

# 면접조사자료와 사망등록자료 간 교육수준 및 직업계층의 신뢰도

김혜련, 강영호<sup>1)</sup>

한국보건사회연구원, 울산대학교 의과대학 예방의학교실<sup>2)</sup>

## Reliability of Education and Occupational Class: A Comparison of Health Survey and Death Certificate Data

Hye-Ryun Kim, Young-Ho Khang<sup>2)</sup>

Korea Institute for Health and Social Affairs, Department of Preventive Medicine, University of Ulsan College of Medicine<sup>2)</sup>

**Objectives :** This study was done to evaluate the reliability of education and occupational class between using the health survey and the death certificate data.

**Methods :** The 1998 National Health and Nutrition Examination Survey (NHANES) was conducted on a cross-sectional probability sample of South Korean households, and it contained unique 13-digit personal identification numbers that were linked to the data on mortality from the Korean National Statistical Office. The data from 269 deaths were used to estimate the agreement rates and the Kappa indices of the education and occupational class between using the NHANES data and the death certificate data.

**Results :** The simple and weighted Kappa indices for education were 0.60 (95% CI=0.53-0.68) and 0.73 (95% CI=0.67-0.79) respectively, if the educational level was grouped into five categories: no-formal-education, elementary-school, middle-school, high-school and college or over. The overall agreement rate was 71.9% for these educational groups. The magnitude of reliability, as measured by the overall agreement rates and Kappa indices, tended to increase with a decrease in the

educational class. The number of non-educated people with using the death certificate data was smaller than that with using the NHANES data. For the occupational class (manual workers, non-manual workers and others), the Kappa index was 0.40 (95% CI=0.30-0.51), which was relatively lower than that for the educational class. Compared with the NHANES, the number of non-manual workers for the deceased who were aged 30-64 tended to be increased (8 to 12) when using the death certificate data, whereas the number of manual workers tended to be decreased (59 to 41).

**Conclusions :** The socioeconomic inequalities in the mortality rates that were based on the previous unlinked studies in South Korea were not due to a numerator/denominator bias. The mortality rates for the manual workers and the no-education groups might have been underestimated.

*J Prev Med Public Health 2005;38(4):443-448*

**Key words :** Reliability, Education, Occupation, Death certificate, South Korea

## 서론

교육수준, 직업 등의 사회경제적 위치 (socioeconomic position)에 따른 사망률 차이(즉, 불평등)는 서구 국가를 중심으로 많이 보고되었다. 국가에서 생검통제의 일 부분으로 전체 국민의 사망 자료를 구축하고 있기 때문이다. 즉, 인구센서스에서 교육수준, 직업 등의 사회경제적 위치 변수를 이용하여 사망률 계산에서의 인구수(분모)를 구하고, 분자에 해당하는 사망자

수를 사망등록자료에서 얻어, 사망률 차이를 계산할 수 있다. 이 같은 방법은 비연계자료 연구(unlinked study)라는 이름을 갖고 있는데 [1], 영국의 경우 오랜 역사적 전통을 갖고 있어 1980년 Black Report의 사망률 불평등을 기술하는 방법으로 활용되기도 하였다 [2]. 미국의 경우에도 같은 방식을 이용하여 사망률 불평등을 보고하였고 [3], 유럽에서의 국가간 비교에서도 활용하였다 [4]. 우리 나라에서도 이 방법으로 사망률에서의 사회계층간 불평등 양상

을 보고한 연구가 과거 몇 차례 수행되었고 [5,6], 최근 연구량이 증가하고 있다 [7-11].

그런데, 인구센서스 자료와 사망등록자료를 이용한 비연계자료 연구들은 연구대상자의 사망여부에 대한 추적이 개인별로 이루어지지 않은 연구이다. 그러므로 사망률 추정 과정상의 분모, 분자 정보의 부정확성으로 인한 분모-분자 비틀림(numerator-denominator bias) [1]의 가능성이 상존한다. 이 때문에, 비연계자료 연구를 수행한 우리 나라 연구자들은 교육수준, 직업 변수의 신뢰도가 어느 정도가 되는

지에 대한 의문을 제기한 바 있다 [7,9,10,12]. 물론 사회경제적 위치 변수 이외에도 성별, 연령 변수에서의 신뢰도도 검토되어야 하지만, 이들의 신뢰도는 일반적으로 매우 높기 때문에, 실제로는 사회경제적 위치 지표의 신뢰도가 논란거리이다. 특히, 우리 나라의 사망등록자료에는 성, 연령 정보를 갖고 있는 주민등록번호가 있어, 만약 인구센서스에서의 성별, 연령 조사가 문제점을 갖고 있지 않다면, 이들 변수에서의 신뢰도는 매우 높을 수 있는 하부구조를 가지고 있다. 하지만, 교육수준과 직업계층과 같은 변수는 성격이 다르다. 예를 들어, 인구센서스 자료의 학력에 비하여 사망등록자료에서의 학력이 낮게 보고되는 경향이 있다면, 실제로는 교육수준에 따른 사망률 불평등이 없거나 작더라도, 불평등이 있거나 큰 것처럼 추정될 위험이 있다. 물론 교육수준은 실제보다 높게 보고되는 경향이 있고 [13], 특히 면접조사자료에 비하여 사망등록자료는 사후에 가족 등이 보고하기 때문에 교육수준이 높게 보고되는 경향이 있기 때문에 [14,15], 우리나라에서 관찰된 교육수준별 사망률 불평등 [7-11]의 크기가 과대 평가된 것이 아니라는 판단이 가능하지만, 실제로 그러할지는 신뢰도 연구 결과로서 뒷받침되어야 한다.

우리 나라에서 사망등록자료의 직업, 교육수준에 있어서 신뢰도는 Kong 등 [16]과, Son [17]이 평가한 바 있다. 하지만, 이들 연구들은 사망등록자료의 사회경제적 위치 지표를 사후에 재확인하는 방식으로 수집된 자료를 이용한 연구이다. 즉, 사전 면접 조사자료에서의 사회경제적 위치 지표를 사후 사망등록자료에서 확인하는 방법이 아니었기 때문에, 사회경제적 사망률 불평등을 다룬 비연계연구 결과가 어느 정도 정확할지에 대한 정보를 제공하기에는 한계가 있다. 이 점에서 「1998년도 국민건강영양조사」 자료는 매우 중요한 의의를 갖는다. 즉, 「1998년도 국민건강영양조사」의 건강검진조사표에 기재된 각 개인의 주민등록번호를 이용하여, 통계청의 사망등록 자료와 연계함으로써, 면접조사자료와 사망등록자료 간 사회경제적 위치 지표의 신뢰도를 평가할 수 있다. 국민건강

영양조사에서는 인구센서스와 비슷한 방식(가구주 면접조사)으로 사회경제적 위치 정보가 수집되기 때문에, 「1998년도 국민건강영양조사」 사망추적 자료에서의 신뢰도는 기존 비연계자료 연구에서의 오류의 방향과 크기를 파악하는 데에 매우 좋은 판단 근거가 될 수 있다. 이러한 평가의 결과는 센서스 자료와 사망등록자료를 이용한 사망률 불평등의 모니터링 가능성에 대한 판단근거를 제공할 수 있다. 예를 들어, 면접조사 자료와 사망등록자료간의 사회경제적 위치 지표의 신뢰도가 상당한 수준이라면, 센서스 자료와 사망등록자료를 이용한 사망률 모니터링이 과학적 근거를 갖고 이루어질 수 있고, 또한 향후 사망률 불평등 감소를 위한 건강목표 설정에 있어서도 센서스 자료와 사망등록자료를 이용한 연구결과가 도움을 줄 수 있을 것이다.

이 연구의 목적은 「1998년도 국민건강영양조사」의 교육, 직업 변수와 사망추적 자료에서의 교육, 직업 변수를 이용하여, 사회경제적 위치 변수의 신뢰도를 평가하는 데에 있다.

### 대상(재료) 및 방법

연구 자료는 <한국보건사회연구원>에서 1998년 11월부터 12월까지 실시한 「1998년도 국민건강영양조사」의 대상자들 중, 건강검진조사표에 주민등록번호 13자리 숫자가 기재되어 있어서 2003년도 말까지 통계청 사망 확인서비스를 통하여 사망 여부가 추적된 자료이다. 「1998년도 국민건강영양조사」의 건강·검진조사를 받은 대상자는 9,771명이었는데, 이 중 주민등록번호 13자리가 모두 있는 사람은

9,602명이었고, 이 중 주민등록번호 유효성 검사 프로그램을 통하여 유효한 주민등록번호를 가진 사람은 8,559명이었다. 30세 미만 사망자수가 매우 적고, 30세 미만 연령층에서는 정규교육이 종료되지 않을 수 있기 때문에, 이 연구에는 30세 이상 5,607명의 주민등록번호가 이용되었다. 1999년부터 2003년까지 사망한 264명(남자 160명, 여자 104명) 중에서 사망등록자료상 교육수준이 미상으로 기재된 1명을 제외하여 263명의 자료를 분석하였다.

교육수준과 직업 모두 분류기준을 어떻게 하느냐에 따라 신뢰도에 차이를 보인다. 교육수준 분류의 경우, 「1998년도 국민건강영양조사」에서는 '안다녔음'과 '학교는 안다녔지만 국문해독이 가능함' 항목을 무학으로 분류하였다. '전문대'와 '대학원 이상'은 대학으로 구분하였다. 최종적으로 무학, 초등학교, 중학교, 고등학교, 대학 이상의 5단계로 구분하였다. 무학과 초등학교를 합하여 4단계, 무학/초등학교와 중학교/고등학교를 각각 하나의 군으로 합한 3단계 구분에서의 연구 결과도 제시하였다. 「1998년도 국민건강영양조사」 조사지침에 따르면 [18], 중퇴 등의 경우에는 이전 학력으로 기재하도록 하고 있어, 졸업을 기준으로 교육수준을 기재하도록 한 사망등록자료의 기준과 동일하다. 사망등록자료에서의 교육수준은 무학, 초등학교, 중학교, 고등학교, 대학 이상으로 구분되어 있다. 직업 계층은 한국표준 직업분류 상의 대분류 1~5를 비육체노동자(non-manual worker), 대분류 6~9를 육체노동자(manual worker)로 구분하고, 나머지 주부(가사), 무직, 학생, 군인 및 미상을 기타(others)로 분류하였다. 「1998년도 국민건강영양조사」자료와 사망등록자료 간의

**Table 1.** Comparison of educational level between 1998 National Health and Nutrition Examination Survey data and 1998-2003 death certificate data among 263 decedents

		1998 National Health and Nutrition Examination Survey data (From Korea Institute for Health and Social Affairs)					
		No education	Elementary	Middle	High	College or higher	Total
1999-2003 Death certificate data (From National Statistical Office)	No education	88 ( 78.6)	9 ( 11.7)	1 ( 3.4)	0 ( 0.0)	0 ( 0.0)	98
	Elementary	22 ( 19.6)	56 ( 72.7)	4 ( 13.8)	6 ( 18.2)	1 ( 8.3)	89
	Middle	1 ( 0.9)	10 ( 13.0)	14 ( 48.3)	5 ( 15.2)	0 ( 0.0)	30
	High	1 ( 0.9)	2 ( 2.6)	10 ( 34.5)	21 ( 63.6)	1 ( 8.3)	35
	College or higher	0 ( 0.0)	0 ( 0.0)	0 ( 0.0)	1 ( 3.0)	10 ( 83.4)	11
Total		112 (100.0)	77 (100.0)	29 (100.0)	33 (100.0)	12 (100.0)	263

**Table 2.** Percentage agreement and Kappa values (95% confidence intervals, CI) for educational level: reliability between 1998 National Health and Nutrition Examination Survey data and 1999-2003 death certificate data

	No. of subject	Percentage agreement(%)	Simple Kappa (95% CI)	Weighted Kappa (95% CI)
<b>Five educational groups (no education, elementary, middle, high, and college<sup>a</sup>)</b>				
All	263	71.9	0.60 (0.53-0.68)	0.73 (0.67-0.79)
Sex				
Male	159	65.4	0.54 (0.45-0.64)	0.70 (0.62-0.78)
Female	104	81.7	0.63 (0.49-0.77)	0.68 (0.54-0.81)
Age groups				
30-64	104	65.4	0.54 (0.41-0.66)	0.67 (0.56-0.77)
65 or over	159	76.1	0.59 (0.48-0.69)	0.70 (0.61-0.79)
<b>Four educational groups (elementary or less, middle, high, and college<sup>a</sup>)</b>				
All	263	83.7	0.64 (0.55-0.73)	0.74 (0.66-0.82)
Sex				
Male	159	78.0	0.63 (0.53-0.73)	0.73 (0.65-0.82)
Female	104	92.3	0.50 (0.22-0.78)	0.57 (0.28-0.85)
Age groups				
30-64	104	74.0	0.60 (0.48-0.73)	0.69 (0.57-0.80)
65 or over	159	89.9	0.60 (0.44-0.75)	0.71 (0.56-0.85)
<b>Three educational groups (elementary or less, middle or high, and college<sup>a</sup>)</b>				
All	263	89.4	0.75 (0.67-0.84)	0.78 (0.69-0.86)
Sex				
Male	159	86.8	0.75 (0.66-0.85)	0.78 (0.66-0.89)
Female	104	93.3	0.55 (0.26-0.85)	NA*
Age groups				
30-64	104	82.7	0.69 (0.56-0.82)	0.71 (0.58-0.84)
65 or over	159	93.7	0.74 (0.59-0.90)	0.78 (0.64-0.92)

\*Weighted Kappa was not computed because there was no women with college<sup>a</sup> education.

교육수준과 직업의 신뢰도를 일치율 (overall agreement rate)과 카파값(Kappa)을 통하여 평가하였다.

## 결 과

### 1. 교육수준의 신뢰도

「1998년도 국민건강영양조사」에서 보고된 교육수준과 사망등록자료의 교육수준을 비교한 결과 전체의 71.9% (189명)가 교육수준이 일치하는 것으로 나타났다 (Table 1)

「1998년도 국민건강영양조사」에서 무학으로 보고된 사람들 중에서 88명 (78.6%)은 사망등록자료에서도 무학으로 기재되어 있었지만, 나머지 21.4%는 상위 교육수준으로 보고되었다. 반면 「1998년도 국민건강영양조사」에서 초등학교 학력을 가진 사람 중 9명 (11.7%)과, 중학교 학력자 중 1명 (3.4%)은 사망등록자료에는 무학으로 기재되었다. 결과적으로 「1998년도 국민건강영양조사」에서는 112명의 무학자가 있었지만, 사망등록자료에서는 98명의 무학자 (112명의 87.5%)가 있는 것으로 나타났다. 이러한 양상은 연령군을 바꾸어

분석하여도(예: 연구대상자를 30~64세 연령군으로 하여 분석) 비슷한 양상이었다.

이처럼 「1998년도 국민건강영양조사」에서의 무학자 수보다 사망등록자료에서의 무학자 수가 적었지만, 초등학교, 중학교, 고등학교 학력에서는 이와 반대의 양상으로, 사망등록자료에서의 적게 보고된 무학자 수를 상쇄하는 양상이었다. 즉, 「1998년도 국민건강영양조사」에서는 초등학교, 중학교, 고등학교 학력자가 각각 77명, 29명, 33명이었지만, 사망등록자료에서는 각각 89명, 30명, 35명이었다.

전체 263명의 사망자에 대한 「1998년도 국민건강영양조사」 자료와 사망등록자료 간 교육수준(5단계)의 단순 카파값 (simple Kappa)은 0.60 (95% CI=0.53-0.68), 가중 카

파값 (weighted Kappa)은 0.73 (95% CI=0.67-0.79)으로 나타났다 (Table 2).

또한 교육수준 분류법을 달리하여, 성·연령별로 제시한 결과에서는 전반적으로 교육수준 단계가 적을수록 일치율과 카파값은 높아지는 양상이었다. 즉, 교육수준을 초등학교 및 무학, 중학교, 고등학교, 대학 이상의 4단계로 나누었을 때, 전체의 83.7%가 「1998년도 국민건강영양조사」 자료와 사망등록자료 간 교육수준이 일치하였고, 단순 카파값은 0.64 (95% CI=0.55-0.73)이었는데, 중학교 학력과 고등학교 학력을 합쳐 교육수준을 3단계로 나누었을 때는 각각 89.4%와 0.75 (95% CI=0.67-0.84)로 일치율과 카파값이 증가하는 양상이었다. 연령별로는 30~64세 연령대에 비하여 65세 이상 연령에서의 일치율이 높은 양상이었지만, 카파값으로 평가하였을 때, 노년층에서의 신뢰도가 통계적으로 유의하게 높은 양상은 아니었다.

### 2. 직업계층의 신뢰도

「1998년도 국민건강영양조사」에서 보고된 직업계층과 사망등록자료의 직업계층을 비교한 결과 전체의 70.3% (185명)가 직업계층이 일치하는 것으로 나타났다 (Table 3).

「1998년도 국민건강영양조사」에서 비육체노동자로 보고된 12명의 사망자 중에서 4명 (33.3%)만이 사망등록자료에서 비육체노동자로 구분되었고, 절반인 6명은 기타로 분류가 이루어졌다. 사망등록자료에서 기타로 분류된 6명 중 4명은 무직이었다. 한편 「1998년도 국민건강영양조사」에서 육체노동자로 보고된 88명의 사람 중에서 50명 (56.8%)은 사망등록자료에서 육체노동자로 분류되었지만, 7명 (8.0%)은 비육체노동자로, 31명 (35.2%)은 기타로 분

**Table 3.** Comparison of occupational class between 1998 National Health and Nutrition Examination Survey data and 1998-2003 death certificate data among 263 decedents

		1998 National Health and Nutrition Examination Survey Data (From Korea Institute for Health and Social Affairs)			
		Non-manual	Manual	Others	Total
1999-2003 Death certificate data (From National Statistical Office)	Non-manual	4 ( 33.3)	7 ( 8.0)	2 ( 1.2)	13
	Manual	2 ( 16.7)	50 ( 56.8)	30 ( 34.4)	82
	Others	6 ( 50.0)	31 ( 35.2)	131 ( 80.4)	168
	Total	12 (100.0)	88 (100.0)	163 (100.0)	263

류되었다. 기타로 분류된 31명 중에서 28명은 무직이었다.

이처럼 「1998년도 국민건강영양조사」와 사망등록자료에서의 비육체노동자 및 육체노동자의 직업계층이 일치하는 경우가 높지는 않았지만, 각 집단별 「1998년도 국민건강영양조사」와 사망등록자료 간의 사망자수의 차이는 크지 않은 것으로 나타났다. 즉, 「1998년도 국민건강영양조사」에서는 비육체노동자, 육체노동자, 기타 집단이 각각 12명, 88명, 163명이었는데, 사망등록자료에서는 각각 13명, 82명, 168명이었다. 하지만, 대상 연령군을 바꾸어 분석하면 다른 양상이었다. 30세~64세의 사망자 104명을 대상으로 할 경우, 「1998년도 국민건강영양조사」에서는 비육체노동자, 육체노동자, 기타 집단이 각각 8명, 59명, 37명이었는데, 사망등록자료에서는 각각 12명, 41명, 51명이었다. 즉, 상대적으로 사망등록자료에서 비육체노동자의 수는 증가한데 반해, 육체노동자의 수는 크게 감소하는 양상이었다.

전체 263명의 사망자에 대한 「1998년도 국민건강영양조사」 자료와 사망등록자료 간 직업계층(3단계)의 단순 카파값(simple Kappa)은 0.40 (95% CI=0.30-0.51)으로 나타나 교육수준의 카파값에 비하여 낮은 수준이었다 (Table 4). 여성 사망등록자료에서 비육체노동자로 분류된 대상이 없어, 카파값을 계산하지 못하였기 때문에 성별 카파값의 비교는 어려웠다. 연령별로는 30~64세 연령군에서 일치율은 58.7%, 카파값이 0.30 (95% CI=0.15-0.45), 65세 이상에서는 일치율이 78%, 카파값은 0.40 (95% CI=0.24-0.57)로, 노년층에서 일치율이 높은 양상이었지만, 카파값의 차이는 통계적으로 유의하지 않았다.

## 고 찰

신뢰도는 측정 방법에 따라 크게 영향을 받을 수 있다. 사망등록 자료와 「1998년도 국민건강영양조사」 지침서에서 공히 학력 작성 기준은 졸업 기준으로, 직업은 한국표준직업분류 상의 대분류에 따라 작성하고 있으므로 기준의 차이에 따른 문제

**Table 4.** Percentage agreement and Kappa values (95% confidence intervals, CI) for occupational class: reliability between 1998 National Health and Nutrition Examination Survey data and 1999-2003 death certificate data

	No. of subjects	Percentage agreement(%)	Simple Kappa (95% CI)
All	263	70.3	0.40 (0.30-0.51)
Sex			
Male	159	61.6	0.32 (0.19-0.45)
Female	104	83.7	NA*
Age groups			
30-64	104	58.7	0.30 (0.15-0.45)
65 or over	159	78.0	0.40 (0.24-0.57)

\*Kappa was not computed because there was no women with non-manual occupation in the death certificate data.

는 없었다.

「1998년도 국민건강영양조사」와 사망등록자료에서 단순 카파값으로 평가한 교육수준의 신뢰도는 교육수준을 5단계로 할 경우 0.60, 4단계로 할 경우 0.64, 3단계로 할 경우 0.75이었다. Landis와 Koch [19]는 카파값을 이용한 신뢰도 평가 기준을 제시한 바 있는데, 카파값이 0.00 미만인 경우는 '나쁨 (poor)', 0.00~0.20은 '약함 (slight)', 0.21~0.40은 '보통 (fair)', 0.41~0.60은 '중간 (moderate)', 0.61~0.80은 '상당한 수준 (substantial)', 0.81~1.00은 '거의 완벽 (almost perfect)'으로 평가하였다. 이 기준을 적용할 경우, 교육수준에서의 신뢰도는 '상당한 수준'에 있다고 할 수 있다. 또한 가중 카파값을 적용한 경우에는 교육수준의 신뢰도가 더 높아졌다. 즉, 교육수준에 따른 기준 비연계연구 결과 [7-11], 발견된 사망률 불평등이 분모-분자 비뒤틀림에 따른 오류 때문일 가능성은 낮다고 할 수 있다. 즉, 실제 현상을 반영한다고 할 수 있다.

반면 직업계층의 신뢰도는 교육수준보다는 낮은 수준이었다. 전체 집단에 대한 단순 카파값은 0.40으로 '보통(fair) 수준'의 일치도에 해당하였으며, 남자와 30~64세 연령대에서는 0.30으로 더 낮았다.

교육수준은 실제보다 높게 보고되는 경향이 있다 [13]. 하지만, 비연계자료를 이용한 사회경제적 사망률 불평등 연구에 있어서 현실적으로 더 중요한 문제는 인구센서스 자료와 사망등록자료 간의 교육수준의 신뢰도이다. 교육수준에 있어서 조사 자료와 사망등록자료 간의 신뢰도를

다른 외국 연구로는 Shai와 Rosenwaike [14], Sorlie와 Johnson [15], Rosamond 등 [20]의 연구가 있다. Shai와 Rosenwaike [14], Sorlie & Johnson [15]의 연구에서는 조사 자료에 비하여 사망등록자료에서 교육수준이 높게 보고되는 경향이 있다고 보고하였지만, Rosamond 등 [20]의 연구에서는 오히려 사망등록자료에서 교육수준이 약간 낮게 보고되는 경향을 보고하였다. 이 연구에서는 조사 자료에 비하여 사망등록자료에서 교육수준이 높게 보고되는 경향을 보였는데, 무학이 초등학교나 중학교 학력으로 기재되는 양상이 뚜렷하였다.

일치율을 보면, Shai와 Rosenwaike [14]의 연구에서는 교육수준을 6단계로 나눌 경우 일치율이 68%였고, Rosamond 등 [20]의 연구에서는 교육수준을 고졸 미만, 고졸, 대학 이상의 3단계로 나눌 경우 일치율이 84%였다. 각 연구들의 연령 구조와 학력 분포의 차이가 있어 직접적인 비교는 어렵지만, Rosamond 등 [20]의 연구처럼 3단계로 구분할 경우의 일치율은 89.4%로, 이 조사에서의 교육수준의 일치율 71.9%~89.4%와 비슷한 수준이었다.

한편, 직업에 있어서 조사 자료와 사망등록자료 간의 신뢰도 문제는 직업별 사망률(특히 암 사망률) 측정과 관련하여 다수의 연구가 미국을 중심으로 이루어진 바 있다 [21-25]. 이들 연구 결과에 따르면, 일치율은 산업 분야의 경우 50~80%, 직업의 경우 이보다 낮은 약 45~70%를 보였다 [24]. 물론 이러한 일치율은 직업을 어떻게 분류하느냐에 따라 증감이 있게 된다.

교육수준과 직업의 일치도에 대한 우리나라 연구로는 다소 오래 전의 Kong 등 [16]의 연구와 비교적 최근의 Son [17]의 연구를 들 수 있다. Kong 등 [16]은 1981년에 사망 신고된 994명을 대상으로 약 1년 후인 1982년 9월에 해당 사망자 가정을 방문하여 사망자의 교육수준, 직업, 결혼상태, 사망원인, 사망장소 등에 대한 정보를 얻은 것이다. 연구 결과, 학력의 경우 일치율은 79.1%이었고, 직업의 일치율은 77.6%이었다. Son [17]은 우리나라 사망등록자료와 근로복지공단에 산재 사망에 따른 보상신청을 받기 위하여 신청한 산재사망

자료를 이용하여 6,866명에서의 직업, 교육, 사망원인의 신뢰도를 평가하였다. 연구 결과 직업계층을 비육체노동자와 육체노동자로 나눌 경우의 일치율은 78.4%, 카파값은 0.49로 나타났고, 교육수준의 경우 4단계 (초등학교 및 무학, 중학교, 고등학교, 대학교)로 나눈 경우, 일치율은 46.9%, 카파값은 0.25인 것으로 나타났다. 이들 국내 연구의 결과와 본 연구 결과를 직접 비교하기에는 연구대상이 상이한 점(농촌 주민, 산재대상자 vs. 전국 대표 표본에서의 사망자), 연령대가 다른 점, 교육수준 및 직업 분포의 차이 등의 문제가 있다. 하지만, 전반적으로 기존 국내 연구의 결과 [16,17]는 직업에서의 일치 정도가 이 연구보다 높은 양상이었다. 이러한 양상은 자료 수집 방법상의 차이에서 기인한 측면이 클 것으로 보인다. 「1998년도 국민건강영양조사」의 사망추적 자료를 사용한 연구는 면접조사에서 직업이 먼저 조사된 후 사망등록자료에서 직업이 다시 확인되는 방식을 거쳤다. 그러므로 사망 당시에는 당초의 직업과는 다른 직업을 가졌거나 무직으로 사회이동 (social mobility)이 일어났을 가능성이 있다. 반면, Kong 등 [16]과 Son [17]의 연구의 경우, 사망등록자료의 사망자 직업을 사후에 재확인하는 방식으로 수집된 자료를 토대로 직업에서의 일치도를 보았기 때문에, 응답자(또는 산재신고자)들은 사망자의 사망 당시 직업을 비교적 정확하게 응답할 수 있었을 것이다. 특히 우리나라의 사망신고서 상의 직업은 '사망의 원인이 되는 질병 또는 사고가 발생한 때의 직업'이므로, 산재로 사망한 사람의 경우 사망등록자료와 산재사망자료 간의 직업의 일치도가 높을 가능성이 많다.

앞서 언급하였듯이, 사회경제적 위치에 대한 신뢰도 연구 결과는 기존 비연계자료 연구 결과의 오류(오류의 방향과 정도)에 대한 정보를 제공할 수 있다. 기존 비연계자료 연구 결과가 타당성을 갖기 위해서는 인구센서스와 사망등록자료가 100%에 이르는 완전성을 갖추어야 하고, 두 자료간의 성별, 연령, 사회경제적 위치 지표의 신뢰도가 확보되어야 한다. 인구센서

스와 사망등록자료의 완전성에 영향을 주어 사회경제적 사망률 불평등의 크기를 좌우할 수 있는 요인들로는, 이민자(해외/국내), 해외 유학생, 해외 근로자 등이 있을 수 있지만, 법무부의 이민 자료 등을 검토한 결과, 이들 요인의 효과는 그리 크지 않을 것으로 보고된 바 있다 [10]. 또한 일반적으로 성별과 연령 자료의 신뢰성은 매우 크다. 결국 비연계자료 연구 결과의 타당성에 가장 큰 영향을 주는 요인은 인구센서스와 사망등록자료 간의 사회경제적 위치 지표(교육수준, 직업 등)의 신뢰도 문제이다.

사회경제적 위치 지표의 신뢰도를 평가하기 위해서는 사망등록자료에서 사망자의 사회경제적 위치를 인구센서스에서 사회경제적 위치와 비교하는 방법이 최선이다. 하지만, 인구센서스 개인 자료에 대한 접근성 문제와 함께, 인구센서스 자료에는 주민등록번호가 없어 이를 사망등록자료와 연계하는 것은 실제로는 어렵다. 따라서 인구센서스와 유사한 방식으로 이루어지는 면접조사자료의 교육, 직업 지표와 사망등록자료의 동일 지표를 비교하는 방법이 현실적으로 좋은 방법이다. 교육수준에서의 일치도에 대한 본 연구 결과는, 기존 비연계연구 결과 [7-11]가 이상적인 수준은 아니지만 상당한 정도 신뢰성이 있을 것이라는 점을 시사한다. 본 연구는 기존의 Khang 등의 연구 [10]에서와 같이 교육수준을 무학, 초등학교 졸업, 중학교 졸업, 고등학교 졸업, 대학 이상의 5 단계로 나누었는데, 무학의 경우 「1998년도 국민건강영양조사」보다 사망등록자료에서의 사망자가 적은 것으로 나타나 (112명→98명), 기존 Khang 등의 연구 결과 [10]에서 나타난 무학에서의 높은 사망률 상대비는 자료상의 오류라기보다는 실제 상황을 반영한 것이라는 점을 시사한다. 교육수준별 사망률 불평등 수준이 과추계 (overestimation)되기 위해서는 면접조사에서 무학보다 상위 교육수준에 해당되는 경우가 사망등록자료에서 보다 적게 보고되어야 할 것이다. 한편, 본 연구결과 사망등록자료에서의 적게 보고된 무학자 수를 초등학교와 중학교 학력에서 상쇄하는 양

상이었으므로, 기존 연구 [10]에서의 초등학교, 중학교 학력자의 상대비는 다소 과대추정되었을 가능성이 있다.

한편, 비연계자료를 이용한 직업계층에 따른 사망률 차이 [7-9]는 다소 과소추정되었을 가능성이 있다. 연구 결과, 30세~64세의 사망자 104명 중, 「1998년도 국민건강영양조사」에서는 비육체노동자, 육체노동자, 기타 집단이 각각 8명, 59명, 37명이었는데, 사망등록자료에서는 각각 12명, 41명, 51명으로, 사망등록자료에서 비육체노동자의 수는 증가한데 반해, 육체노동자의 수는 감소하는 양상이었다. 결과적으로 연계자료가 아닌 비연계자료를 이용하여 비육체노동자/육체노동자 간 사망률 불평등을 연구한다면, 원래의 사망 위험의 상대비보다 낮은 상대비를 산출하게 될 것이다. 「1998년도 국민건강영양조사」의 사망추적 자료를 이용하여, 30~64세 연령군에서의 교육수준과 직업에서의 사망 위험의 상대비를 추정한 연구 [26]에 따르면, 기존 비연계자료 연구 결과 [9]와 달리, 성, 연령과 교육수준을 보정한 후에도 육체노동자에서의 사망 위험 상대비가 비육체노동자에 비하여 3.49배 (95% CI=1.33-9.17) 높은 것으로 나타났다.

이 연구에 사용한 자료는 몇 가지 한계를 가지고 있다. 사망자의 수가 많지 않아 보다 세분화한 연령군에서의 신뢰도가 평가되지 못하였고, 직업 세부 분류(13대 분류 등)를 이용한 신뢰도 평가도 이루어지지 못하였다. 또한 사망 원인별 사회경제적 사망률 불평등 결과 [7,10]의 타당성을 평가하기 위해서는, 기존 연구처럼 [17] 사망 원인별로 사회경제적 위치의 신뢰도가 평가되어야 하지만, 이를 위해서는 보다 많은 사망자가 확보되어야 한다.

이러한 한계를 극복하기 위하여, 「1998년도 국민건강영양조사」를 앞으로 지속적으로 사망 여부를 추적하여, 보다 많은 사망자를 확보하는 방법이 있다. 하지만, 이는 몇 가지 한계에 부딪힐 수밖에 없는 데, 직업은 시간에 따라 바뀔 확률이 커지며, 교육 또한 부분적으로 바뀔 가능성이 있다. 또한 사망 원인별, 직업 세부 분류에서의 신뢰도를 평가하기 위해서는 추적해

야 할 기간이 길어질 수밖에 없다는 단점이 있다. 이 상황에서 보다 직접적으로 보다 많은 사망자를 대상으로 신뢰도 평가를 하는 것이 향후 연구의 과제인데, 사망등록자료에서의 사망자와 인구센서스의 교육수준과 직업을 연계하여 평가하는 연구가 그것이다. 예를 들어, 특정 시군구 지역의 사망자를 해당 지역의 이전 인구센서스 자료와 연계하는 방법을 이용할 경우, 두 자료를 연계하는 것이 상대적으로 수월할 수 있을 것이다. 이렇게 할 경우, 보다 많은 사망자 수의 확보가 가능하고, 기존의 비연계자료 연구 결과의 타당성 평가에 필요한 사회경제적 위치의 신뢰도를 이 연구보다 직접적인 방식으로 연구할 수 있다.

## 결론

「1998년도 국민건강영양조사」와 사망등록자료의 교육수준 및 직업계층 분류를 비교한 결과, 교육수준의 신뢰도가 상당한 수준에 있는 것으로 나타났다. 직업계층에서의 신뢰도는 교육수준보다 낮은 수준이었다. 교육수준의 경우, 교육수준 분류를 단순화할수록 신뢰도는 증가하였다. 연구 결과를 토대로 할 때, 교육수준에 따른 사망률 불평등을 보여주고 있는 기존 연구 결과들은 분자-분모 비뮌림에 의한 것이 아닌 실제 사망률 차이를 반영하는 것으로 판단된다. 그러므로 센서스 자료와 사망등록자료를 이용한 교육수준별 사망률 불평등의 모니터링이 지속될 필요가 있다.

## 참고문헌

- Macintyre S. The black report and beyond: What are the issues? *Soc Sci Med* 1997; 44: 723-745
- Townsend P, Davidson N, Whitehead M. Inequalities in Health: the Black Report and the Health Divide. 2nd ed. London: Penguin Books; 1992
- Kitagawa EM, Hauser PM. Differential mortality in the United States: A study in socioeconomic epidemiology. Cambridge: Harvard University Press; 1973
- Fox J. Health Inequalities in European Countries. Aldershot: Gower; 1989
- Kwon TH. The trends and patterns of mortality and health in the Republic of Korea (Asian Population Studies Series No. 76). Bangkok: Economic and Social Commission for Asia and the Pacific; 1986
- Kim TH. Mortality transition in Korea: 1960-1980. Seoul National University: The Population and Development Studies Center; 1990
- Son M. Occupational class and health: The differentials in mortality, morbidity and workplace injury rates by occupation, education, income and working conditions in Korea [dissertation]. London: London School of Hygiene & Tropical Medicine, Univ; 2001
- Son M. The relationships of occupational class, educational level and deprivation with mortality in Korea. *Korean J Prev Med* 2002; 35(1): 76-82 (Korean)
- Son M, Armstrong B, Choi JM, Yoon TY. Relation of occupational class and education with mortality in Korea. *J Epidemiol Community Health* 2002; 56: 798-799
- Khang YH, Lynch JW, Kaplan GA. Health inequalities in Korea: Age- and sex-specific educational differences in the 10 leading causes of death. *Int J Epidemiol* 2004; 33(2): 299-308
- Khang YH, Lynch JW, Yoon SC, Lee SI. Trends in socioeconomic health inequalities in Korea: Use of mortality and morbidity measures. *J Epidemiol Community Health* 2004; 58: 308-314
- Son M. Commentary: Why the educational effect is so strong in differentials of mortality in Korea? *Int J Epidemiol* 2004; 33(2): 308-310
- Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988; 10: 87-121
- Shai D, Rosenwaike I. Errors in reporting education on the death certificate: some findings for older male decedents from New York State and Utah. *Am J Epidemiol* 1989; 130(1): 188-192
- Sorlie P, Johnson NJ. Validity of education information on the death certificate. *Epidemiology* 1996; 7: 437-439
- Kong SK, Lim JK, Kim MK. Mortality and causes of death. Seoul: Korea Institute for Population and Health; 1980 (Korean)
- Son M. A comparison of occupation, education, and cause of death from national death certificates and deaths data due to workplace injuries from WELCO in Korea. *Korean J Epidemiol* 2002; 23: 44-51 (Korean)
- Ministry of Health and Welfare of Korea. Interim report on the 1998 National Health and Nutrition Survey. Kwachon: Ministry of Health and Welfare of Korea; 1999 (Korean)
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-174
- Rosamond WD, Tyroler HA, Chambless LE, Folsom AR, Cooper L, Conwill D. Educational achievement recorded on certificates of death compared with self-report. *Epidemiology* 1997; 8: 202-204
- Swanson GM, Schwartz AG, Burrows RW. An assessment of occupation and industry data from death certificates and hospital medical records for population-based cancer surveillance. *Am J Public Health* 1984; 74: 464-467
- Gute DM, Fulton JP. Agreement of occupation and industry data on Rhode Island death certificates with two alternative sources of information. *Public Health Reports* 1985; 100: 65-72
- Lilienfeld DE. Occupation on the death certificate: To use or not to use, that is the question [Editorial]. *Am J Ind Med* 1988; 14: 119-120
- McLaughlin JK, Mehl ES. A comparison of occupational data from death certificates and interviews. *Am J Ind Med* 1991; 20: 335-342
- Andrews KW, Savitz DA. Accuracy of industry and occupation on death certificates of electric utility workers: Implications for epidemiologic studies of magnetic fields and cancer. *Bioelectromagnetics* 1999; 20: 512-518
- Khang YH, Kim HR. Relationship of education, occupation, and income with mortality in a representative longitudinal study of South Korea. *Eur J Epidemiol* 2005; 20: 217-220