

## 노인 생활습관 설문서의 신뢰도 및 타당도 평가 연구<sup>1)</sup>

박병주<sup>1</sup>, 김대성<sup>1</sup>, 구혜원<sup>1</sup>, 배종면<sup>2</sup>

서울대학교 의과대학 예방의학교실<sup>1</sup>, 제주대학교 의과대학 예방의학교실<sup>2</sup>

= Abstract =

### Reliability and validity study of a life style questionnaire for elderly people

Byung-Joo Park<sup>1</sup>, Dae Sung Kim<sup>1</sup>, Hye Won Koo<sup>1</sup>, Jong Myon Bae<sup>2</sup>

*Department of Preventive Medicine, College of Medicine, Seoul National University<sup>1</sup>*

*Department of Preventive Medicine, College of Medicine, Jeju National University<sup>2</sup>*

The study was done to determine the reliability and validity of a life style questionnaire for the elderly. The questionnaires were sent to 16,524 elderly people who were beneficiaries of Korean Medical Insurance Corporation in Pusan. Among the completed 9,139 questionnaires, 200 were randomly sampled and retested. Finally, 110 duplicates were collected. Weighted kappa-value and Pearson correlation coefficients were estimated to measure the reliability. Validity coefficient was estimated by using reliability coefficient. In self-self responses, reliability coefficients of the most of items were over 0.6 except some physical activity related item. Relatively high reliability was observed in smoking, alcohol related items and anthropometric items. In self-proxy responses, most of the physical activity related items were found to be less reliable than self-self responses. Smoking and alcohol related items were consistently reliable. Male showed higher validity in food related item than female. On the other hand, some of the physical activity related items and smoking and alcohol related items were less valid in male than female. With regard to bias of proxy respondents, offsprings tended to underestimate the frequency of 'house cleaning' and 'kitchen work' and overestimate the height of them. In conclusion, the life style questionnaire was found to be reliable in the most of items. But, some items related with physical activity were found to be somewhat less reliable. Sexual difference on the validity was identified in

---

1) 본 연구는 1995년도 서울대학교병원 일반연구비 지원에 의해 이루어진 것임

some items. With regard to bias of proxy respondents, offsprings tended to have bias in part of items of housework and anthropometry.

**Key words** : elderly Korean, life-style, questionnaire, reliability, validity

## I. 서 론

우리 나라 노인인구의 증가 속도는 세계의 어느 나라보다도 빠른 것으로 알려져 있으며(송건용, 1994) 의료보험수진자료에서도 노인의 진료건수는 급속도로 증가하고 있다(이가옥 등, 1994). 따라서 최근 들어 노인인구를 대상으로 한 의학연구들이 활발히 추진되고 있다. 노인들을 대상으로 의학연구를 수행하는 데 있어서 질병의 발생뿐 아니라 그 위험요인 혹은 교란변수인 흡연, 음주, 육체적 활동량, 식습관 등의 생활습관을 정확하게 파악하는 것이 매우 중요하다. 이러한 생활습관을 파악하는 것은 직접 관찰하는 것이 가장 이상적인 방법이나 연구대상이 대규모 인구집단일 때는 현실적으로 불가능하므로 설문조사를 통하여 간접적으로 필요한 정보를 얻을 수 밖에 없다. 설문서는 응답자와 설문조사자 및 설문지 자체의 특성에 의하여 신뢰도와 타당도가 좌우되며 이것이 검정되지 않으면 연구에 사용할 수 없다. 외국의 경우 노인인구에서의 성, 연령, 응답자의 종류, 설문서의 항목별로 신뢰도와 타당도에 관한 많은 연구가 있으나 우리나라에서는 노인인구를 대상으로 한 설문서의 신뢰도 평가는 노인성 치매 등의 특수한 분야에서 일부 이루어졌을 뿐(박종한 등, 1995), 일상생활습관에 관한 설문서의 신뢰도 연구는 거의 이루어진 바가 없다. 노인의 경우 연구대상이 치매, 이환, 고령 등의 문제로 인하여 청장년에 비하여 설문서의 신뢰성을 확보하기가 어려우므로 신뢰도 및 타당도연구가 더욱 필요한 계층이다. 또한 노인의 경우 본인이 직접 응답하지 않고 자녀, 배우자 등의 대리응답자가 응답하는 경우가 있는데 이로 인한 신뢰도의 저하나 비틀림(bias)의 문제는 현재까지 국내에서는 연구된 바가 없다. 노인을

대상으로 설문조사를 시행할 때에는 불가피하게 대리응답자를 사용할 수 밖에 없는 경우가 많으므로 대리응답자에서의 신뢰도와 비틀림여부를 파악하는 것은 노인인구를 대상으로 올바른 의학적 연구를 수행하는데 필수적이다.

따라서 본 연구는 65세 이상의 노인 인구를 대상으로 생활습관 전반에 관한 설문서의 신뢰도를 파악하며, 파악된 신뢰도로부터 타당도를 추정하여 그것의 응답자별, 성별차이를 파악하고, 대리응답자가 가질 수 있는 비틀림을 파악함으로써 노인인구에서의 생활습관 설문서의 유용성을 평가하는데 그 목적이 있다.

## II. 연구방법

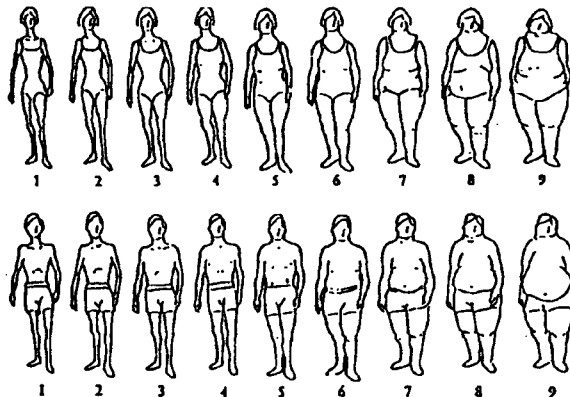
부산지역의 공무원 및 사립학교교직원 의료보험관리공단의 피보험자 및 피부양자중 65세이상인 16,524명을 대상으로 생활습관에 관한 설문서를 발송하여 이 중 9,139명으로부터 설문서를 회수하였다. 설문서의 신뢰도 및 타당도를 평가하기 위해서 그 중 200명을 단순무작위추출하여 3개월 후에 다시 설문서를 발송하여 112명으로부터 설문서를 회수하였다. 설문서의 항목의 대부분이 응답결손인 2건은 제외하여 110명이 최종분석대상이 되었다. 연구대상의 연령은 65세에서 91세(평균 71.9세)의 범위였으며, 남자가 24명 여자가 86명으로 여자가 많았다. 설문서의 내용으로는 건강상태, 약물복용, 흡연, 음주, 식생활습관, 수면시간, 육체적 활동, 인적사항 등의 총 47개 세부문항으로 구성되었다(표 1).

설문서 자체의 신뢰도를 파악하기 위하여 1차와 2차 모두 본인이 응답한 경우에 제한하여 신뢰도 계수를 산출하였으며 대리응답의 신뢰도를 파악하기 위하

**표 1. 본 연구에서 사용된 노인 생활습관 설문서의 구성 및 항목수**

범주	설문내용
주관적 건강상태	자신의 건강상태는 어떻다고 느끼십니까?
활동상태	평소에 혼자서 걷거나 활동하시는데 불편함은 있습니까?
흡연습관	담배를 피우십니까? 모두 몇 년 동안 담배를 피우십니까? 하루에 얼마나 담배를 피우십니까? (※갑, 개피) 3년 전에는 담배를 피우셨습니까? 3년 전에는 하루에 평균 얼마나 피우셨습니까?
음주습관	평소에 술을 드십니까? 몇 년 전에 끊으셨습니까? 모두 몇 년 동안 술을 마셨습니까? 1 주일에 평균 몇 번이나 마셨습니까? (※술 종류별 응답표를 제시) 3년 전에 술을 얼마나 드셨습니까? (※술 종류별 응답표를 제시)
식생활습관	음식의 간을 어떻게 드시는 편이십니까? 매운 음식을 잘 드시는 편입니까? 채식을 주로 드십니까, 육식을 주로 드십니까?
수면습관	하루에 평균 몇 시간씩 주무십니까? 잠을 못 이루시거나 자다가 새벽에 깨시는 경우가 있습니까?
집밖에서의 활동	운동을 얼마나 하시는지 알아보려고 합니다? (※운동종류별 응답표를 제시) 산보, 약수터가기, 자전거타기, 게이트 볼, 골프, 달리기, 정구, 등산, 기타
집안에서의 활동	운동을 얼마나 하시는지 알아보려고 합니다? (※운동종류별 응답표를 제시) 정원일, 청소, 손빨래, 아기보기, 부엌일, 맨손체조, 기타
체격	현재의 키와 몸무게는 얼마나 되십니까? 다음의 그림*중에서 현재 자신의 몸매과 가장 가깝다고 생각되는 그림에 동그라미 표시를 해주십시오.
산과력	총 몇 명의 아이를 낳았습니까? (※ 이상 등의 6 문항)

\*: 그림 1.을 참고



**그림 1. 본 연구에 사용된 체격그림(Must 등, 1993 ;배종면 등, 1996)**

여 한번은 본인이, 다른 한번은 대리응답자가 응답한 경우에 제한하여 신뢰도를 산출하고 타당도계수와 그 신뢰구간을 추정하였다. 신뢰도 계수는 명목척도 및 두 개의 범주로 구성된 순위척도에서는 kappa 값을 구하였으며 세 개 이상의 범주로 구성된 순위척도에서는 weighted kappa 값을 사용하였고 연속변수에 대해서는 Pearson 상관계수를 구하였다. 신뢰도 계수( $\kappa$  값, weighted  $\kappa$  값, r 값)를 이용하여 타당도 계수(validity coefficient)를 추정하는 데 있어서 첫 번째 test( $X_1$ )의 오차와 두 번째 test( $X_2$ )의 오차의 분산이 동일하며, 이 오차가 참값과 상관관계가 없으며, 첫 번째 test의 오차와 두 번째 test의 오차가 상관관계가 없을 때 이를 참값에 대한 parallel tests 라고 하며 이런 경우 타당도 계수는 다음과 같이 산출하였다(Bruce 등, 1992; Allen 등, 1979).

$$\rho_{TX_1} = \rho_{TX_2} = \sqrt{\rho_{X_1X_2}}$$

$\rho_{TX_1}$  : test  $X_1$ 의 타당도 계수  
 $\rho_{TX_2}$  : test  $X_2$ 의 타당도 계수  
 $\rho_{X_1X_2}$  : test  $X_1, X_2$ 의 상관계수

혹은, 첫 번째 test의 오차가 두 번째 test의 오차보다 분산이 작으며, 즉 두 test의 정밀도(precision)에 차이가 있으며 이 오차가 참값과 상관관계가 없고, 첫 번째 test의 오차와 두 번째 test의 오차가 상관관계가 없을 때 타당도 계수는 다음과 같이 산출한다.

$$\rho_{X_1X_2} < \rho_{TX_1} < \sqrt{\rho_{X_1X_2}}$$

$\rho_{TX_1}$  : test  $X_1$ 의 타당도 계수  
 $\rho_{X_1X_2}$  : test  $X_1, X_2$ 의 상관계수

본 연구에서는 이상의 방법을 사용하여 설문대상자의 응답자양상별, 성별로 신뢰도를 이용하여 타당도를 추정하였으며 각 항목에서 타당도의 차이를 파악하였다. 대리응답자의 설문결과에 존재할 수 있는 비뮌림을 파악하기 위하여 한번은 본인이 다른 한번은 대리응답자가 응답한 경우에 한하여 비연속변수에 대하여는 윌콕슨 부호화 순위검정, 연속변수에 대하여는

paired t-test를 사용하여 쌍체검정을 실시하였다. 대리응답자의 대부분이 자녀이었고, 대리응답자별로 그 비뮌림의 양상이 다를 수 있으므로 쌍체검정에 사용된 대리응답자는 자녀로 제한하였다.

설문서의 응답자별 분포는 1차 응답에서는 본인이 39명으로 제일 많았고 며느리를 포함한 자녀 35명, 손자 3명, 기타 33명이었다. 2차 응답에서는 자녀가 44명 본인이 29명 배우자가 1명, 손자 1명, 기타가 35명이었다. 응답자를 본인응답과 대리응답으로 구분하면 표 2와 같았다.

표 2. 1차 설문응답자와 2차 설문응답자의 응답자별 분포

단위 : 명(%)

	2차 설문		계
	본인응답	대리응답	
1차 설문 본인응답	18 (16.4)	21 (19.1)	39 (35.5)
대리응답	11 (10.0)	60 (54.5)	71 (64.5)
계	29 (26.4)	81 (73.6)	110 (100)

### Ⅲ. 연구결과

설문지의 항목별 신뢰도를 파악하기 위하여 1차 응답과 2차 응답에서 본인이 직접 응답한 18명에 대하여 비연속 변수에서는 weighted kappa 값을, 연속변수에서는 피어슨의 상관계수를 구하여 신뢰도 계수를 구하였다(표 3). 대부분의 항목이 0.6 이상의 신뢰도 계수를 가졌으나, ‘산보’(0.33), ‘정월일’(0.28), ‘맨손체조’(0.31), ‘활동상태’(0.33) 등과 같은 육체적 활동과 관계 있는 항목의 일부에 있어서는 낮은 신뢰도를 보였다. 반면 현재의 ‘현재까지의 흡연여부’(0.82), ‘현재까지의 흡연양’(0.84), ‘3년전의 흡연여부’(1.00), ‘지난 1년간의 음주량’(0.97), ‘3년전의 음주량’(0.86) 등의 흡연 및 음주와 관련있는 항목과 ‘키’(0.94), ‘몸무게’(0.88), ‘체격그림’(0.97)과 같은 체격과 관련있는

**표 3. 본인-본인 응답의 항목별 신뢰도**

신뢰도 계수	항 목
0.21 - 0.40	활동상태 산보, 정원일, 맨손체조
0.41 - 0.60	3년전의 음주여부 숙면여부 손빨래
0.61 - 0.80	주관적 건강상태 수면시간*, 야간수면시간* 3년전의 흡연양* 음주여부 짜게 먹는 정도, 맵게 먹는 정도, 육식선호정도 약수터가기, 청소, 부엌일 첫아이를 낳은 나이*, 막내를 낳은 나이*
0.81 - 1.00	몸무게*, 키*, 체격그림 현재까지의 흡연여부 및 흡연양*, 3년전의 흡연상태 지난 1년간의 음주량*, 3년전의 음주량* 주간수면시간*, 폐경나이*, 낳은 아이의 수*

\* : 피어슨의 상관계수

항목, ‘폐경나이’(0.99), ‘낳은 아이의 수’(0.82)와 같은 산부인과력의 일부항목, 등에서는 높은 신뢰도를 보였다.

한번은 본인이 응답하고 다른 한번은 대리응답자가 응답을 한 32명에 대하여 이상과 같은 방법으로 신뢰도 계수를 산출하였다(표 4). ‘청소’(0.28), ‘맨손체조’(0.33), ‘손빨래’(0.39) 등의 육체적 활동과 관련있는 일부의 변수에서 낮은 신뢰도를 보였다. 특히 ‘청소’ 항목에서는 본인이 두 번 응답한 경우의 0.68에 비하여 상당히 낮은 신뢰도를 보였다. ‘주관적 건강상태’도 본인이 두 번 모두 응답한 경우의 0.72에 비하여 본인-대리응답에서 0.32로 신뢰도의 저하가 심하게 나타났다. ‘첫아이를 낳은 나이’ 역시 신뢰도 계수가 0.70에서 0.33으로 떨어져 큰 차이를 보였다. 반면 흡연 및 음주와 관련있는 대부분의 변수에서 0.6이상의 높은 신뢰도를 보였다. ‘키’(0.79), ‘몸무게’(0.81), ‘체격그림’(0.91)과 같이 체격과 관련된 항목에서도 0.6 이상의 신뢰도를 보였으나 ‘키’의 경우 본인이 두 번

**표 4. 본인-대리응답간의 항목별 신뢰도**

신뢰도 계수	항 목
0.21 - 0.40	주관적 건강상태 청소, 맨손체조, 손빨래 첫아이를 낳은 나이
0.41 - 0.60	활동상태 정원일, 부엌일 맵게 먹는 정도, 육식선호정도 야간수면시간
0.61 - 0.80	현재까지의 흡연양 음주여부 짜게 먹는 정도 수면시간, 주간수면시간 산보 키* 폐경나이*, 낳은 아이의 수*, 막내를 낳은 나이*
0.81 - 1.00	약수터 가기, 현재까지의 흡연여부, 3년전의 흡연상태 및 흡연양 지난 1년간의 음주량, 3년전의 음주여부 및 음주량 몸무게*, 체격그림

\* : 피어슨의 상관계수

응답한 경우의 신뢰도 0.94에 비하여 상대적으로 본인-대리응답의 경우 낮은 신뢰도(0.79)를 보였다.

표5는 Allen 등(1979)의 방법을 사용하여 신뢰도를 이용한 타당도 추정의 결과 중 응답의 양상별로 타당도의 유의한 차이가 있는 항목들이다. 식생활과 관련이 있는 항목의 일부에서 본인-대리응답의 타당도가 상대적으로 낮게 나타났으며 육체적 활동과 관련된 항목 중 ‘청소’ 항목에서는 본인-본인 응답보다 본인-대리응답에서 타당도가 낮았으나 산보, 정원일에서는 오히려 본인-대리응답에서 더 타당도가 높게 나타났다. 대리응답자의 타당도가 본인이 두 번 응답한 경우와 비슷하게 높게 나타나서 대리응답으로 인한 타당도의 저하가 없다고 생각되는 항목은 표6과 같다. 흡연여부, 흡연량 등의 흡연과 관련된 항목과 음주여부, 음주량 등의 음주에 관련된 항목에서 대리응답자의 응답이 비교적 타당도가 높다고 나타났다. 키의 타당도는 본인에 비하여 대리응답자에서의 타당도가 조금 낮았으나 체격그림에서 대리응답자의 타당도는 본인

표 5. 본인-본인 응답과 본인-대리응답의 타당도 추정에서 차이가 나는 항목

항 목	타당도계수*		
	본인-본인(N=18)	본인-대리(N=32)	
대리응답자에서 낮은 타당도를 보이는 항목	육식선호정도	0.76~0.87	0.56~0.75
	맷게먹는 정도	0.69~0.83	0.40~0.63
	야간수면	0.76~0.87	0.53~0.73
	첫아이를 낳은나이	0.70~0.84	0.33~0.57
	주관적 건강상태	0.72~0.85	0.32~0.57
	청소	0.68~0.82	0.28~0.53
대리응답자에서 높은 타당도를 보이는 항목	산보	0.33~0.57	0.62~0.79
	정원일	0.28~0.53	0.54~0.73
	3년전 음주여부	0.44~0.66	0.87~0.93

\* :  $\rho_{X_1X_2} < \rho_{TX_1} < \sqrt{\rho_{X_1X_2}}$   $X_1X_2$

표 6. 대리응답자의 타당도가 본인응답과 비교하여 비슷하거나 높게 나타난 항목

항 목	타당도계수*	
	본인-본인(N=18)	본인-대리(N=32)
흡연여부	0.82~0.91	0.97~0.98
흡연양	0.84~0.92	0.80~0.89
3년전의 흡연여부	1.00~1.00	0.97~0.98
3년전의 흡연양	0.63~0.79	0.96~0.98
음주여부	0.65~0.81	0.68~0.82
음주량	0.97~0.98	0.85~0.92
3년전의 음주여부	0.44~0.66	0.87~0.93
3년전의 음주량	0.86~0.93	0.84~0.92
몸무게	0.88~0.94	0.81~0.90
키	0.94~0.97	0.79~0.89
체격그림	0.97~0.98	0.91~0.95
총 수면시간	0.60~0.77	0.76~0.87
주간수면시간	0.84~0.92	0.72~0.85
막내를 낳은 나이	0.77~0.88	0.76~0.87
낳은 아이의 수	0.82~0.91	0.75~0.87

\* :  $\rho_{X_1X_2} < \rho_{TX_1} < \sqrt{\rho_{X_1X_2}}$

응답과 비슷하게 높게 나타났다.

성별에 따라서 타당도를 추정된 결과, 식생활과 관련된 항목에서는 여자보다 남자에서 타당도가 높았고 ‘약수터 가기’, ‘정원일’, ‘산보’ 등의 일일 활동과 관련된 항목과 음주, 흡연 등의 항목에서는 여자에서 타

당도가 높았다. 그 외의 항목에서는 뚜렷한 차이가 없었다(표7).

대리응답자가 가질 수 있는 비뚤림을 파악하기 위하여 본인과 대리응답자가 번갈아 응답한 경우에 한하여 비연속변수에 대하여는 윌콕슨 부호화 순위검정, 연속변수에서는 paired t-test를 실시하였다(표8). 번갈아 응답한 경우는 총 32명이나 이중에서 대리응답자가 배우자, 손자, 미기재의 경우는 그 수가 적고, 대부분을 차지하는 자녀의 대리응답양상을 정확히 파악하기 위하여 자녀외의 대리응답자의 경우는 분석에서 제외하여 비뚤림을 파악하기 위한 쌍체분석의 대상자는 19명이었다. 대부분의 항목에서 본인과 자녀의 응답사이에는 통계적으로 유의한 차이를 보이지 않았으나 ‘청소’ 항목(p-value : 0.04)에서 본인의 응답이 자녀의 응답보다 청소를 더 자주하는 것으로 나타났다. ‘키’ 항목(p-value : 0.03)에서는 자녀의 응답보다 본인의 응답이 키가 더 작은 것으로 나타났고 통계적 유의성이 약하기는 하지만 ‘부엌일’(p-value : 0.06)에서 본인응답이 자녀의 응답보다 부엌일의 빈도가 더 많은 것으로 나타났다.

표 7. 추정된 타당도가 성별에 따라 차이를 보이는 항목

항 목		타당도계수*	
		남자(N=24)	여자(N=86)
남자가 여자보다 타당도가 높게 나타난 항목	짜게 먹는 정도	0.86 ~ 0.93	0.57 ~ 0.76
	낮잠시간	0.79 ~ 0.89	0.55 ~ 0.74
	육식 채식 선호정도	0.79 ~ 0.89	0.49 ~ 0.70
	뭇게 먹는 정도	0.73 ~ 0.85	0.51 ~ 0.72
	등산	0.53 ~ 0.73	0.03 ~ 0.18
여자가 남자보다 타당도가 높게 나타난 항목	3년전의 음주여부	0.53 ~ 0.73	0.86 ~ 0.93
	3년전의 흡연양	0.49 ~ 0.70	0.80 ~ 0.89
	약수터 가기	0.48 ~ 0.70	0.73 ~ 0.85
	흡연양	0.47 ~ 0.68	0.94 ~ 0.97
	정원일	-0.02**	0.47 ~ 0.69
	산보	-0.04**	0.48 ~ 0.69

\* :  $\rho_{X_1X_2} < \rho_{TX_1} < \sqrt{\rho_{X_1X_2}}$

\*\* : validity could not be estimated because of negative value

표 8. 본인과 자녀가 번갈아 응답한 경우의 쌍체분석

	본인		자녀		p-value
	mean	mode	mean	mode	
청소	2.7	4	1.7	0	<0.05*
부엌일	2.4	4	1.7	0	0.06*
	mean	s.d.	mean	s.d.	
키	155.7	6.4	158.0	7.0	<0.05**

\*: 윌콕슨 부호화 순위 검정

\*\* : paired t-test

#### IV. 고 찰

본 연구에 사용된 설문서의 본인-본인 응답의 신뢰도에서 대부분의 항목이 0.6이상의 신뢰도를 보였다. 특히 흡연 및 음주와 관련된 항목, 체격지수와 관련된 항목에서는 0.8이상의 높은 신뢰도를 보였다. 그러나 육체적 활동과 관련있는 일부 변수에서 특히 낮은 신뢰도를 보였는데 여기에는 두 가지의 가능한 원인이 있을 것으로 생각된다. 첫째, 설문서 자체의 신뢰성의 문제이며 이러한 문제는 보다 신뢰성 높은 설문서 혹

은설문방법의 개발을 통하여 해결할 수 있다. 둘째, 설문과 재설문사이의 기간은 일반적으로 1개월 정도가 적당하다고 추천되고 있으나 본 연구의 검사-재검사 간격은 4.5개월로 비교적 길었으며 또한 첫째, 설문서 자체의 신뢰성의 문제이며 이러한 문제는 보다 신뢰성 높은 설문서 혹은 설문방법의 개발을 통하여 해결할 수 있다. 둘째, 설문과 재설문사이의 기간은 일반적으로 1개월 정도가 적당하다고 추천되고 있으나 본 연구의 검사-재검사 간격은 4.5개월로 비교적 길었으며 또한 연구대상이 고령이며 이환 및 여러 가지 요인에 의하여 육체적 활동에 변화가 많은 점을 고려하면 그 기간에 실제로 육체적 활동에 변화가 있었을 가능성도 배제할 수 없다. 반면 흡연이나 음주와 관련된 항목의 신뢰도는 높게 평가되었으며 이 항목에서는 본인과 대리응답자가 교대로 응답한 경우에도 여전히 높은 신뢰도를 보여주었다.

McLaughlin(1987) 등의 National Health and Nutrition Examination Survey(NHANES I)에 바탕을 둔 대리응답자의 응답양상에 대한 연구에 의하면 다른 항목에 비하여 흡연여부와 흡연량에서의 신뢰도가 높

게 나타났다고 하였는데 이러한 결과가 우리나라에서도 확인되었다고 하였다. 본인과 대리응답자가 교대로 응답한 경우 신뢰도계수가 0.6 이하인 항목이 11개로 본인이 두 번 응답한 경우의 7개에 비하여 조금 많았다. 11개항목중 육체적 활동과 관련된 항목이 6개로 절반을 차지하여 육체적 활동에 대한 항목에서 특히 대리응답과 본인응답의 차이가 많음을 알수 있었다.

신뢰도를 이용한 타당도 추정결과에서 대리응답자의 타당도는 본인의 타당도에 비하여 6개의 항목에서 낮게 나타났으며 육체적 활동과 관련있는 변수중 '산보'와 '정원일'에서는 대리응답자의 타당도가 오히려 높게 나타나서 대리응답자로 인한 타당도의 저하는 심각하지 않은 것으로 판단되었다. '키' 항목의 타당도는 본인에 비하여 대리응답자의 타당도가 낮게 나타났으나 '체격그림'은 본인의 경우와 비슷하게 높은 타당도를 보였다. 따라서 노인인구에서의 비판도를 대리응답자를 통한 설문서로 평가하고자 할 때 키와 몸무게를 직접 물어보는 것 외에도 체격그림을 활용하는 것이 더 타당한 결과를 얻을 수 있을 것이다.

노인인구를 대상으로 한 설문조사에서 남성보다 여성에서 신뢰도가 높다는 연구가 있다(Boult 등, 1994). 본 연구에서는 신뢰도로부터 타당도를 추정하여 성에 따른 차이를 파악하였다. 여성에서 남성보다 '3년전의 음주여부', '현재까지의 흡연양', '3년전의 흡연양', '약수터 가기', '정원일', '산보'에서 타당도가 높았으며, 남성에서 여성보다 '짜게 먹는 정도', '맵게 먹는 정도', '육식 채식선호 정도', '낮잠시간', '등산'에서 타당도가 높게 나타났다. 남성의 경우 식생활 습관의 세가지 항목에서 여성보다 타당도가 높게 나타난 것과 음주와 흡연의 일부항목과 육체적 운동과 관련된 일부항목에서는 남성보다 여성에서 타당도가 높게 나타난 것은 우리나라의 문화적 특징에 기인하였을 가능성이 있으며 이러한 부분에 대하여는 추가적인 연구가 필요할 것이다.

역학적 연구에서는 신뢰도뿐만 아니라 응답에 비뚤림이 존재하는 지를 파악하는 것도 중요하다. 대리응답자의 사용은 연구대상의 확보를 용이하게 하며 따

라서 대표성에 있어서의 장점을 가지고 있지만 비뚤림의 존재 가능성으로 인하여 그 사용에 주의를 요한다. 예를 들면 대리응답자는 직업력에 대해서 연구대상 본인보다 덜 보고하는 것으로 알려져 있다(Blot 등, 1978). 배우자 특히 부인들은 남편의 흡연과 커피음용을 과대보고하며(Kolonel 등, 1977), 수동흡연은 적게 보고한다고 한다(Cummings 등, 1989). 남자들은 자신의 아내에 대하여 음식섭취량을 적게 보고하나, 여자들이 보고한 남편의 음식섭취량에서는 그러한 경향을 발견하지 못했다고 한다(Humble 등, 1984). 반면 대리응답자와 본인응답사이의 차이에 대한 다른 연구에서는 의료이용력과 식이조사에서 뚜렷한 차이가 없었다는 보고도 있다(Herrmann, 1985). 외국의 경우 대리응답자는 노인의 육체적 기능을 과소평가하는 경향이 있다는 연구가 있으며(Dorevitch 등, 1992) 흡연양, 가족내 흡연자수, 커피섭취량, 일회음식섭취량에서 비뚤림이 존재한다는 보고가 있다(Bruce 등, 1992). 본 연구에서는 '청소'와 '부엌일'과 같은 가사노동에서 본인의 응답보다 자녀의 응답이 더 빈도가 낮게 나타났으며 '키' 항목에서 본인보다 자녀의 응답이 더 크게 나타났다. 그 외 흡연과 음주 및 식생활 습관과 관련하여서는 본인과 자녀사이의 유의한 차이는 확인할 수 없었다.

본 연구의 제한점으로는 첫째, 110명에 대하여 검사-재검사 방법을 사용한 설문서의 신뢰도 평가 연구로서 그 연구대상수가 설문서의 신뢰도를 각 항목별, 성별, 대리응답자별, 연령별로 세분화하여 파악하기에는 충분히 크지 않았다. 둘째, 검사와 재검사의 기간에는 1개월 정도의 시간간격이 필요하다고 알려져 있으나(Bruce 등, 1992), 본 연구의 검사와 재검사와의 간격은 약 4-5개월정도로서 추정된 재현도와 타당도는 실제보다 과소평가되었을 것이다. 셋째, 연구의 주목적이 대리응답자의 응답양상을 파악하기 위하여 설계된 것이 아니었으므로, 대리응답자의 응답특성의 파악에 한계가 있었다. 즉 대리응답자의 응답양상을 제대로 파악하기 위해서는 동일한 시점에서 동일한 방법으로 본인과 대리응답자 양군에서 자료를 수집하여야 하는



데 본 연구에서는 시간적으로 상당히 차이가 나고 모든 연구대상에서 대리응답자의 자료를 수집하지 못하였다.

본 연구결과로서 노인을 대상으로 한 생활습관 설문서의 신뢰도는 전반적으로 우수하여 역학적 조사에 사용하기에 적합한 것으로 나타났다. 그러나 육체적 활동량과 관련된 항목들은 비교적 신뢰도가 낮게 나타났다으므로, 육체적 활동량을 측정하기 위하여 보다 신뢰도 높은 설문서의 개발이 필요할 것이다.

## V. 요약

부산지역의 공무원 및 사립학교 교직원 의료보험관리공단의 피부양자중 65세 이상인 16,524명에 대하여 생활습관에 대한 설문서를 발송하여 이중 9,139명으로부터 설문서를 회수하였다. 본 연구는 여기에 사용된 설문서의 신뢰도를 평가하기 위하여 이중 200명을 단순 무작위 표본추출하여 다시 설문서를 발송하였으며 회수된 설문서 110건에 대하여 항목별, 응답자별 신뢰도를 평가하였다.

1) 본인이 두 번 모두 응답한 경우에 있어서 대부분의 항목이 0.6 이상의 신뢰도 계수를 가졌으나, '산보', '정원일', '맨손체조', '활동상태'와 같은 육체적 활동과 관계 있는 항목의 일부에 있어서는 낮은 신뢰도를 보였다. 반면 현재의 '현재까지의 흡연여부', '현재까지의 흡연양', '3년전의 흡연여부', '지난 1년간의 음주량', '3년전의 음주량' 등의 흡연 및 음주와 관련된 항목과 '키', '몸무게', '체격그림'과 같은 체격과 관련있는 항목, '폐경나이', '낳은 아이의 수'와 같은 산부인과력의 일부항목 등에서는 높은 신뢰도를 보였다.

2) 한번은 본인이 다른 한번은 대리응답자가 응답한 32건에 대하여 신뢰도 계수를 계산하였다. 두 번 다 본인이 응답한 경우에 비하여 신뢰도 계수가 크게 하락한 항목은 '청소', '주관적 건강상태', '첫아이를 낳은 나이', '폐경나이', '땀계 먹는 정도', '육식선호 정도', '야간수면시간' 등이었으며 본인이 두 번 응답

한 경우에 비하여 본인-대리응답에서 신뢰도계수가 더 높게 나타난 항목은 '3년전의 음주여부', '3년전의 흡연량', '산보', '정원일', '활동상태' 등이었다.

3) 응답자의 성별 타당도 분석결과 식생활과 관련된 항목에서는 여자보다 남자에서 타당도가 높았고, '약수터 가기', '정원일', '산보' 등의 일일 활동과 관련된 항목과 음주, 흡연 등의 항목에서는 여자에서 타당도가 높았다.

4) 대리응답자가 가질 수 있는 비뚤림을 파악하기 위하여 쌍체분석을 실시하였다. 분석대상은 본인과 자녀가 번갈아 응답한 19건에 제한하였다. '청소'와 '부엌일'에서 본인의 응답이 자녀의 응답보다 빈도가 더 높은 것으로 나타났다. '키' 항목에서는 자녀보다 본인의 응답이 키가 더 작은 것으로 나타났다.

## 참고문헌

- 박종한, 조성완, 손형석. 노인기능 상태 평가의 신뢰도. 신경정신의학 1995 ; 34(2) : 475-481
- 배종면, 안윤옥. 비만도 측정 역학조사도구로서의 신체외형도 유용성 평가. 한국역학회지 1996 ; 18(1) : 101-7
- 송건용. 노인인구의 증가에 따른 사회학적 접근. 대학의학협회지 1994 ; 37(10) : 1148
- 이가옥, 등. 노인부양에 관한 연구. 한국보건사회연구원, 1990, 쪽 52
- Allen MJ, Yen WM. Introduction to Measurement Theory. Monterey, Brook/Cole, 1979, pp. 1-117
- Blot WJ, Harrington JM, Toledo A, et al. Lung cancer after employment in shipyards during World War II. NEJM 1978 ; 299 : 620-4
- Boult L, Boult C, Pirie P, Pacala JT, et al. Test-retest reliability of a questionnaire that identifies elders at risk for hospital admission. J Am Geriatr Soc 1994 ; 42(7) : 707-11
- Bruce KA, Emily W, Rodolfo S. Principles of exposure measurement in epidemiology. Oxford Univ. Press 1992, pp. 78-114
- Cummings KM, Markello SJ, Mahoney MC, et al. Measurement of lifetime exposure to passive smoke. Am J Epidemiol 1989 ; 130 : 122-32
- Dorevitch MI, Cossar RM, Bailey FJ, et al. The accu-

- racy of self and informant ratings of physical functional capacity in the elderly. *J Clin Epidemiol* 1992 ; 45(7) : 791-8
- Gloth FM 3rd, Walston J, Meyer J, Pearson J et al. Reliability and validity of the Frail Elderly Functional Assessment questionnaire. *Am J Phys Med Rehabil* 1995 ; 74(1) : 45-53
- Herrman N. Retrospective information from questionnaire I. Comparability of primary respondents and their next-of-kin. *Am J Epidemiol* 1985 ; 121 : 937-47
- Humble CG, Samet JM, Skipper BE. Comparison of self- and surrogate-reported dietary information. *Am J Epidemiol* 1984 ; 119 : 86-98
- Jackson N, Little J, Wilson AD. Comparison of diet history interview and self completed questionnaire in assessment of diet in an elderly population. *J Epidemiol Community Health* 1990 ; 44 (2) : 162-9
- Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. *Am J Epidemiol* 1977 ; 106 : 476-84
- McLaughlin JK, Dietz MS, Mehl ES, et al. Reliability of surrogate information on cigarette smoking by type of informant. *Am J Epidemiol* 1987 ; 126 : 144-6
- Must A, Willett WC, Dietz WH. Remote recall of childhood height, weight, and body build by elderly subjects. *Am J Epidemiol* 1993 ; 138 : 56-64
- Pickle LW, Brown LM, Blot WJ. Information available from surrogate respondents in case-control interview studies. *Am J Epidemiol* 1983 ; 118 : 99-108
- Rivero E, Carrillo E, Portella E, et al. Evaluation of the design and reliability of a questionnaire to identify elderly individuals in need for social and health services. *Med Clin (Barc)* 1993 ; 101 (18) : 688-92
- Rocca WA, Fratiglioni L, Bracco L, et al. The use of surrogate respondents to obtain questionnaire data in case-control studies of neurologic disease. *J Chr Disease* 1986 ; 39 : 907-12
- Rogot E, Reid DD. The validity of data from next-of-kin in studies of mortality among migrants. *Int J Pub Health* 1975 ; 4 : 51-54
- Stein AD, Lederman RI, Shea S. The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample. *Am J Public Health* 1993 ; 83 : 1768-72
-