보건행정학회지 : 제 21 권 제 1 호
Korean J. of Health Policy & Administration
Vol. 21, No. 1, March. 2011 ; 77-92
DOI : 10.4332/KJHPA.2011.21.1.077

▧ 원 저 ▧

연구를 위한 건강보험 청구자료 요구 및 이용 요인분석

이정아[*], 오주환[*†], 문상준[*,**], 임준태[***], 이진석[*,***], 이진용[****], 김 윤[*,***]
서울대학교 의학연구원 의료관리학연구소[*], 질병관리본부 만성질환관리과[**], 서울대학교 의과대학 의료관리학교실[***], 건양대학교 의과대학 예방의학교실[*****]

<Abstract>

Assessment of Needs and Accessibility Towards Health Insurance Claims Data

Jung-A Lee[*], Juhwan Oh[*†], Sangjun Moon[*,**], Jun Tae Lim[***], Jin-Seok Lee[*,***], Jin-Yong Lee[****], Yoon Kim[*,***]
Institute of Health Policy and Management, Medical Research Center, Seoul National University[*], Division of Chronic Disease Control, Korea Center for Disease Control & Prevention[**], Department of Health Policy and Management, Seoul National University College of Medicine[***], Department of Preventive Medicine, College of Medicine, Konyang University[****]

Objectives : This study examined the health policy researchers' needs and their accessibility towards health insurance claim datasets according to their academic capacity.

Methods : An online questionnaire to capture relevant proxy variables for academic needs, accessibility, and research capacity was constructed based on previous studies. The survey was delivered to active health policy researchers through three major scholarly associations in South Korea. Seven-hundred and one scholars responded while the survey as open for 12

* 접수 : 2011년 2월 8일, 최종수정 : 2011년 3월 28일, 게재확정 : 2011년 3월 29일
† 교신저자 : 오주환, 서울대학교 의학연구원 의료관리학연구소, 서울시 종로구 연건동 28,
전화번호 : 02-740-8361, Fax : 02-743-2009, E-mail : oh328@snu.ac.kr
* 이 논문은 2010년도 건강보험심사평가원 "건강정보서비스 중장기 마스터플랜 수립 연구"의 연구비 지원을 받았음.

days (starting on December 20th, 2010). Descriptive statistics and logistic regression analysis were carried out.

Results : Regardless of the definition for operational needs, the prevalent needs of survey respondents were not met with the current provision of claim data. Greater research capacity was shown to be correlated with increased demand for claim data along with a positive correlation between attempts to obtain claim datasets and research capacity. A greater research capacity, however, was not necessarily correlated with better accessibility to the claim data.

Conclusions : The substantial unmet need for claim data among the healthcare policy research community calls for establishing proactive institutions which could systematically prepare and make available public datasets and provide call-in services to facilitate proper handling of data.

Key words : Health insurance claim data, Unmet needs, Accessibility, Academic capacity

I. 서 론

우리나라는 전국민 단일 건강보험제도로 운영되고 있어, 방대한 양의 건강보험 청구자료가 지속적으로 축적되고 있다. 이 청구자료에는 진료비 청구내역 대부분이 전산화된 형태로 수록되어 있어, 통계청의 사망자료나 암등록 자료 등의 다양한 자료와 연계한 의학연구를 수행할 수 있다. 실제 국립암센터에서는 통계청의 사망자료를 연계하여 내원한 암 환자의 비용 코호트 연구를 2004년부터 시행하고 있다(국립암센터, 2010). 또한, 국민건강보험 청구자료에는 의료기관에서 제공하는 서비스 관련 정보가 수록되어 있어, 패널 자료나 지역사회건강조사 자료를 연계한다면 국민건강과 관련된 보건정책을 수립하는 등의 정보 활용 측면에서 그 가능성과 유용성이 클 것이다(김재용, 2007). 외국에서는 건강보험자료의 한계를 인정하면서도 그 잠재적 가능성에 주목하며 1990년대 초부터 행정자료를 이용한 의료의 질 평가와 환자 안전에 대한 연구를 시작으로 청구자료의 활용에 대한 많은 논의가 이루어지고 있고, 이를 활용하는 연구들이 급격히 확대되고 있다(Iezzoni, 2002 ; Pandiani & Banks, 2003 ; Zhan & Miller, 2003 ; Roos 등, 2004).

통계청에서 실시하는 인구센서스 자료와 질병관리본부에서 실시하는 국민건강영양조사 자료는 인터넷을 통해 누구라도 접근 가능한 형태로 제공하고 있어, 연구기관들이 이 자료가 필요한 경우 손쉽게 활용할 수 있다. 이러한 집단적 자료는 우리나라 국민의 건강관련행태의 변화, 진료행태에 영향을 주는 요인과 건강상태의 파악 등 국가나 지역 단위의 분석이 가능하다(대한예방의학회, 2008). 따라서 전국민 건강보험제도에 의해 집적된 건강정보에 대해서도 개인 자료뿐만 아니라 집단적 자료로의 광범위한 활용에 대한 체계적인 연구와 노력이 필요하다.

그러나 국민건강보험 청구자료는 자료의 접근이 제한되어 일부 공공기관들을 중심으로 제한된 분석만이 시도되고 있으며, 이 자료를 활용한 연구를 하고자 하는 연구자에게 접근이 순조롭지 않아 청구자료가 정책연구 자료로서 충분히 활용되지 못하고 있다. 정부나 공공기관에서 발주된 연구를 수행하는 경우, 일정한 검토과정을 거쳐 개인정보를 제외한 자료가 건강보험심사평가원이나 건강보험공단을 통해 제한적으로 제공되고는 있다. 하지만, 자료를 구득하여도 자료의 변수나 특성을 충분히 파악하지 못하여 부정확한 결론을 내거나 분석 자체를 포기하는 경우 등 자료 활용에 대한 지원이 부족한 것으로 알려져 있다(김재용, 2007). 이런 정책연구 환경에서 보다 구체적인 국민건강보험 청구자료에 대한 요구도, 자료 제공의 적절성, 자료 구득의 장애요인, 청구자료의 특성에 대한 설명이 적절히 연구자에게 제공되고 있는지에 대한 조사연구가 필요하지만 아직까지 많지 않은 실정이다.

한편, 이런 요구도를 조사하는 방법에는 면접 설문조사, 전화 설문조사를 포함한 여러 가지 방법이 있는데, 그중에서 온라인 설문조사는 비용 효과적인 조사 방법으로 알려져 최근 연구자들의 관심을 끌고 있다. 인터넷을 통한 온라인 설문조사는 기존의 설문조사 방식에 비해 빠른 속도와 낮은 비용으로 조사를 가능하게 하며, 응답자의 답변 내용에 따라 설문내용을 조정하는 상호 대화적(interactive)인 기능도 반영되어 있다(Weible & Wallace, 1998 ; Sheehan & McMillan, 1999 ; Bethell 등, 2004 ; Kaplowitz 등, 2004). 응답자들이 조사 장소에 나오지 않아도 되므로 바람직한 선택에 대해 긍정적으로 반응하는 효과(social desirability)가 낮고 조사원에 의한 오프라인 조사보다 솔직한 답변을 제공해 준다는 장점도 있다(Paolo 등, 2000). 국내에서는 인터넷의 발전에 따라 일부 마케팅 영역에서 온라인 설문조사 비중이 높아지고 있으나 보건의료 분야에서는 온라인 설문조사를 통한 연구가 상대적으로 적은 편이다.

이에 이 연구에서는 온라인 설문조사 방법을 이용하여 연구자들의 국민건강보험 청구자료에 대한 요구도와 자료제공 상황을 연구역량과의 관계를 토대로 분석함으로써, 청구자료가 연구역량과 수요에 맞게 적절히 제공되어 왔는지를 파악하여, 향후 국민건강보험 청구자료 제공 서비스 개선을 위한 근거 자료를 제공하고자 한다.

II. 연구방법

1. 조사대상 및 자료수집

우리나라 보건의료부문 연구자 집단의 청구자료 요구도와 자료 제공 양상을 측정하기 위해 한국보건행정학회(2,300명), 대한예방의학회(800명), 한국의료QA학회(4,100명) 회원을 대상으로 온라인 설문조사를 실시하였다. 설문지는 온라인 설문조사 전문사이트를 활용하여 구성하였고, 조사기간은 2010년 12월 20일부터 2010년 12월 31일까지 12일 동안이었다. 사전예고 메일을 17일(금) 한국보건행정학회 회원에게 발송하였고 3일 후인 20일(월) 자료제공 서비스 개선을 위한 조사 본 메일을 발송하였다. 같은 방법으로 20일(월) 대한예방의학회와 한국의료QA학회 회원에게 사전예고 메일을 발송하였고, 다음 날인 21일(화) 본 메일을 발송하였다. 수집된 총 응답 수는 701명이었다.

2. 조사내용

이 연구에서 조사한 변수들로는 청구자료의 요구도와 구득 경험에 관한 내용으로 구성되어 있으며, 이는 크게 인구사회학적 특성, 연구자 역량관련 요인, 청구자료 구득 과정별 장애요인, 자료 제공 문제의 해결 방안이 포함되었다. 설문도구는 선행연구의 대한예방의학회(2008)에서 개발한 조사도구를 활용하여 1차 설문지를 작성하였으며, 공동연구자들의 내부 회의를 거쳐 최종 설문도구를 확정하였다.

사용된 종속변수는 4문항으로 설문 내용은 지난 5년간 청구자료를 이용한 연구를 구상한 경험이 있는지 여부, 지난 5년이내 청구자료가 필요한 연구 구상시마다 해당 자료를 요청하여 잘 구득하였는지 여부(자료 요청한 적 없음/요청 포기/요청했으나 매번 구득 못함/요청시마다 구득함), 자료를 파악하고 분석하는데 받은 설명 자료만으로 어려움이 없었는지 여부와 사용이 많은 변수를 정기적으로 공용데이터 셋으로 만들어 간편한 절차를 통해 자료를 제공할 경우 연구에 사용할 의향이 있는지 여부로 구성하였다. 이 연구에서는 4개의 종속변수를 다음과 같이 조작적으로 정의하여 8개의 종속변수로 분석하였다. 청구자료에 대한 요구도(needs), 청구자료에 대한 접근이 어려워 연구 요구도가 없어진 대상자를 포함한 청구자료의 요구도(controlled needs), 청구자료 요청 시도(trial), 청구자료에 대한 접근이 어려워 연구 요구도가 없어진 대상자를 요구는 있으나 구득 시도를 하지 않는 역의 인과관계(reverse causality effect)가 있는 것으로 간주하여 광의로 정의한 청구자료 요청 시도(controlled trial), 청구자료 구득(taking), 항상 구득 가능했던 경우만을 '구득 성공'으로 보는 협소한 정의에 따른 청구자료 구득(strict taking), 청구자료 변수에 대한 이해 정도(understanding), 공용데이터 셋 사용 의사(open data set)로 구분하여 분석하였다.

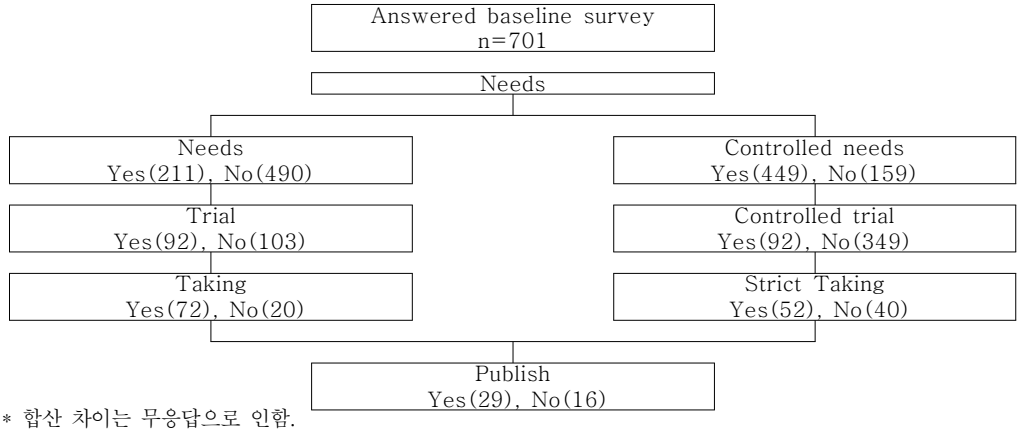
독립변수 중 연구자 '역량 관련 요인'의 대리지표(proxy)로 최종 학위, 연구 경력과 함께

지난 1년간의 교신 또는 1저자 논문, 공저자 논문, 2차 자료 활용 논문, 학술발표 횟수, 2차 자료 활용 학술발표 횟수, 2차 자료 활용 연구보고서 편수를 0편, 1편(회) 이상으로 나누어서 살펴보았고, 공저논문의 경우는 최근 1년간 3편 이상과 미만으로 나누어 분석하였다. '인구사회학적 특성요인'으로 성, 연령, 전공 분야, 소속 기관 등을 조사하였다. '청구자료 구득 과정별 장애요인'으로는 청구자료를 활용한 연구 구상 여부, 청구자료를 활용한 연구 미구상 사유, 청구자료 요청 경험 여부, 청구자료 요청 거절 시점, 청구자료 구득 경험 여부, 청구자료 구득 포기 사유, 청구자료를 파악하고 분석하는데 설명 자료의 어려움 정도, 청구자료 분석 여부, 분석 포기 사유, 2차 자료를 활용한 논문 및 연구보고서 공표 여부를 살펴보았다.

로지스틱 회귀분석에는 연구자의 역량 관련 요인(8문항), 청구자료 구득 과정별 장애요인(10문항), 자료 제공 문제의 해결 방안(3문항) 등으로 구성하였다.

3. 분석방법

이 연구에서는 수집된 701명의 조사 결과에 대한 기술통계 분석을 통해 일반적 특성에 따른 연구대상자의 분포와 청구자료에 대한 구득 장애요인을 확인하였다. 자료에 대한 통계분석은 SPSS 15.0 for Windows를 이용하였으며, 연구자의 청구자료에 대한 요구도, 청구자료 구득 시도, 청구자료 구득, 청구자료 변수에 대한 이해 정도와 공용데이터 셋 사용 의사를 종속변수로 하고, 연구자 역량 관련 요인과의 관계를 분석하였다. 연구자의 역량 요인을 각각 독립변수로 하여 로지스틱 회귀분석을 실시하였다. 각 종속변수는 청구자료를 활용한 연구를 구상해 본 적이 있는지 여부에 따라 청구자료 요청 경험, 청구자료 구득 경험, 분석 및 공표 여부의 문항으로 넘어가는 조건부 확률을 사용하여, 청구자료 구득 시도에 대한 분석은 자료 요구도가 있다고 답한 응답자만 분석하고, 자료 구득 여부는 자료 구득 시도를 했다는 응답자만 분석하였다(그림 1).



* 합산 차이는 무응답으로 인함.

[그림 1] 온라인 설문조사 분석 대상자

III. 연구결과

1. 응답자의 일반사항

조사 응답자의 일반적 현황을 살펴본 결과는 표 1과 같다. 설문항목의 마지막으로 조사된 인구사회학적 특성 변수들에 대한 선택적 무응답이 전반적으로 많았다. 응답자만의 분포를 보면, 여성이 남성보다 많았고, 연령 분포는 40대가 가장 많았으며, 30대가 다음을 차지하였다. 응답자의 전공 분야는 '의학/약학/간호학/치의학/한의학' 전공자가 가장 많았고, '정책학/보건학'이 다음을 차지하였다. 소속 기관의 경우, 대학 소속이 가장 많았고, 공공 연구기관이 다음을 차지하였다. 최종 학위가 박사 미만인 경우가 박사 이상인 경우보다 조금 더 많았으며, 연구 경력에서는 5년 미만인 경우가 5년 이상인 경우보다 많았다.

표 1. 응답자의 일반현황

(단위: 명, %)

구 분	명(%)	구 분	명(%)
성별		소속 기관	
남	117(16.7)	대학	180(25.7)
여	164(23.4)	공공 연구기관	47(6.7)
무응답	420(59.9)	민간 연구기관	26(3.7)
연령		의료기관	27(4.0)
≤29	24(3.4)	무응답	420(59.9)
30–39	100(14.3)	최종 학위	
40–49	114(16.3)	박사 미만	182(26.0)
50–59	40(5.7)	박사 이상	99(14.1)
≥60	3(0.4)	무응답	420(59.9)
무응답	420(59.9)	연구 경력	
전공 분야		5년 미만	266(37.9)
정책학/보건학	105(15.0)	5년 이상	213(30.4)
의학/약학/간호학/치의학/한의학	142(20.3)	무응답	222(31.7)
경제학/경영학	20(2.9)		
통계학/역학	11(1.6)		
인문사회학	3(0.4)		
무응답	420(59.9)		

2. 연구환경 및 역량 특성

연구 환경 및 역량에 관한 질문에서는, 최근 1년 동안 교신 또는 1저자 국내외 논문이 있는 응답자가 전체의 26.2%였으며, 공저자 논문편수는 3편 이상인 경우가 16.3%를

차지하였다. 2차 자료를 이용한 논문이 있는 응답자가 18.7%를 차지하였고, 27.8%의 응답자가 1회 이상의 국내외 학술발표를 지난 한 해 동안 하였으며, 2차 자료를 이용한 학술발표 경험이 지난 1년간 있던 경우가 12.1%였다. 지난 1년간 2차 자료를 이용한 연구보고서를 출판한 적 있는 응답자는 13.7%였다.

청구자료에 대한 요구가 없었다는 응답자가 69.9%로 많았고, 청구자료를 이용한 연구를 구상해 보지 않은 이유로는 해당 자료를 이용한 연구에 관심이 없는 경우가 많았으며(39.8%), 자료에 대해 파악하기 곤란하거나, 자료구득이 어렵다고 알고 있어서, 혹은 자료구조파악이 곤란해서 등의 이유로 연구구상을 하지 않았다는 응답이 각각 36.6%, 12.0%, 8.3%였으며, 분석 장소 제한으로 인한 경우도 12.5%였다. 청구자료 요청한 사람의 78.3%가 한 번이라도 구득해 본 적이 있었다. 한편, 청구자료 구득을 포기한 이유로는 '분석에 필요한 변수를 제외하고 제공해서'라는 응답이 포기한 이유 전체 중에서 70.0%로 가장 높았다. 받은 자료에 대한 설명 요청이 쉽지 않아 자료파악을 못해서 분석을 포기한 경우도 자료를 받은 연구자의 34.7%에 달하였다. 자료를 받았던 경우, 64.4%에서 논문 및 연구보고서를 공표한 것으로 나타났다. 청구자료를 연구에 활용하는데 필요정보 부족, 변수설명 부족 등 자료에 대한 접근성으로 인해 연구의 구상이 전개되지 못하고 포기되는 경우를 접근성 때문에 연구요구가 사라진 사례로 정의하여 이를 적극적으로 청구자료를 활용한 연구요구가 있는 경우로 적극적으로 정의할 때, 701명 중 449명이 연구요구가 있으나, 실제 매번 자료를 구득하였다고 답한 경우는 52명에 지나지 않았고, 연구결과를 출판하였다고 응답한 경우는 29명에 불과했다. 통상적인 연구요구도 정의에 의해 볼 때, 211명이 연구요구가 있었으나, 72명만이 한 번 이상 자료를 얻었던 적이 있었다고 답하였고, 29명이 출판하였다고 응답하였다.

향후 청구자료 제공 문제의 해결 방안으로는 자료신청자에 대한 청구비용의 증액을 통해 해야 한다는 응답보다는, 기관의 인력증원, 장비보강, 예산증액을 통해 해결해야 한다는 응답이 70.0%로 더 많았으며, 향후 공용데이터 셋 이용의사에 대해서는 65.6%가 이용하고 싶다고 응답하였다. 응답자가 생각하는 합리적인 자료제공 소요시간으로는 1개월이 28.0%로 가장 높았다.

3. 연구자의 역량에 따른 자료 요구도

다중 로지스틱 회귀분석을 통해 각각의 종속변수에 영향을 미치는 역량요인의 영향력을 살펴본 결과는 다음과 같다.

먼저 청구자료 요구도(needs)의 경우, 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서, 최종 학위가 박사 미만인 경우보다 박사 이상인 경우, 연구 경력이 5년 미만인 경우보다 5년 이상인 경우, 공저자 논문이 3편 미만인 경우보다 3편 이상인 경우, 2

표 2. 응답자의 연구 환경 및 역량 특성

(단위: 명, %)

구 분	명(%)	구 분	명(%)
교신 또는 1저자 국내외 논문 편수*		연구 구상 여부	
0편	206(29.4)	없었음	490(69.9)
1편 이상	184(26.2)	있었음	211(30.1)
무응답	311(44.4)	연구 미구상 사유‡	
공저자 논문 편수*		해당 자료에 대한 연구 관심 없어서	159(39.8)
3편 미만	276(39.4)	자료에 대해 구체적으로 알 방법 없어서	146(36.6)
3편 이상	114(16.3)	자료 구득 쉽지 않다고 알고 있어서	48(12.0)
무응답	311(44.4)	자료 구득해도 자료구조 파악 어려워서	33(8.3)
2차 자료를 이용한 논문 편수*		자료 분석장소 제한으로	50(12.5)
0편	259(36.9)	자료 요청 경험 여부†	
1편 이상	131(18.7)	요청 안함	103(48.8)
무응답	311(44.4)	요청함	92(43.6)
국내외 학술발표 횟수*		무응답	16(7.6)
0편	171(24.4)	자료 요청 거절 시점†	
1회 이상	195(27.8)	즉시 거절	17(18.5)
무응답	335(47.8)	심의결과 후 거절 통보	16(17.4)
2차 자료를 이용한 학술발표 횟수*		두 경우 모두 있었음	8(8.7)
0편	281(40.1)	무응답	51(55.4)
1회 이상	85(12.1)	자료 구득 경험 여부†	
무응답	335(47.8)	구득 못함	20(21.7)
2차 자료를 이용한 연구보고서 편수*		구득함	72(78.3)
0편	270(38.5)	자료 이해정도†	
1편 이상	96(13.7)	별 어려움 없었음	20(27.8)
무응답	335(47.8)	전화 설명요청으로 극복할 수 있었음	24(33.3)
구득 포기 사유‡		자료설명요청 쉽지 않아 그냥 분석함	27(37.5)
분석자료 장소 제한에 응하기 어려워서	7(23.3)	무응답	1(1.4)
분석에 필요한 변수를 제외하고 제공해서	21(70.0)	해결 방안	
자료 제공까지 걸리는 시간이 너무 길어서	7(23.3)	인력증원, 장비 보강, 예산증액	491(70.0)
Paper version으로 자료가 제공되어서	1(3.3)	자료 신청자의 자료 이용료 인상	39(5.6)
이용 수수료가 너무 비싸서	4(13.3)	기타	14(2.0)
자료 분석 여부†		무응답	157(22.4)
분석 안함(못함)	25(34.7)	공용데이터 셋 이용의사	
분석함	45(62.5)	아니오	99(14.1)
무응답	2(2.8)	예	460(65.6)
분석 포기 사유‡		무응답	142(20.3)
본인의 분석능력 부족	5(7.0)	합리적인 자료제공 소요시간	
자료(변수) 설명 지원 서비스 부족	22(31.0)	1주	144(20.5)
자료 구득 이후 분석 필요성 없어짐	6(8.5)	2주	142(20.3)
본인의 분석시간 부족	11(15.5)	1개월	196(28.0)
무응답	27(38.0)	2개월	42(6.0)
논문 및 보고서 공표 여부†		6개월	11(1.6)
공표 안함	16(35.6)	무응답	166(23.7)
공표함	29(64.4)		

* 2010년 1년의 실적

† 조건부 확률

‡ 복수응답

차 자료 이용 논문이 없는 경우보다 1편 이상인 경우, 2차 자료 활용 학술발표가 없는 경우보다 1회 이상인 경우, 2차 자료 활용 연구보고서가 없는 경우보다 1편 이상인 경우 청구자료에 대한 요구도가 높은 것으로 나타났다.

표 3. 자료 요구도와 자료 요청 시도 여부에 대한 회귀분석

Variables	Needs		Controlled needs*		Trial		Controlled trial†	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
최종 학위								
박사 미만	1		1		1		1	
박사 이상	2.64	1.46–4.80	1.52	0.69–3.34	5.03	2.0–12.611	5.58	2.61–11.91
연구 경력								
5년 미만	1		1		1		1	
5년 이상	2.17	1.38–3.42	1.74	1.05–2.91	2.44	1.18–5.07	3.04	1.64–5.66
교신/1저자 논문								
0편	1		1		1		1	
1편 이상	1.49	0.93–2.41	1.03	0.59–1.80	3.90	1.72–8.84	3.33	1.74–6.39
공저자 논문								
3편 미만	1		1		1		1	
3편 이상	2.01	1.20–3.36	1.77	0.89–3.50	3.13	1.43–6.84	2.92	1.55–5.48
2차 자료 이용 논문								
0편	1		1		1		1	
1편 이상	2.34	1.43–3.83	2.04	1.06–3.91	1.41	0.68–2.91	1.89	1.02–3.50
학술발표 횟수								
0편	1		1		1		1	
1회 이상	1.48	0.90–2.41	1.07	0.61–1.90	2.18	1.01–4.72	2.25	1.19–4.25
2차 자료 이용 학술발표								
0편	1		1		1		1	
1회 이상	2.71	1.53–4.77	1.69	0.81–3.52	1.81	0.83–3.94	2.53	1.31–4.91
2차 자료 이용 연구보고서								
0편	1		1		1		1	
1편 이상	2.61	1.51–4.49	1.56	0.77–3.16	3.22	1.49–6.96	4.46	2.34–8.49

OR: Odds Ratio, CI: Confidence Interval

All values were adjusted for age, gender, affiliation, and major filed.

* Controlled needs meant adjusted accessibility.

† Controlled trial meant adjusted reverse causality effect.

청구자료에 대한 접근이 어려워 연구 구상이 더 진행되지 않은 대상자를 접근도로 인한 요구도 감소로 정의하여 이를 요구도가 있는 것으로 간주한 조정된 청구자료 요구도(controlled needs)의 경우, 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서 연구

자의 역량요인 중 연구 경력이 5년 미만인 경우보다 5년 이상인 경우, 2차 자료 활용 논문이 없는 경우보다 1편 이상인 경우, 청구자료에 대한 조정 요구도가 높은 것으로 나타났다.

4. 연구자의 역량에 따른 자료 요청 시도

청구자료 요청 시도의 경우, 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서, 최종 학위가 박사 미만인 경우보다 박사 이상인 경우, 연구 경력이 5년 미만인 경우보다 5년 이상인 경우, 교신 또는 1저자 논문이 없는 경우보다 1편 이상인 경우, 공저자 논문이 3편 미만인 경우보다 3편 이상인 경우, 국내외 학술발표 횟수가 없는 경우보다 1회 이상인 경우, 2차 자료 활용 연구보고서가 없는 경우보다 1편 이상인 경우, 자료 구득 시도를 많이 한 것으로 나타났다.

표 4. 자료 구득 여부, 자료의 이해 정도와 공용데이터 사용 의사 여부에 대한 회귀분석

Variables	Taking		Strict taking*		Understanding		Open data set	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
최종 학위								
박사 미만	1		1		1		1	
박사 이상	0.96	0.18–4.95	3.02	0.38–24.10	0.41	0.44–3.85	2.45	0.75–8.03
연구 경력								
5년 미만	1		1		1		1	
5년 이상	1.39	0.32–6.16	1.27	0.27–5.94	0.40	0.63–2.57	1.47	0.78–2.77
교신/1저자 논문								
0편	1		1		1		1	
1편 이상	2.66	0.53–13.39	2.95	0.47–18.42	0.71	0.12–4.36	1.65	0.77–3.55
공저자 논문								
3편 미만	1		1		1		1	
3편 이상	1.22	0.32–4.62	2.45	0.51–11.78	0.72	0.13–3.96	2.69	0.96–7.53
2차 자료 이용 논문								
0편	1		1		1		1	
1편 이상	12.04	2.07–70.14	12.62	2.10–75.70	1.08	0.19–6.29	3.08	1.13–8.39
학술 발표 횟수								
0편	1		1		1		1	
1회 이상	4.21	0.90–19.67	1.08	0.24–4.97	4.75	0.68–33.45	2.04	0.92–4.51
2차 자료 이용 학술발표								
0편	1		1		1		1	
1회 이상	4.12	0.82–20.82	3.45	0.67–17.84	0.85	0.16–4.61	3.33	0.95–11.63
2차 자료 이용 연구보고서								
0편	1		1		1		1	
1편 이상	12.77	2.04–79.81	4.46	0.82–24.28	6.65	0.73–60.49	1.74	0.62–4.86

OR: Odds Ratio, CI: Confidence Interval
All values were adjusted for age, gender, affiliation, and major filed.
* Strict taking exclude optional responses.

청구자료에 대한 접근이 어려워 연구 구상이 진행되지 않은 대상자를 이 자료를 필요로 하는 연구의 요구는 있으나 구득 시도를 하지 않은 것으로 간주하고, 이에 따라 광의로 정의한 조정 청구자료 요청 시도(controlled trial)의 경우, 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서 연구자의 역량요인 모두가 종속변수에 유의미한 영향을 미치는 것으로 나타났다. 즉, 최종 학위, 연구 경력, 교신 또는 1저자 논문, 공저자 논문, 2차 자료 활용 논문, 국내외 학술발표, 2차 자료 활용 학술발표, 2차 자료 활용 연구보고서 등의 전반적인 역량요인 변수가 종속변수인 조정 청구자료 요청시도(controlled trial)에 유의미한 영향을 미쳤다. 이를 보다 구체적으로 살펴보면, 최종 학위가 박사 미만인 경우보다 박사 이상인 경우, 연구 경력이 5년 미만인 경우보다 5년 이상인 경우, 교신 또는 1저자 논문이 없는 경우보다 1편 이상인 경우, 공저자 논문이 3편 미만인 경우보다 3편 이상인 경우, 2차 자료 활용 논문이 없는 경우보다 1편인 경우, 국내외 학술발표 횟수가 없는 경우보다 1회 이상인 경우, 2차 자료 활용 학술발표가 없는 경우보다 1회 이상인 경우, 2차 자료 활용 연구보고서가 없는 경우보다 1편 이상인 경우, 청구자료 구득 시도를 많이 한 것으로 나타났다. 즉, 8가지 연구역량 변수 모두에서 연구역량이 높은 경우 구득 시도를 많이 한 것으로 나타났다.

5. 연구자의 역량에 따른 자료 구득

연구자의 역량에 따른 청구자료 구득을 분석한 결과는 표 4와 같다. 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서, 연구자의 역량요인 중 2차 자료 활용 논문이 없는 경우보다 1편 이상인 경우, 2차 자료 활용 연구보고서가 없는 경우보다 1편 이상인 경우 자료 구득을 많이 한 것으로 나타났다. 그러나 2차 자료 연구실적 외 최종학위, 연구경력, 교신저자/제1저자 논문 출판실적, 3편 이상의 공저자 논문 출판실적, 학술발표 실적 등 다른 6가지 연구역량 기준에 따른 자료구득 여부는 유의한 관련성이 없게 나타났다.

청구자료를 요청할 때마다 항상 구득하였던 연구자들만 청구자료 '구득 성공'으로 협소하게 정의한 청구자료 구득(strict taking)의 경우, 성, 연령, 소속 기관, 전공 분야 등을 보정한 상태에서 연구자의 역량요인 중 2차 자료 활용 논문이 없는 경우보다 1편 이상인 경우 청구자료 구득을 많이 한 것으로 나타났다. 그러나 이 협소한 정의에 따른 자료구득의 경우도 2차 자료 연구실적 외 상기 6가지 다른 연구역량 기준에 따른 자료구득 여부에서는 유의한 관련성이 나타나지 않았다.

6. 자료에 대한 이해 정도와 공용데이터 셋 사용 의사

연구자의 역량에 따른 청구자료에 대한 이해 정도와 공용데이터 셋 사용 의사를 종속변수로 투입한 모형을 살펴본 결과, 성, 연령, 소속 기관, 전공 분야 등을 보정한 건강보

험 청구자료 특성에 대한 이해 정도는 연구자의 역량과 관련이 없는 것으로 나타났다. 공용데이터 셋 사용 의사는 연구역량이 높을수록 더 많은 것으로 나타났다. 이러한 결과를 성, 연령, 소속 기관, 전공 분야 등을 보정한 후에 살펴본 결과, 2차 자료 활용 논문이 1편 이상인 연구자에서 공용데이터 셋 사용 의사가 더 많은 것으로 나타났다.

IV. 고 찰

이 연구는 연구자의 청구자료에 대한 미충족 필요와 청구자료가 필요한 연구자에게 잘 제공되고 있는지를 파악하고자 하였다. 주요 결과로는 청구자료에 대한 연구자의 요구에 비해, 자료제공의 충족률이 충분하지 않았으며, 이는 연구 역량이 높은 연구자의 청구자료에 대한 요구도 충족률도 그렇지 않은 연구자에 비해 높지 않은 것으로 나타났다. 주요 결과와 그 의미를 하나씩 검토해 보면, 첫째, 연구자의 역량이 높을수록 청구자료에 대한 요구도가 높았다. 마찬가지로 청구자료의 접근성을 보정한 후 분석한 결과도 연구자의 역량이 높을수록 요구도가 높은 것으로 나타났다. 그러나 청구자료에 대한 요구도가 높아도 분석에 필요한 변수를 제외하고 제공하거나 분석자료의 장소 제한, 제공할 때까지 걸리는 오랜 시간 등으로 인하여 역량 있는 연구자들이 자료 구득을 포기하는 것으로 나타났다. 이는 건강보험 청구자료에 대한 접근성이 높아진다면 역량 있는 연구자의 연구가 양적으로나 질적으로 활성화될 수 있음을 의미한다.

둘째, 요구도와 마찬가지로 연구자의 역량이 높을수록 청구자료 구득 시도를 더 많이 하였다. 또한, 청구자료의 접근성 문제로 요구도가 없어진 대상자를 청구자료 구득을 시도하지 않은 대상자에 포함하여 보수적으로 분석한 결과에서도 역량 있는 연구자들이 자료를 구득하려는 시도도 더 높은 것으로 나타나, 청구자료에 대한 구득시도가 연구역량과 관련이 많다는 것이 비교적 일관되게 확인되었다. 그중에서도 박사학위 이상인 연구자와 2차 자료를 활용한 연구보고서 실적이 1편 이상인 연구자의 요구도 교차비 값이 더욱 큰 것으로 나타나 다른 방법으로 측정한 연구 역량기준에 비해 박사 이상 연구자나 2차 자료를 활용한 연구보고서 출판경험이 있는 연구자가 2차 자료형태인 진료 청구자료를 구득하려는 시도가 더 많은 것으로 나타났다.

셋째, 여러 가지 측면의 연구역량을 기준으로 한 분류에서, 2차 자료를 활용한 논문과 2차 자료를 이용한 연구용역보고서 실적이 있는 연구자의 경우에서는 청구자료 구득을 했던 경우가 더 많았던 것으로 나타났다. 이러한 결과는 청구자료 구득 여부가 요청할 때마다 달랐다고 응답한 연구자를 포함하지 않은 분석에서도 같은 결과를 보였는바, 연구자의 역량이 높다고 하여 자료를 반드시 구득할 수 있는 것은 아니지만, 연구용역에 의한 연구를 하는 경우 혹은 2차 자료 분석 경험이 있는 연구자들, 이 둘의 경우에서는 청구자

료의 접근이 비교적 수월한 것으로 해석된다. 또한, 별도의 표로 제시하지 않았으나, 상관관계 분석에서 연구자의 역량이 높으면 2차 자료를 활용한 연구를 많이 한다는 것을 유추할 수 있었다. 그러나 2차 자료를 활용한 연구보고서를 발행하는 연구자가 청구자료를 구득하였다 하여 반드시 구득 자료를 활용하여 논문을 작성하는지 여부에 대한 추가적인 분석은 표본 수 제한으로 이루어지지 못하였다. 이 결과에서 나타난 바와 같은 기존의 자료 제공환경은 정책적 수요가 있는 연구에 더 우호적이었다. 그러나 더욱 선진적인 연구 환경을 이루려면, 향후 정책적 수요와 다소 다른 학문적 수요 혹은 반대 견해를 검증하고자 하는 연구에 대해서도, 다양한 정책적, 학술적 연구 성과가 만들어질 수 있도록, 일정 수준의 기본적인 연구 역량이 갖추어져 있는 연구자가 기획하는 대부분의 연구에 대해 지금보다 더욱 원활하게 건강보험 청구자료가 제공될 필요가 있다.

다른 한편, 연구자의 역량이 높아도 청구자료를 파악하고 분석하는 데 어려움이 있는 것으로 나타나 자료제공 서비스 외에도 청구자료에 대한 설명이나 상담 서비스도 잘 이루어져야 함을 시사하고 있다. 미국의 ResDAC(Research Data Assistance Center)에서는 CMS(Centers for Medicare and Medicaid Services) 자료제공 이후에도 분석할 때 변수 설명 등을 유선으로 상담을 해주거나, 세미나나 워크숍을 하여 데이터 셋의 이해를 돕고 있다. 결국, 이러한 서비스의 결과는 CMS 자료의 접근성을 높이고 메디케어 및 메디케이드 수혜자와 건강관리 프로그램의 발전으로 이어지고 있다(Conway & VanLare, 2010 ; Rantz & Connolly, 2004). 또한, 청구자료 설명에 대한 적절한 서비스는 청구자료 사용의 질적 수준을 보장하고 효율성을 증대시키며 청구자료 요청 시 심의위원회의 요구사항에 맞는 연구계획서 준비를 가능하게 함으로써, 연구자들의 청구자료 구득 성공률을 높일 수 있다. 마지막으로 많은 연구자 특히, 2차 자료를 활용한 논문 실적이 많은 연구자에서 변수를 정기적으로 공용데이터 셋으로 만들어 간편한 절차를 통해 제공하는 Open data set 방식에 긍정적인 의견을 보였는바 한국형 ResDAC 개발 및 입원, 외래, 시계열 자료로 확대하는 등 장단기 서비스 전략이 절실히 필요하다. 이와 관련하여 미국의 HCUP(Healthcare Cost & Utilization Project) 자료는 크게 주 단위 자료 3개와 국가 단위 자료 3개로 총 6개의 데이터베이스로 구성하여 각각의 자료를 동의서 작성 후 홈페이지를 통해 구입이 가능하도록 구축되어 있다. 또한, 자료마다 제공 변수가 100개 이상이며 타 기관의 자료와도 연계 가능하다는 점에서 활용도가 높다고 할 수 있다(Steiner 등, 2002).

온라인 설문조사의 회신율은 연구 주관기관이나 설문조사 내용, 표본의 성격, 응답자에게 주는 이익의 정도, 독촉 방법에 따라 많은 차이를 보이겠지만 본 연구에서 온라인 설문에 관한 응답 양상을 살펴본바, 회신율에 있어서는 한국보건행정학회를 기준으로 사전예고 메일 발송 후 본 메일을 발송하였을 때 1일까지가 17.7%이었다. 본 연구에서의 회신율은 본 메일 발송 후 2일째부터는 학회별로 조사 시기가 겹치게 되어 정확한 회신율을 추정할 수 없었다. 이러한 온라인 설문조사 결과는 기존에 알려진 바와 같은 대표성

의 문제에도 불구하고(O'nell & Penrod, 2001), 본 연구에서는 전화 설문조사의 회신율과 큰 차이가 없음을 보여주고 있다. 실제로 Roster 등 (2004)과 Parks 등 (2006)은 응답자의 행동 예측 면에서 온라인 조사가 전화 조사보다 우수한 결과를 보이는 것으로 보고하기도 하였다. 또한, 온라인 설문조사와 우편 설문조사의 회신율을 비교한 Kaplowitz 등 (2004)이 제시한 바대로 본 연구에서는 사전예고 메일을 발송하였는데, 우편 혹은 전화설문 조사에 필적할 만한 응답수준으로 보인 결과를 고려해 볼 때, 이 예고 메일은 면접자가 직접 상대하지 않음에도 불구하고, 온라인 설문조사에 대해 응답자들이 응답을 보다 적극적으로 할 수 있도록 기여한 요인의 하나였을 것으로 추정된다. 본 연구의 온라인 조사결과 초기 2일간의 응답률이 매우 높았던 점을 고려해 볼 때, 온라인 설문조사를 이용한 단기간의 비용효과적인 서베이를 원하는 연구를 구상하는 연구자들은 가능한 금요일에 사전예고 메일을 발송하고 월요일에 본 설문 메일을 발송하는 것을 추천할 수 있을 것으로 보인다.

본 연구는 몇 가지 제한점을 갖고 있다. 첫 번째, 온라인 설문조사의 문제점으로 지적되고 있는 응답자 대표성의 문제와 선택적 참여(self-selection), 응답자의 부재(noncontact)와 설문조사 일반이 갖고 있는 거부(refusal) 등 무응답 오차(nonresponse error)에 의한 자료 바이어스 가능성이 있을 수 있다. 그러나 본 연구에서는 학회에 가입하지 않은 연구자가 그다지 많지 않다고 불 수 있다면, 일반인 대상 온라인 조사연구와는 달리 연구결과를 크게 왜곡할 정도의 대표성에 따르는 바이어스 문제가 발생하지는 않았을 것으로 추정한다. 또한, 이메일 계정이 있는 보건의료부문 연구자를 대상으로 분석하였으므로, 특정 연구자 계층이 빠질 가능성이 크다고 생각하기는 어려울 것으로 보인다. 두 번째는 본 연구에서는 청구자료의 요구도, 요청 여부, 구득 여부와 마찬가지로 청구자료를 활용한 분석 경험이나 논문 및 연구보고서의 공표 여부 등은 설문도구에 포함되어 있었고, 조건부 확률로 분석이 계획되어 있었으나, 두 변수의 표본 수가 제한되어 회귀분석에 따른 평가가 이루어지지 못했다. 세 번째는 면접자가 없는 온라인 설문조사 방법을 이용하였기에, 한 해 동안 청구자료를 필요로 하는 연구의 구상이 1회 이상인 연구자들이 있었을 것임에도 불구하고, 구상횟수단위로 반복되는 루프형 설문방식으로 접근할 시, 응답자의 중도포기 가능성이 크다고 예측하여, 연구구상마다 그것을 단위로 반복 측정하지 못하고 한해의 경험을 하나의 변수로 포괄하는 질문으로 연구를 진행하였기에 자료의 병합(aggregation)에 따른 바이어스(bias)가 있을 수 있다. 네 번째로는 정책 연구자 커뮤니티만을 대상으로 한 조사여서, 익명성 보장이 불충분할 것으로 판단한 응답자들의 부담감이 작동한 것으로 보인다. 이로 인한 것으로 추정되는, 인구사회학적 변수에서만의 선택적 무응답을 한 응답자가 많았다. 인구사회학적 변수 무응답을 하나의 더미 변수로 그 특징을 모델에 고려하였으나, 그럼에도 불구하고 응답자의 일반특성이 다 조사된 경우보다 바이어스 가능성이 있다. 마지막으로 사전에 확정된 분석계획에 따라 분석을 시행하였으나 다중 검정(multiple testing)의 가능성을 배

제할 수 없으므로 연구결과의 해석에 주의를 요한다.

이런 몇 가지 제한점에도 불구하고 이 연구에서 밝혀진 연구자들의 청구자료에 대한 요구도와 여러 장애요인의 분석 결과는 향후 청구자료를 관리하는 기관인 심사평가원이 정책연구자에게 연구목적으로 필요한 청구자료를 더욱더 적극적으로 제공할 수 있는 근거가 될 것이다. 또 이는 곧 질적으로 향상된 연구의 증가로 이어지게 될 것이다. 이 연구결과를 계기로 해당 기관이 최근 보여주고 있는 연구수요에 부합한 보다 적극적인 자료제공 체계를 구축하기 위한 다양한 노력 (예를 들면, 자료 설명회 개최, 인력, 장비 보강 노력, 공용데이터 셋 제공을 위한 학회와의 협약 등)이 모쪼록 보건정책학계의 지지를 받아 보다 나은 정책 자료 환경이 가까운 시일 내 얻어질 수 있게 되길 기대한다.

감사의 글

본 연구의 설문 조사에 도움을 준 한국보건행정학회, 대한예방의학회, 한국의료QA학회 임원진, 사무국 및 회원분들께 깊은 감사를 드립니다. 온라인 설문조사의 장점과 한계에 관한 주요 자문과 이와 관련된 기존 연구를 검토하는데 도움을 주신 서울대학교 보건대학원 조병희 교수께 감사의 말씀을 드립니다.

참고문헌

국립암센터. 암 진료의 질 평가와 형평성 및 보장성 평가체계 개발 연구. 2010.

김재용. 근거기반 보건의료의 활성화를 위한 건강보험자료 활용: 현황과 과제. 대한예방의학회 동계 심포지움: 대한예방의학회; 2007. p. 1–28.

대한예방의학회. 진료정보의 적정활용 방안에 대한 연구; 2008.

Bethell C, Fiorillo J, Lansky D, Hendryx M, Knickman J. Online consumer surveys as a methodology for assessing the quality of the United States health care system. Journal of Medical Internet Research 2004;6(1):e2.

Conway P, VanLare J. Improving Access to Health Care Data. JAMA 2010;304(9):1007–1008.

Healthcare Cost and Utilization Project(HCUP). Data Use Agreement Training. Available at: http://www.hcup-us.ahrq.gov/tech_assist/dua.jsp (accessed at January 11, 2011)

Iezzoni L. Using Administrative Data to Study Persons with Disabilities. Milbank

Quarterly 2002;80(2):347-379.

Kaplowitz M, Hadlock T, Levine R. A comparison of web and mail survey response rates. Public opinion quarterly 2004;68(1):94-101.

O'NEIL K, Penrod S. Methodological variables in Web-based research that may affect results: Sample type, monetary incentives, and personal information. Behavior Research Methods, Instruments, & Computers 2001;33(2): 226-33.

Pandiani J, Banks S. Large data sets are powerful. Psychiatric Services 2003;54(5):745-746.

Paolo A, Bonaminio G, Gibson C, Partridge T, Kallail K. Response rate comparisons of e-mail-and mail-distributed student evaluations. Teaching and Learning in Medicine 2000;12(2):81-84.

Parks K, Pardi A, Bradizza C. Collecting data on alcohol use and alcohol-related victimization: A comparison of telephone and Web-based survey methods. Journal of Studies on Alcohol and Drugs 2006;67(2):318-323.

Rantz M, Connolly R. Measuring nursing care quality and using large data sets in nonacute care settings: State of the science. Nursing Outlook 2004;52 (1):23-37.

Roos L, Menec V, Currie R. Policy analysis in an information-rich environment. Social Science & Medicine 2004;58(11):2231-2241.

Roster C, Rogers R, Albaum G, Klein D. A comparison of response characteristics from web and telephone surveys. International Journal of Market Research 2004;46:359-374.

Sheehan K, McMillan S. Response variation in e-mail surveys: An exploration. Journal of Advertising Research 1999;39(4):45-54.

Steiner C, Elixhauser A, Schnaier J. The healthcare cost and utilization project: an overview. Effective clinical practice 2002;5(3):143-151.

Weible R, Wallace J. The impact of the Internet on data collection. Marketing Research 1998;10(3):19-23.

Zhan C, Miller M. Administrative data based patient safety research: a critical review. British Medical Journal 2003;12(Suppl 2).