래물을 국내 연구자에게 제공한다고 하니(<http://www.cdc.go.kr/CDC/contents/CdcKrContentView.jsp?cid=24606&viewType=CDC&menuIds=HOME001-MNU1136-MNU1223-MNU1348>) 많은 관심을 가져볼 만 하다.

4. 국민건강보험공단 청구자료원을 이용한 임상연구 결과의 예시

지금까지 소개한 청구 자료원을 이용하여 한국의 만성콩팥병 및 말기신부전 현황을 분석한 결과를 Table 1과 같이 정리하였다. 만성콩팥병의 유병률이 높고 증가도 빠르게 진행하고 있지만 제

대로 인식되지 못 하여 현재 치료중인 환자의 비율이 상대적으로 매우 낮다. 또한, 우리나라 국민 사망원인 중 주요한 원인이며, 개인적으로 또는 국가적으로 지출해야 할 의료비용도 매우 높고 의료비 증가율이 가파른 질병이다. 투석치료를 시작하는 말기신부전 환자의 사망율은 더욱 높아 일반인구집단 사망율에 비하여 비교하는 표준화 사망비가 10.9이며, 의료비용도 투석전 단계의 만성콩팥병 환자에 비하여 폭발적으로 증가된다. 이러한 연구결과들은 환자 개인별 위험인자를 발굴하여 예방 및 조기 치료를 할 수 있도록 하는 데에도 도움이 되지만, 국가 의료정책 수립에도 귀중한 근거가 될 수 있을 것으로 생각한다.

결론

공공 데이터 개방 정책에 따라 심사평가원과 국민건강보험공단의 대규모 자료원을 이용한 다양한 연구 진행이 가능하게 되었다. 이들은 전국민의 의료 이용을 기반으로 한 자료원이기 때문에 우리나라의 실제 보건 의료 상황에 대한 실태를 파악하기에는 무엇보다도 좋은 자료이다. 본 자료원들에서는 행위별 수가제 하의 세부 의료이용 내역 확인이 가능하고 약물 및 처치에 대한 자세한 정보가 포함되어 있는 것도 장점이다. 대상자가 전국민이기 때문에 드물게 발생하는 질병이나 사건에 대한 연구가 가능하다. 그러나, 의료이용은 개인정보에 속하고 이에 대한 보호 대책 마련이 필수적이므로 자료 이용에는 엄격한 제한이 있고 자료 이용에 대한 사전 심사와 사용 과정 중 책임이 뒤따른다. 또한, 질병군별 포괄 수가제 대상 질병의 경우에는 구체적 진료내역이 없는 단점이 있고, 급여가 인정된 의료이용만 포함되므로 비급여 항목 사용에 대한 연구는 불가능하다. 또한, 상병코드로 입력된 진단명이 부정확할 수 있으므로 질병의 정의를 명확하게 하려는 노력이 선행되어야 한다.

그럼에도 불구하고, 본 자료들의 특성, 환자의 의료이용 행태, 질병의 고유특성을 잘 이해하고, 진료과정 및 임상환경, 병원의 전산망과 청구과정, 건강보험급여제도에 대한 명확한 이해가 있다면 학술 목적이나 정책에 반영할 수 있는 매우 귀중한 연구 결과를 얻을 수 있을 것으로 생각된다.

References

1. Jeong HS. Korea's National Health Insurance: lessons from the past three decades. *Health Aff* 2011;30:136-144.
2. Kwon S. Thirty years of National Health Insurance in South Korea: lessons for achieving universal health care coverage. *Health Policy Plan* 2009;24:63-71.
3. Korea Health Industry Development Institute. Major statistics in National Health Insurance, 2015 [Internet]. Cheongju: Korea

- Health Industry Development Institute; 2016 [cited 2017 Apr 12]. Available from: <http://www.nhis.or.kr/bbs7/boards/B0074/17042?boardKey=29&sort=sequence&order=desc&rows=10&messageCategoryKey=&pageNumber=1&viewType=generic&targetType=12&targetKey=29&status=&period=&startdt=&enddt=&queryField=&query=>.
4. Kim H, Kim KH, Park K, Kang SW, Yoo TH, Ahn SV, et al. A population-based approach indicates an overall higher patient mortality with peritoneal dialysis compared to hemodialysis in Korea. *Kidney Int* 2014;86:991-1000.
 5. Lee S, Ryu JH, Kim H, Kim KH, Ahn HS, Hann HJ, Cho Y, et al. An assessment of survival among Korean elderly patients initiating dialysis: a national population-based study. *PLoS One* 2014;9:e86776.
 6. Jo MW. An evaluation of national screening program for chronic kidney disease. Cheongju, Korea: Centers for Disease Control & Prevention; 2016.
 7. Ji E, Kim YS. Prevalence of chronic kidney disease defined by using CKD-EPI equation and albumin-to-creatinine ratio in the Korean adult population. *Korean J Intern Med* 2016;31:1120-1130.
 8. Health Insurance Review and Assessment Service. HIRA report: [Internet]. Seoul: *Health Insurance Review and Assessment Service*; 2011 [cited 2017 Apr 12]. Available from: http://www.mohw.go.kr/front_new/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&PAGE=1&CONT_SEQ=250266&SEARCHKEY=TITLE&SEARCHVALUE=%B8%B8%BC%BA%BD%5%BA%CE%CO%FC.
 9. National Health Insurance Service; Health Insurance Review & Assessment Service. 2014 National Health Insurance statistical yearbook. Seoul: National Health Insurance Service; 2015.
 10. Kim H, Kim KH, Ahn SV, Kang SW, Yoo TH, Ahn HS, et al. Risk of major cardiovascular events among incident dialysis patients: a Korean national population-based study. *Int J Cardiol* 2015;198:95-101.
 11. Ryu JH, Kim H, Kim KH, Hann HJ, Ahn HS, Lee S, et al. Improving survival rate of Korean patients initiating dialysis. *Yonsei Med J* 2015;56:666-675.
 12. Choi H, Kim M, Kim H, Lee JP, Lee J, Park JT, et al. Excess mortality among patients on dialysis: comparison with the general population in Korea. *Kidney Res Clin Pract* 2014;33:89-94.
 13. Health Insurance Review and Assessment Service. HIRA report: [Internet]. Seoul: Health Insurance Review and Assessment Service; 2015 [cited 2017 Apr 12]. Available from: http://www.mohw.go.kr/front_new/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&page=1&CONT_SEQ=318523&SEARCHKEY=TITLE&SEARCHVALUE=%C7%F7%BE%D7%5%F5%BC%AE%20%C8%AF%C0%DA%BE%C8%C0%FC%20%C1%F6%C7%A5%BF%A1%20%B6%D1%B7%C7%7%D1%20%B0%B3%BC%B1%20%C8%BF%B0%FA%20%B3%AA%C5%B8%B3%AA.