

의료보험자료 상병기호의 정확도 추정 및 관련 특성 분석 -법정전염병을 중심으로-

신의철¹, 박용문¹, 박용규², 김병성³, 박기동⁴, 맹광호¹

가톨릭대학교 의과대학 예방의학교실¹, 의학통계학교실²,
경희대학교 의과대학 가정의학교실³, 보건복지부 보험국 보험관리과⁴

= Abstract =

Estimation of Disease Code Accuracy of National Medical Insurance Data and the Related Factors

Euichul Shin¹, Yong-mun Park¹, Yong gyu Park², Byung Sung Kim³,
Kidong Park⁴, Kwang-ho Meng¹

*Department of Preventive Medicine¹, Department of Health Statistics², The Catholic University of Korea,
Department of Family Medicine, The Kyunghi University³,
Health Insurance Management Division, Ministry of Health & Welfare⁴*

This study was undertaken in order to estimate the accuracy of disease code of the Korean National Medical Insurance Data and disease the characteristics related to the accuracy. To accomplish these objectives, 2,431 cases coded as notifiable acute communicable diseases (NACD) were randomly selected from 1994 National Medical Insurance data file and family medicine specialists reviewed the medical records to confirm the diagnostic accuracy and investigate the related factors.

Major findings obtained from this study are as follows :

1. The accuracy rate of disease code of NACD in National Medical Insurance data was very low, 10.1% (95% C.I. : 8.8-11.4).
2. The reasons of inaccuracy in disease code were 1) claiming process related administrative error by physician and non-physician personnel in medical institutions (47.0%), 2) input error of claims data by key punchers of National Medical Insurer (31.3%) and 3) diagnostic error by physicians (21.7%).
3. Characteristics significantly related with lowering the accuracy of disease code were location and level of the medical institutions in multiple logistic regression anal-

ysis. Medical institutions in Seoul showed lower accuracy than those in Kyonggi, and so did general hospitals, hospitals and clinics than tertiary hospitals. Physician related characteristics significantly lowering disease code accuracy of insurance data were sex, age group and specialty. Male physicians showed significantly lower accuracy than female physicians; thirties and forties age group also showed significantly lower accuracy than twenties, and so did general physicians and other specialists than internal medicine/pediatric specialists.

This study strongly suggests that a series of policies like 1) establishment of peer review organization of National Medical Insurance data, 2) prompt nation-wide expansion of computerized claiming network of National Medical Insurance and 3) establishment and distribution of objective diagnostic criteria to physicians are necessary to set up a national disease surveillance system utilizing National Medical Insurance claims data.

Key words : Accuracy, National Medical Insurance data, Notifiable acute communicable disease

I. 서 론

질병 통계는 한 나라의 보건의료 정책을 수립하고 보건의료사업의 효과를 비교하고 평가하기 위해서 꼭 필요한 자료이다. 또한 최근에는, 지역보건법의 실시로 각 지방자치단체 스스로도 보건의료사업을 기획·실시하고 평가하게 요구됨으로써 지역주민의 질병상태를 파악할 수 있는 통계의 필요성이 더욱 커졌다(보건복지부, 1997).

이렇게 질병양상을 파악하기 위한 자료에는 법정질병 신고자료, 생생통계자료, 등록(registry)자료, 각종 조사자료 그리고 보험자료와 같은 행정자료 등이 있으며(Teutsch & Churchill, 1994) 그 중에 의료보험자료는 우리나라에서 가장 빈번히 사용되는 질병통계자료 중의 하나이다. 그외에도 의료보험자료는 보건 의료정책을 수립하고 그 효과를 평가하기 위해 빈번히 사용되어 왔으며 그 예로는 포괄수가제도, 의료이용과 관련된 연구(유승흠 등, 1988) 등이 있으며 국가가 국민보건을 증진해야 하는 의무가 강조되는 요즘에는 그 유용성이 더욱 증가하고 있다.

우리나라에서 의료보험자료가 보건의료연구의 기초자료로 빈번히 쓰이는 이유는 전 국민에 대한 의료

이용정보를 담고 있다는 점, 연구자의 별도의 정보수집 노력 없이도 얻을 수 있는 기구축 자료라는 점, 그리고 자료의 전산화 과정이 필요 없다는 점등으로 사료된다.

그러나, 의료보험자료의 사용에는 제한적인 면이 있다. 그 중 가장 문제되는 것은 자료의 정확성 여부이며, 특히 최근에 의료보험자료 상병기호의 정확성에 관한 문제가 제기되고 있다. 그러나 이에 대해서는 정확한 실태와 그 원인에 대한 연구가 미비한 실정이다.

이에 저자들은 법정전염병을 대상으로 질병통계자료로서의 우리나라 의료보험자료에 대해

- (1) 상병기호의 정확도를 추정해보고,
- (2) 상병기호가 부정확한 경우, 그 사유를 알아보며
- (3) 상병기호의 정확도에 영향을 미치는 특성에는 어떤 것들이 있는지를 알아보기 위해 이 연구를 실시하였다.

II. 연구자료 및 방법

1. 연구자료

연구자료로 1994년 한 해 동안 의료기관이 보험급

여를 위해 청구한 환례들에 대해 의료보험자가 재입력한 자료 중 주상병 혹은 부상병 기호가 법정전염병인 전산화일을 이용하였다. 연구대상 질환을 법정전염병으로 한 것은 진단에 객관적인 기준이 정의되어 있어(CDC, 1990) 연구자가 의료보험자료의 상병기호가 정확한지를 알아보기 위해서 의료기관에서 해당 환례를 객관적으로 검토할 수 있기 때문이었다.

제1, 2종 법정전염병 중 상병기호를 정확히 분류할 수 없는 후천성면역결핍증을 제외하고, 1994년 한해 동안 전국의 의료기관에서 의료보험자에게 보험급여를 청구한 자료는 모두 372,151건으로 이들의 종별 질병별 분포는 표1과 같다.

표 1. 제1, 2종 법정전염병 의료보험자료의 종별 질병별 분포, 1994년

종별	질병명	건수(%)	질병명	건수(%)
제1종 법정전염병				
	디프테리아	833 (0.2)	발진티푸스	4,299 (1.2)
	세균성이질	33,235 (8.9)	장티푸스	16,109 (4.3)
	콜레라	11,130 (3.0)	파라티푸스	1,119 (0.3)
	페스트	1,807 (0.5)	황 열	3,361 (0.9)
제2종 법정전염병				
	공수병	6,966 (1.9)	렙토스피라증	5,398 (1.5)
	말라리아	4,428 (1.2)	발진열	836 (0.2)
	백일해	17,445 (4.7)	성홍열	11,160 (3.0)
	수막구균성수막염	1,682 (0.5)	아메바성이질	8,468 (2.3)
	유행성이하선염	100,083 (26.9)	유행성출혈열	1,268 (0.3)
	일본뇌염	2,428 (0.7)	채귀열	666 (0.2)
	쯔쯔가무시병	5,972 (1.6)	파상풍	7,042 (1.9)
	폴리오	3,928 (1.1)	홍 열	122,488 (32.9)
계				372,151 (100.0)

2. 연구방법

이 연구는 크게 조사대상 선정, 선정된 조사대상 환례에 대한 의료기관 방문 및 의무기록 검토, 의료보험자료 상병기호에 대한 정확도 추정, 그리고 부정확 사유 및 관련 특성 분석 등으로 구분된다.

1) 조사대상 선정(표본 추출)

조사대상은 총 372,151건 연구자료 중에서 연구자가 직접 의료기관을 방문하여 의무기록을 검토함으로써 의료보험자료의 정확성을 확인할 수 있도록 연구자 소재기관과 가까운 서울, 경기지역의 의료기관으로 국한하였고 진단의 객관성이 비교적 떨어지는 홍역, 유행성이하선염을 제외하였다. 이러한 작업으로 선택된 자료는 41,862건이며 이를 지역별, 의료기관 등급별로 층화하고 그 분포비에 따라 조사대상을 표본추출하였다. 이때 의료기관에서 확인해야 할 대상 환례수를 의료기관 한 곳 당 최소한 40건, 그리고 최대 50건이 되도록 하여, 최종적으로 의료기관 50곳으로부터 2,431건의 조사대상을 무작위 추출하였다. 조사대상 의료기관과 의무기록의 지역별 의료기관등급별 분포는 표2와 같다.

표 2. 조사대상 의료기관과 의무기록의 지역별 의료기관등급별 분포

지 역	서 울		경 기		계	
	의료기관등급	의무기록수	의료기관등급	의무기록수	의료기관등급	의무기록수
의원급	16	783	9	424	25	1,207
병원급	3	146	1	50	4	196
종합병원급	6	284	4	194	10	478
3차병원급	10	500	1	50	17	550
계	35	1,713	15	718	50	2,431

2) 의료기관 방문 및 의무기록 검토

의료보험자료 상병기호의 진위를 파악하기 위해 의료기관을 방문, 의무기록을 검토하였으며 그를 위해서는 법정전염병 진단의 객관적 기준을 정의할 필요가 있었는데, 이를 위해 미국 질병관리센터(Centers for Disease Control and Prevention, 1990) 의 질병감시를 위한 증례정의를 바탕으로 하여 우리나라의 제1, 2종 법정전염병에 해당되는 부분을 발췌하고, 내과 및 가정의학과 전문의의 도움을 받아 한글로 알맞게 재편집하였다.

이를 근거로 조사대상 의무기록 2,431건에 대해 가정의학 전문의 12인으로 구성된 의무기록 검토팀이 해당 의료기관을 방문하여 관련 기록을 검토하였다. 이때 보건복지부 방역과, 서울특별시청 의약과, 경기도청 보건과 그리고 관할보건소를 통해 조사대상 의료기관에 협조 지원을 요청하였으며 이러한 협조관계의 구축은 조사에 절대적인 도움이 되었다.

3) 정확도 추정

의료보험자료 상병기호의 정확도는 조사대상 자료 중 의무기록의 판별이 가능한 것 중에 실제 진단명이 법정전염병으로 의료보험자료와 상병기호가 일치하는 경우의 비율을 구하였다.

이는 식1과 같으며, 이렇게 산출한 의료보험자료 상병기호의 정확도에 대한 신뢰도는 이항분포에 의한 확률의 95% 신뢰구간으로 나타내었다(Cochran, 1977).

$$\frac{\text{실제 법정전염병 건수}}{\text{법정전염병 상병기호의 의료보험자료 중 의무기록 판별가능 건수}} \times 100 \dots \text{식1}$$

4) 부정확 사유조사

의무기록 검토결과 의료보험자료의 상병기호가 부정확한 경우에 대해 의무기록 검토자가 그 사유를 면담조사하였다. 사유항목은 크게 진단오류와 행정오류로 구분하였는데 진단오류는 의사의 진단과정 중에 발생하는 것이고, 행정오류는 진단과정 이후 상병기호를 기재·입력하거나 이에 대한 보험급여를 청구하는 과정에서 발생하는 일반 사무와 관련된 것으로 정의하였다. 따라서 행정오류는 의사, 의료기관내 의사 외 관련 직원, 그리고 의료보험자에 의해 발생할 수 있었다.

조사대상의 상병기호가 부정확한 사유의 해당 항목이 두 가지 이상인 경우는 중복표시 하도록 하였다.

5) 부정확 관련특성 분석

의료보험자료의 상병기호가 부정확한데에는 어떠

한 특성들이 관련되는지 알아보기 위해 의료기관 관련 특성과 의사 관련 특성들을 알아보고, 이들 특성들에 대해 단변량 분석과 다변량 로지스틱 회귀분석(multiple logistic regression analysis)을 SAS(SAS Institute Inc., 1994)를 이용하여 실시하였다.

이때 의료기관 관련 특성 변수는 의료기관의 소재 지역(서울/경기), 등급(의원/병원·종합병원/3차병원) 그리고 소유주(공공/법인/개인)의 정보를 의료보험 자료 중 요양기관 화일로부터 알아보았다. 의사관련 특성 변수는 의사의 성, 연령군(20대/30·40대/50대 이상) 그리고 전공과목(내과·소아과/기타 전공과/일반의)의 정보를 의료기관 방문조사시 파악하였다.

III. 연구결과

1. 의료보험자료 상병기호의 정확도 추정

조사대상 자료 2,431건에 대해 의료기관을 직접 방문한 결과 전체 50개 의료기관 중 병원급 의료기관 1곳과 3차병원급 의료기관 1곳을 제외한 총 48개 의료기관에 대해 방문 조사가 가능하였다. 2개 의료기관에 대해 방문조사를 실시하지 못한 것은 해당 의료기관이 업무전산화를 진행 중에 있어서 의무기록 열람이 불가능했거나, 의무기록 검토에 비협조적이었기 때문이었다.

조사가능 48개 의료기관의 의무기록 2,339건에 대해 의무기록 검토팀이 의료기관을 직접 방문하여 조사한 결과는 표 3과 같다.

표3에서 보듯이 조사대상 자료 총 2,339건 중 의료기관에 해당 의무기록이 없는 경우가 150건(6.4%)이었다. 나머지 의무기록이 있었던 2,169건에 대해 의무기록을 검토하였는데 그 내용을 판독할 수 없는 경우는 20건(0.9%)이었다. 결국 전체 조사대상 자료의 7%인 2,169건에 대해서만 의무기록 내용을 판별하여 의료보험자료 상병기호의 정확도를 판단할 수 있었다.

판별가능 의무기록 2,169건에 대해 검토한 결과, 미 국질병관리센타의 진단기준에 부합되는 경우는 단지

표 3. 조사대상자료의 의무기록 검토 결과

의무기록 존재 판별	의료보험자료		건수	(%)	(%)
	실진단명	상병기호			
있음	가능	법정전염병	정확	218	(10.1)
	불가능	비법정전염병	부정확	1,951	(89.9)
소 계			2,169	(100.0)	(92.7)
없음	불가능		20	(0.9)	
총계			150	(6.4)	
총계			2,339		(100.0)

218건으로, 판별가능 의무기록 중 실제 법정전염병 건수의 비율(식 1 적용)인 의료보험자료 상병기호의 정확도는 10.1%, 95% 신뢰구간은 8.8-11.4%였다.

2. 의료보험자료 상병기호의 부정확 사유

의무기록 검토 결과 의료보험 자료의 상병기호가 부정확한 것으로 확인된 자료 1,951건에 대해 그 사유를 조사한 결과는 표4와 같다.

표4에서 보듯이 의료보험자료의 상병기호가 부정확한 사유는 진단오류가 423건(21.7%), 행정오류가 1,528건(78.3%)을 차지했다.

표 4. 의료보험자료 상병기호의 부정확 사유

항 목	정 의	건수	(%)
진단오류	의사: 진단오류	423	(21.7)
행정오류	의사: 기재오류	0	(0.0)
	비의사: 의료기관 혹은 의료보험자의 기재 및 전산 입력오류	1,528	(78.3)
계		1,951	(100.0)

표5는 행정오류 1,528건에 대해 그 사유를 조사한 결과이다. 조사당시 의무기록 검토자가 중복 표시를 하도록 하였다. 표에서 보듯이 행정오류의 사유 중 가장 빈번한 것은 비의사의 '급여심사 기준에 맞추려' 항목으로 1,097건(71.8%)이었다. 그 다음도 역시 비의사의 행정오류로 '상병기호를 기재·전산 입력할 때

표 5. 행정오류의 사유별 분포(중복 표시 허용)

항 목	내 역	건수	(%)
의사의 진단에 자신 없어서		24	(1.6)
오 류	진단명을 기재하지 않아서	53	(3.5)
	배제진단 코드를 사용해서	333	(21.8)
	보험급여 심사 기준에 맞추기 위해	30	(2.0)
	잘못된 코딩 습관 때문에(코딩에 무관심)	146	(9.6)
	기타	110	(7.2)
비의사	진료의사가 기재한 진단명을 잘못 인식해서	108	(7.1)
오 류	의무기록지 내용을 잘못 인식해서	17	(1.1)
	급여 심사기준에 맞추려(관련 검사지 등)	1,097	(71.8)
	상병기호의 단순 기재 혹은 전산 입력을 잘못해서	770	(50.4)
	기타	53	(3.5)
계		1,528	(100.0)

의 단순오류'였으며 770건(50.4%)을 차지했다. 의사의 행정오류로서 '배제진단 코드를 사용해서'는 333건(21.8%), '코딩에 무관심해서 등과 같이 잘못된 코딩 습관 때문에'는 146건(9.6%)로 그 다음 순위를 차지했다.

비의사의 행정오류 사유 중에서 상병기호의 단순 기재·전산 입력 오류는 의료기관의 관련 직원이나 의료보험자의 전산입력담당자에 의해 일어날 수 있었다. 따라서 이 연구에서 이러한 기재·전산 입력오류 770건에 대해 오류의 발생장소를 조사하기 위해 해당 환례의 의료보험 청구명세서와 의료보험자료의 상병기호를 비교하였으며 그 결과는 표 6과 같다.

표 6. 의료기관 청구명세서와 의료보험자료의 상병기호 비교 결과

청구명세서 상병기호	의 미	건수	(%)	(%)
확인여부				
가 능	일 치	의료기관 기재·	105	(21.3)
		전산 입력 오류		
	불일치	의료보험자 전산	387	(78.7)
		입력 오류		
계			492	(100.0) (63.9)
불가능			278	(36.1)
계			770	(100.0)

표6에서 보듯이 단순한 기재·전산 입력 오류에 의한 행정오류 770건 중에서 해당 의료기관에서 청구명세서 확인이 가능한 것은 492건(63.9%)이었으며, 이 중에서 의료기관 청구명세서의 상병기호와 의료보험자료의 상병기호가 서로 일치하여 실진단명에 대한 상병기호를 의료기관에서 잘못 기재·입력한 것으로 판단된 경우는 105건(21.3%)이었으며, 상병기호가 서로 일치하지 않아 의료보험자의 입력오류로 판단된 경우는 387건(78.7%)이었다. 이 결과로 미루어 보아 의료보험자료의 상병기호가 부정확한 경우, 의료보험자의 입력오류율의 추정치는 식(2)에 의해서 31.1%인 것으로 산출된다.

$$\begin{aligned} & \text{부정확 의료보험자료 상병기호 중 보험자 입력오류율} = \\ & \text{행정오류율} \times \text{기재} \cdot \text{입력오류율} \times \text{청구명세서 불일치율} \\ & \dots\dots\dots \text{식 2} \end{aligned}$$

3. 의료보험자료 상병기호의 정확도 관련특성

의료보험자료의 상병기호가 정확한지 그렇지 아니한지에 영향을 주는 특성들에 대해 알아보기 위해서 의료기관과 의사특성 변수들을 가지고 의료보험자료 상병기호의 정확도에 대해 단변량 및 다변량 로지스틱 분석을 실시하였다. 분석 결과 각 특성 변수의 odds ratio와 95% 신뢰구간은 표7과 같다.

이 표에서 보듯이 우선 의료기관 특성의 의료보험자료 상병기호의 정확도에 대한 영향을 보면 의료보험자료 상병기호의 정확도는 의료기관의 소재지역이 서울인 경우가 경기지역인 경우에 비해 1.5배 유의하게 낮았다. 의료기관 등급별로는 3차병원급 의료기관에 비해서 병원·종합병원급이 8.8배, 의원급 의료기관이 3.5배 의료보험자료 상병기호의 정확도가 유의

표 7. 의료보험자료 상병기호의 정확도에 대한 관련 특성들의 odds ratio 및 95% 신뢰구간

특성	정 확 도		단변량 분석	다변량 분석
	부정확(%)	정확	O.R.(95% C.I.)	O.R.(95% C.I.)
의료기관				
지역	경기+	86	1.0	1.0
	서울	132	1.5*(1.1-2.0)	4.3*(2.9-6.4)
등급	3차병원+	104	1.0	1.0
	병원·종합병원	19	8.8*(5.3-14.6)	13.3*(7.5-23.5)
	의원	95	3.5*(2.6-4.7)	5.0*(3.0-8.6)
소유주	공공+	93	1.0	
	법인	28	2.8*(1.8-4.3)	
	개인	97	3.2*(2.3-4.3)	
의사				
성	여자+	104	1.0	1.0
	남자	110	3.6*(2.7-4.8)	4.9*(3.4-6.9)
연령군	20대+	82	1.0	1.0
	30,40대	76	6.8*(4.8-9.6)	3.0*(1.8-4.8)
	50대이상	49	5.6*(3.8-8.2)	1.1 (0.6-2.0)
전공과목	내과, 소아과+	186	1.0	1.0
	기타 전공과	27	2.4*(1.6-3.7)	1.8*(1.1-2.9)
	일반	3	4.3*(1.4-13.8)	4.0*(1.2-13.5)

* 유의수준 0.05에서 통계적으로 유의

+ 기준치

O.R. odds ratio

C.I. confidence interval

하게 낮았다. 의료기관 소유주별로는 공공(공립 및 학교법인)인 경우에 비해서 기타법인이나 개인 소유의 의료기관의 경우가 보험자료 상병기호의 정확도가 각각 2.8배, 3.2배 유의하게 낮았다.

의사의 특성에 의한 의료보험자료 상병기호의 정확도 변화를 살펴보면, 남자의사인 경우가 3.6배 유의하게 낮았으며, 연령군은 20대 의사인 경우에 비해 30, 40대 의사가 3.0배 유의하게 낮았다. 50대 이상의 연령군에서도 20대 의사에 비해 의료보험자료 상병기호의 정확도가 5.6배 유의하게 낮았다. 전공과목의 영향은 내과·소아과 전공인 경우에 비해 기타 과목 전공인 경우가 2.4배, 전공과목이 없는 일반의사인 경우도 4.3배 유의하게 의료보험자료 상병기호의 정확도가 저조하였다.

한편 이 표에서는 의료보험자료 상병기호의 정확도에 영향을 미치는 특성들에 대한 다변량 분석(stepwise 다변량 로지스틱 분석) 결과를 각 특성 변수들의 odds ratio 와 95% 신뢰구간으로 나타내 주고 있다.

우선 의료기관 특성에 대해 살펴볼 경우, 의료보험자료 상병기호의 정확도는 경기지역보다 서울지역 의료기관의 경우 4.3배 유의하게 낮았으며, 의료기관 등급별로는 3차병원에 비해 병원·종합병원급 그리고 의원급 의료기관의 정확도가 각각 13.3배, 5.0배 유의하게 낮았다.

의사의 특성으로는 남자의사의 경우가 여자의사보다 의료보험자료 상병기호의 정확도가 4.9배 유의하게 낮았으며, 20대에 비해 30, 40대 연령군의 경우가 3.0배 유의하게 낮았다. 50대 이상의 의사인 경우는 20대 의사인 경우와 비교하여 의료보험자료 상병기호의 정확도가 유사한 수준이었다. 전공과목별로는 내과·소아과 의사에 비해 기타 관련 없는 전공과 의사인 경우에 의료보험자료 상병기호가 1.8배 유의하게 부정확했으며 전공과목이 없는 일반의사인 경우에는 의료보험자료 상병기호가 4.0배 유의하게 부정확했다.

IV. 고 찰

이 연구는 의료보험자료 상병기호의 정확도를 추정해보고 그와 관련된 특성에는 어떤 것들이 있는지를 살펴보는 데 일차적인 목적이 있다. 우선 이 연구결과 의료보험자료 상병기호의 정확도는 10.1%로 매우 저조한 것으로 나타났는데, 이는 암 환자조사(보건복지부, 1996)에서 나타난 의료보험자료 상병기호의 정확도 35.7%와 비교해 보면 다소 차이는 있으나 의료보험자료 상병기호의 정확도가 낮기는 마찬가지이다. 이 연구결과를 외국의 것과 비교해 볼 경우 미국 미네아폴리스주의 한 의료보험자 청구자료 중 진단코드가 고혈압인 경우에 대해 의무기록의 진단명을 조사해 본 결과 일치율은 74%였으며, 진단코드가 고혈압인 경우 중에서도 혈압강화제를 사용한 경우에 대해 의무기록 진단명과 일치율은 97%로 매우 높아(Quam et al, 1993) 우리나라 의료보험자료의 진단코드 정확도는 재고의 여지가 매우 절박한 것으로 사료된다.

이 연구에서 법정전염병에서 의료보험자료 상병기호의 정확도가 좀더 낮은 것은 대상질병과 그 질병을 보는 의료기관 등급, 의사의 전문성 및 관심도 차이 때문인 것으로 생각된다. 즉, 법정전염병은 의사가 신고자체를 꺼리고 따라서 청구도 다른 질병코드로 한 다거나, 암은 그 진료를 종합병원급 이상의 의료기관 그리고 전문성과 관심이 높은 분과전문의가 주로 담당함으로써 질병코드를 좀더 정확히 하고자 하는 노력이 있었던 것으로 생각된다.

연구의 제한점으로는 의료보험자료 상병기호의 정확도에 영향을 미치는 특성에는 어떤 것들이 있는가를 알아볼 때, 관련 특성변수로서 의료기관과 의사의 특성만을 고려하였을 뿐, 질병의 특성과 상병기호의 입력에 관련된 의료보험자의 특성을 고려하지 못한 점이 있다. 그 이유들로서 질병의 특성을 고려치 못한 것은 의료보험자료 상병기호의 정확도가 이 연구대상 자체를 정확성 판단의 객관성을 높이기 위해 법정전염병으로 국한하였기 때문이었으며, 의료보험자의 특

성을 고려치 못한 것은 연구계획 수립과 의료기관 방문조사 단계에서 의료보험자의 입력오류에 의한 의료보험자료 상병기호의 부정확성을 전혀 예상치 못한 결과이다. 따라서 의료보험자료 진단코드 정확성의 관련특성분석을 위한 다변량분석에서 제외변수오류가 있을 수 있다. 이 연구의 다른 제한점으로는 조사대상 자료를 표본추출할 때 의원급과 병원급 이상 의료기관의 추출비를 1:1로 균등히 함으로서 의료보험자료 상병기호의 정확도에 대한 특성 분석결과가 병원급 이상 의료기관의 특성이 많은 것으로 편향되었을 가능성이 있다.

V. 요약 및 결론

이 연구를 종합해보면,

첫째, 1994년도 의료보험청구 재입력 자료 중 무작위 표본 추출한 제1,2종 법정전염병 2,431건에 대한 의무기록을 검토, 상병기호의 진위를 확인해 본 결과 정확도는 10.1%(95% C.I.: 8.8-11.4)에 불과하였다.

둘째, 상병기호가 부정확했던 경우에 대해 그 사유를 조사해 본 결과 가장 큰 부분은 의료기관 내 의사·비의사의 행정오류로서 47.0%를 차지했으며 그 다음은 의료보험자의 입력오류로 31.3%였다. 나머지는 21.7%는 의사의 진단오류에 의한 것으로 추산되었다.

셋째, 상병기호의 부정확성과 관련된 특성을 단변량 및 다변량 로지스틱 회귀분석으로 조사해 보았는데, 의료기관 관련 특성으로는 소재지역이 서울인 경우가 경기지역보다 정확도가 4.3배(95% C.I.: 2.9-6.4) 유의하게 낮았으며, 의료기관 등급은 3차병원에 비해 병원·종합병원이 13.3배(9.5% C.I.: 7.5-23.5), 의원이 5.0배(95% C.I.: 3.0-8.6) 유의하게 정확도가 낮았다. 상병기호 부정확성의 의사관련 특성으로는 성별상 남자 의사인 경우가 여자 의사보다 4.9배(95% C.I.: 3.4-6.9) 부정확했으며, 연령군에 의한 차이는 20대에 비해 30, 40대인 경우가 3.0배(1.8-4.8) 유의하게 부정확하였다. 전문과목별로는 내과·소아과 전공에 비해

기타 전공과는 1.8배(1.1-2.9), 일반의사인 경우는 4.0배(95% C.I.: 1.2-13.5) 유의하게 부정확하였다.

이상의 결과를 바탕으로 연구자들은 의료보험자료의 정확성을 향상시켜 질병양상의 파악에 이용하고 더 나아가 우리나라 보건의료정책의 수립·평가에 활용하기 위해서 다음과 같은 정책을 강력히 추진할 것을 제안한다.

1. 의료보험자료 상병기호 검증기구 설립

이 연구결과 의료보험자료 상병기호가 부정확한 이유의 약 70%는 의료기관에서 일어나는 것으로 의사의 진단오류 그리고 진료비 청구작업과 관련된 여러 행정절차에 의한 것이었다. 따라서 이에 대한 개선방안을 마련하기 위하여 우선 의료기관을 방문하여 의무기록 검토를 포함한 조사를 할 수 있는 전문인력을 확보하고, 관련 오류를 개선토록 해당 의료기관을 조정하거나 오류를 방지할 수 있는 제도적 장치를 개발할 수 있는 기구를 설립해야 한다. 또한 이를 위하여 현행 의료보험연합회의 심사기능을 의사의 참여 폭을 늘리는 등 합리적으로 조정·강화하는 대안도 생각할 수 있다. 이와 유사한 사례로 미국에서는 의료의 질관리를 위해 의료이용감시를 하고 있는데 의료이용 및 의료의 질관리에 관한 동료심사위원회(Utilization and Quality Control Peer Review Organization)이 대표적인 예이다(Korda, 1994; Grant et al, 1996).

2. 현행 의료보험진료비 청구업무 전산화 사업의 신속한 확산

이 연구결과 의료보험자료 상병기호가 부정확한 사유 중 약 30% 정도는 의료보험자의 청구명세서 전산 입력 오류에 의한 것으로 추산되는데, 이는 의료기관의 청구업무를 전산화하는 경우, 개선이 가능하다. 다행히 의료보험연합회에서는 의료보험진료비의 청구와 심사업무의 효율성을 증진시키기 위하여 1994년 6월

부터 한국통신과 함께 '의료보험 EDI 청구사업'을 실시하여 의료기관에서의 전산화 청구자료가 재입력 과정 없이 직접 의료보험자료가 되도록 시도하였으며, 95년 12월에는 수도권 지역 100개 병·의원을 대상으로 시범사업을 실시한 이후, 1998년 2월 현재는 2,251개 요양기관이 참여하고 있다(의료보험연합회, 1998). 따라서 의료보험 진료비 EDI 청구사업의 대상기관을 전국적으로 신속히 확장하도록 정책적으로 배려한다면, 부정확한 의료보험자료 상병기호 중 의료보험자의 입력오류에 의한 부분은 완벽히 해소할 수 있다.

3. 진단기준안 정립 및 홍보

이 연구결과에서 의료보험자료 상병기호가 부정확한 경우의 약 21.7%는 의사의 진단오류에 의한 것이었다. 이 부분을 개선하기 위해서는 먼저 진단에 대한 객관적 기준을 마련해야 한다. 미국 질병관리센터에서는 법정전염병의 진단을 객관화하기 위해 증례정의를 마련하였으며 각 신고대상 전염병을 확인, 유사, 의심 진단으로 분류하고 각각에 대해 임상적, 검사실적 기준을 제시하였다(CDC, 1990). 또한 이렇게 마련한 객관적 증례정의를 의료인이 알고 사용할 수 있도록 교육, 홍보하는 일인데 이는 의과대학에서의 교육, 대중매체를 통한 홍보, 그리고 각종 관련 학회의 연구교육을 통해 지속적으로 추진해야 한다.

4. 의료보험자료를 이용한 질병정보체계 (surveillance system) 구축

이상에서 제시한 것과 같은 과정들을 통해서 의료기관은 보다 객관적으로 진단을 내리고, 관련 정보를 전산화하여 의료보험 청구를 하며, 의료보험자는 수집된 정보가 정확한지 여부에 대해 전문가 집단의 검증 기구를 통해 확인하며, 이렇게 정확성이 검증된 질병정보를 가지고 보건당국은 국민이나 지역사회 주인의 질병상태를 지속적으로 파악하고 이에 적절한 보건정책을 수립·집행·평가하는 이른바 '의료보험자료를

이용한 질병정보체계(surveillance system)'를 구축할 수 있으리라 생각된다.

감사의 글

이 연구의 수행에 많은 도움을 주신 의료보험연합회 및 관리공단, 서울특별시청 의약과, 경기도청 보건과 그리고 관련 보건소에 근무하시는 관계자 여러분께 감사드립니다. 또한 50여 곳 의료기관을 방문하여 직접 의무기록을 검토하고 조사하는데 힘써주신 김용철, 노용균, 박노원, 송상욱, 신호철, 양윤준, 오정열, 이원근, 이 행, 최윤선, 최환석 등 11명의 가정의학과 선생님께도 깊은 감사의 말씀을 드립니다.

참고문헌

- 보건복지부. 1992년도 암환자 조사보고서. 1996
- 보건복지부. 보건·의약 관계법규, 지역보건법. 서울, 계축문화사, 1997, 쪽 281-298
- 유승흠, 조우현, 손명세, 박종연. 지역의료보험 가입자의 외래 이용 변화. 예방의학회지 1988; 21(2): 419-430
- 의료보험연합회. 의료보험 EDI 청구 추진현황. 1997
- Centers for Disease Control. Case definitions for public health surveillance. MMWR 1990; 39: RR-13
- Cochran, WG. Sampling techniques, 3rd Ed. New York, John Wiley & Sons Inc., 1977
- Grant JB, Hayes RP, Pates RD, Elwart KS, Ballard DJ. HCFA's health care quality improvement program: the medical a Informatics challenge. Journal of the American Medical Informatics Association 1996; 3(1): 15-26
- Korda H. Utilization review for Medicaid diagnosis-related group system: practice, innovation, and lessons of experience. American Journal of Medical Quality 1994; 9(2): 54-67
- Quam L, Ellis LB, Venus P, Clouse J, Taylor CG, Leat herman S. Using claims data for epidemiological research. The concordance of claims-based criteria with medical record and patient survey for identifying a hypertensive population. Medical Care 1993; 31(6): 498-507

SAS Institute Inc. SAS/STAT user's guide, Release
version 6, 4th ed. 1994
Teutsch SM, Churchill RE. Sources in routinely col-

lected data for surveillance. Chap3 in Princiles
and Practice of Public Health Surveillance. New
York, Oxford University Press, 1994, pp31-85.
