

한국판 국제 소아천식 및 알레르기 질환 연구 설문지의 신뢰도 및 타당도 연구

최성우, 주영수, 김대성, 김재용, 권호장¹, 강대희, 이상일², 조수현

서울대학교 의과대학 예방의학교실, 단국대학교 의과대학 예방의학교실¹,
성균관대학교 의과대학 소아과학교실²

= Abstract =

Reliability and Validity of the Korean Version of ISAAC Questionnaire

Seong-Woo Choi, Yeong-Su Ju, Dae-Sung Kim, Jae-Yong Kim, Ho-Jang Kwon¹,
Dae-Hee Kang, Sang-II Lee², Soo-Hun Cho

Dept. of Preventive Medicine, Seoul National University College of Medicine,

Dept. of Preventive Medicine, Dankook University College of Medicine¹,

Dept. of Pediatrics, Sungkyunkwan University College of Medicine²

Recent increases of asthma and allergies in childhood made the need for a standardized approach to international and regional comparisons of their prevalence and severity. To address these issues, "International Study of Asthma and Allergies in Childhood (ISAAC)" is currently underway. In Korea, "Nationwide Study of Asthma and Allergies in Korean Children" began in 1995 according to ISAAC protocol. ISAAC written and video questionnaires were used in this survey, but their reliability and validity were not evaluated properly yet.

In this study, our aim was to evaluate the reliability and validity of two kinds of questionnaires and their usefulness in international and regional comparisons. The test and retest of two questionnaires were completed by male(n=110) and female(n=111) middle school students with two and three weeks interval each. Kappa(or weighted kappa) were calculated from each questions and validity coefficients were estimated from those statistics.

In Korean version of written questionnaire, the questions for allergic rhinitis, atopic dermatitis, allergic conjunctivitis, and food allergy proved to have high kappa values (or weighted kappa values) and validity coefficients and they can be used in further

studies without any correction. But some questions about asthma(especially nocturnal cough, wheezing in exercise, and severe asthma) and drug allergy need to be revised for better understanding to study subjects. Video questionnaire has the same degree of reliability and validity when compared to written questionnaire and this is the unexpected result. Accordingly, it also need to be revised to overcome the racial and cultural differences of the study subjects. In conclusion, the Korean version of written and video questionnaires may be considered to be useful methods in international and regional comparisons of asthma and allergic diseases in childhood after correction of some questions.

Key words : ISAAC, questionnaire, reliability, validity

I. 서 론

최근 들어 서구 여러 나라들과 개발도상국들에서 천식과 알레르기 질환으로 인한 사망률과 이환율이 증가하고 있다. 나라들마다 증가하는 정도에는 차이가 있으며, 나라들 간에, 또한 한 나라내에서도 지역 간 천식 및 알레르기 질환의 유병률이 상당한 차이를 보이고 있다(Shaw 등, 1995). 천식 및 알레르기 질환의 원인으로 유전적 요인과 환경적 요인이 중요하게 작용한다는 것은 밝혀져 있지만, 상당수의 연구에도 불구하고 지역간의, 그리고 시간적 유병률의 변동을 일으키는 기전에 대해서는 명확하게 알지 못하고 있는 실정이다(Asher 등, 1995).

인구집단 단위에서의 알레르기 질환의 유병률에 영향을 미치는 인자들에 대한 연구는 알레르기 질환의 감수성이 높은 사람들을 대상으로 하는 개인단위의 연구에 비해서 여러 가지 어려움으로 인해 수행이 활발하지 못했던 것이 사실이다. 이러한 인구집단 단위의 역학연구를 수행하는데 있어서 주요한 장애물의 하나는 인구집단간에 천식과 알레르기 질환의 유병률과 중증도를 비교하는 적절한 방법이 없었다는 것이다. 다른 하나는 비교가능한 자료를 얻고 분석하는데 있어 여러지역을 연결하고 조절하는 협력 프로그램이 없었다는 것이다. 그러한 지역간의, 그리고 한 지역내에서의 시계열적 비교는 개인단위의 연구에서 간과되

기 쉬운 인구집단 수준에서 작용하는 질병의 결정인자를 밝히는데 가치있는 지표를 산출할 수 있다. 그러한 자료를 얻을 목적으로 '국제 소아천식 및 알레르기 질환 연구(International Study of Asthma and Allergies in Childhood; ISAAC)'가 1989년에 처음으로 시작되었다(Asher 등, 1995). 우리나라에서도 대한 소아 알레르기 및 호흡기학회 주관으로 1993년부터 ISAAC에 참여하여 '한국 어린이, 청소년의 알레르기 질환에 관한 전국적 역학조사'가 진행 중에 있다.

ISAAC 연구수행 지침서에 의하면 연구는 3단계로 진행되는데, 그 중 1단계에서는 6~7세의 어린이와 13~14세의 청소년들에서 천식 및 알레르기 질환의 유병률과 중증도를 조사하게 된다. 이때 사용하는 도구로써 ISAAC 협력위원회(ISAAC Coordinating Committee)에서 개발된 영문판 설문지를 기본 골격으로 하여 참여하는 나라들마다 자국언어로 된 설문지들을 공통된 지침에 따라서 개발되었다. 이 설문지들은 표준화가 이루어졌고, 그 타당성이 입증되어 현재 국제적으로 ISAAC을 진행하는데 사용되고 있는 중이다(Shaw 등, 1995).

그러나, 그와 같은 각각의 자국언어로 된 설문지를 사용하다 보니 문제점이 발생하게 되었는데 즉, 국제간의 비교에 있어서 문화와 언어의 차이에 따라서 설문지의 해석상에도 차이가 생기게 된 것이다. 이러한 문제점들은 이미 몇몇 국가들에서 체험을 하였고, 따

라서 이런 국가들에서 설문지를 사용하여 얻어진 자료는 그 신뢰성을 재고하게 되었다(Burney 등, 1989; Crane 등, 1989; Osterman 등, 1991). 따라서 이러한 문화와 언어의 차이에서 오는 비교성의 문제점을 최소화하기 위해서 Wellington Asthma Research Group에 의해 천식문항에 관한 비디오 설문법이 개발되었다. 비디오 설문법은 천식의 증상과 증후에 관해서 설문지 문항에 시청각 자료를 함께 제시하는 방법을 말한다(Shaw 등, 1995). 국내에서도 국제간 공통된 지침을 따라서 한국판 ISAAC설문지(이하 ISAAC설문지)와 비디오 설문법(이하 비디오 설문법)을 개발하였는데, 이는 외국의 ISAAC과 일본의 알레르기 질환 역학 조사내용을 바탕으로 하여 6종류의 알레르기 질환에 관한 내용을 포함하고 있다. 설문지는 초등학생용과 중학생용의 두 종류를 만들어서 초등학생의 경우 부모가 설문지를 작성하고, 중학생의 경우 본인 직접 작성하게 하였다(이상일 등, 1995).

역학연구에 쓰이는 도구로서 설문지는 그 시행은 용이하지만 조사내용이 피검자의 주관성에 좌우된다는 문제가 있다(권호장 등, 1995). 따라서 설문지 문항에서 얻은 정보가 얼마나 진실에 가깝느냐는 타당성과, 얼마나 같은 정보를 재현해서 얻을 수 있는느냐는 신뢰성이 검증되어야만 인구집단에 관한 정보를 얻는데 유용하게 쓰일 수 있다.

ISAAC설문지의 타당성에 관한 연구는 메타콜린 유발검사나 전문의의 진단을 황금기준으로 하여 일부 국외연구에서 시행된 바가 있다(Shaw 등, 1995; Jenkins 등, 1996; Ponsonby 등, 1996). 그리고 신뢰성에 관한 연구는 Shaw 등(1995)이 뉴질랜드에서 13~14세 학생 79명을 대상으로 2주 간격으로 비디오 설문법을 검사-재검사하여 각 항목마다 퍼센트 일치율을 산출한 연구가 있다. 이에 반해 한국판 ISAAC설문지와 비디오 설문법은 '한국 어린이, 청소년의 알레르기 질환에 관한 전국적 역학조사'에 이미 사용된 바 있으나, 그 타당도 및 신뢰도는 아직 제대로 검증되지 못한 실정이다. 따라서 본 연구에서는 같은 대상을 상대로 검사-재검사의 방법으로 한국판 ISAAC설문지와

비디오 설문법의 신뢰도를 검증하고, 산출한 신뢰도를 이용하여 타당도를 추정함으로써 한국판 설문지가 국제간 천식 및 알레르기 질환의 유병률 및 증증도 비교에 적합한가를 평가해 보고자 하였다.

II. 연구방법

1. 연구대상

ISAAC연구의 대상 연령층중 자기 자신이 설문지를 직접 작성할 수 있는 만 13~14세 연령층을 연구대상으로 설정하여 서울지역의 중학교 두 곳을 조사하였다. 97년 5월중에 A중학교 1학년 남학생 3개반 113명과 B중학교 1학년 여학생 3개반 111명을 대상으로 ISAAC설문지와 비디오 설문법을 검사-재검사 하였다. 이들중 검사-재검사 어느 한쪽에만 참가한 남학생 3명과 여학생 1명은 분석에서 제외하여 총 220명(남학생 110명, 여학생 110명)이 분석대상이 되었다.

2. 조사내용

ISAAC설문지의 구성은 이름, 성별, 생년월일, 키/몸무게 등을 묻는 인적사항에 관한 항목과 천식, 알레르기성 비염, 아토피성 피부염, 알레르기성 결막염, 식품 알레르기, 약물 알레르기 등의 6가지 알레르기 질환에 관한 내용을 묻는 항목으로 구성되어 있다. 각각의 알레르기 질환에 관한 문항은 그 질환의 과거력, 지난 1년간의 유병상태, 그 질환의 진단력, 치료력 등을 포함하고 있다.

비디오 설문법은 천식증상을 나타내는 비디오 장면 5가지를 보여 주면서 각 장면마다 3~4가지 질문에 답하게끔 구성되어 있다. 비디오의 장면은 휴식시 증증도의 천명증상을 보이는 장면과 운동중 천명증상을 보이는 장면, 밤에 수면중에 천명증상으로 깨어나는 장면, 밤에 수면중에 심한 기침으로 깨어나는 장면, 휴식시 심한 천명과 호흡곤란 증상을 보이는 장면 등으로 구성되어 있다.

3. 조사방법

학사일정 관계로 A중학교의 경우 2주 간격으로, B중학교의 경우 3주 간격으로 검사-재검사를 수행하였다. ISAAC 표준지침을 따라서 ISAAC설문지의 작성 후 비디오 설문법을 시행하는 순서를 따랐다. ISAAC 설문지의 작성은 본 연구의 연구원이 직접 교실에 참관한 상태에서 학생들에게 설문지 작성요령을 숙지시킨 후 15분에 걸쳐서 이루어졌다. 이어서 10분간에 걸쳐 비디오를 상영하면서 천식에 관한 비디오 설문지를 작성케 하였다.

4. 자료의 분석

첫 번째로 신뢰도(일치율)의 평가지표로써 카파값(또는 가중치를 둔 카파값), 카파값의 95%신뢰구간을 산출하였다. 이분성 명목변수와 다분성 명목변수로 된 문항의 경우 카파값을, 순위척도로 된 문항에 대해서는 가중치를 둔 카파값을 산출하였다(Maclure와 Willett, 1987). 일부 문항들은 검사나 재검사시 응답이 '예', 또는 '아니오' 한가지 밖에 없었고, 따라서 자료의 분포가 1×1표, 1×2표, 2×1표의 형태를 가지므로 이 경우 퍼센트 일치율로써 그 일치정도를 해석하였다(Brennan과 Hays, 1992).

카파값의 95% 신뢰구간을 구하여 신뢰구간이 '0'을 포함하고 있을 경우에는 유의하지 않으므로 우연에 의한 일치로 해석하였다. 카파값의 일치강도를 정량적으로 평가하고자 Jekel 등(1996)이 제안한 평가법을 사용하여 0~0.2: negligible(무시할 만한 수준), 0.2~0.4: minimal(미약한 수준), 0.4~0.6: fair(보통 수준), 0.6~0.8: good(좋은 수준), 0.8~1.0: excellent(상당한 수준)로 평가하였다. 그리고 퍼센트 일치율은 70%: necessary(최소한의 수준), 80%: adequate(적절한 수준), 90%: good(좋은 수준)로 평가하였다(Topf, 1986).

두 번째로 검사-재검사가 동일한 측정방법으로, 그

리고 동일한 사람들을 대상으로 행해졌을 경우(즉 검사, 재검사의 측정오류의 분산이 동일한 parallel test의 경우) 타당도 계수는 일치율인 카파값의 양의 제곱근과 동일하다는 아래의 추정식을 이용하여 타당도 계수를 제시하였다(Allen과 Yen, 1979; Armstrong 등, 1992). 카파값이 음의 값을 가지거나 유의하게 나오지 않은 경우 타당도 계수는 산출할 수가 없었다. 타당도의 정량적 기준은 문헌고찰을 하였으나 발견할 수 없었고, 기존의 연구 결과들에 의해서 0.6 이상이면 '높은 수준'인 것으로 추정하였다(하미나, 1996; 주영수 등, 1997; 박병주 등, 1998).

$$\rho_{TX1} = \rho_{TX2} = \sqrt{\rho_{X1X2}}$$

(ρ_{TX1} : test X1의 타당도 계수

ρ_{TX2} : test X2의 타당도 계수

ρ_{X1X2} : test X1, X2의 상관계수)

III. 연구결과

1. ISAAC설문지의 신뢰도 및 타당도

1) 천식문항

일치율(카파값)은 천식에 관한 10개 문항을 평가하였을 때 -0.01~1.00로 넓은 범위에 걸쳐 있었다. '지난 1년간 천명으로 인한 수면장애의 횟수'를 묻는 문항은 완전한 일치를 보였고, '천식의 진단력 유무'와 '지난 1년간 천명증상 유무'에 관한 문항은 좋은 수준의 일치를 보였다. 그 외 문항들은 모두 미약한 수준 이하의 일치가 관찰되었다. 천식에 관한 5개 문항에서 타당도 계수는 0.49~1.00의 범위였다.

'지난 1년간 천식으로 인한 결석일수'를 묻는 문항은 모든 빈도가 검사시 '결석한 적이 없다', 그리고 재검사시 '결석한 적이 없다'로 집중하여 카파값은 산출할 수 없었으나 1×1표의 형태이고 퍼센트 일치율이 100%이므로 완전한 일치로 해석하였다.

지난 1년간 천식으로 인한 여러 가지 장애들을 묻는 문항들은 모두 천명증상의 과거력이 있는 학생들

표 1. ISAAC설문지 천식문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
천명증상의 과거력 유무	0.36 (0.16, 0.55)	86.0	0.60
지난 1년간 천명증상 유무	0.73 (0.24, 1.21)	88.9	0.85
지난 1년간 천명증상의 횡수	0.57* (-0.12, 1.26)	83.3	—
지난 1년간 천명증상으로 인한 수면장애의 횡수	1.00* (1.00, 1.00)	100.0	1.00
지난 1년간 심한 천식발작의 경험 유무	0.33 (-0.23, 0.90)	66.7	—
지난 1년간 천식으로 인한 결석일수	— —	100.0	—
천식의 진단력 유무	0.66 (0.30, 1.02)	98.5	0.81
지난 1년간 천식 치료력 유무	-0.01 (-0.02, 0.00)	98.1	—
지난 1년간 운동중 천명증상 유무	0.24 (0.04, 0.44)	87.0	0.49
지난 1년간 밤중에 기침증상 유무	0.15 (-0.06, 0.35)	88.9	—

*가중치를 둔 카파값

표 2. ISAAC설문지 알레르기성 비염문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
알레르기성 비염의 과거력 유무	0.56 (0.44, 0.67)	78.8	0.75
지난 1년간 알레르기성 비염증상의 유무	-0.13 (-0.21, -0.05)	76.6	—
지난 1년간 알레르기성 비염과 알레르기성 결막염의 동반여부	0.54 (0.27, 0.82)	80.0	0.74
지난 1년간 알레르기성 비염으로 인한 일상생활의 불편함의 정도	0.70* (0.43, 0.97)	86.7	0.84
알레르기성 비염의 진단력 유무	0.64 (0.43, 0.86)	95.4	0.80
지난 1년간 알레르기성 비염의 치료력 유무	0.65 (0.39, 0.91)	97.0	0.81

*가중치를 둔 카파값

표 3. ISAAC설문지 아토피성 피부염 문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
아토피성 피부염의 과거력 유무	0.58 (0.31, 0.84)	95.9	0.76
아토피성 피부염의 진단력 유무	0.87 (0.69, 1.05)	99.0	0.93
지난 1년간 아토피성 피부염의 치료력 유무	0.83 (0.59, 1.06)	99.0	0.91

표 4. ISAAC설문지 알레르기성 결막염 문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
알레르기성 결막염의 과거력 유무	0.57 (0.41, 0.73)	88.5	0.75
지난 1년간 알레르기성 결막염 증상 유무	0.00 —	89.5	—
알레르기성 결막염에 이환된 눈 부위	0.65 (0.35, 0.96)	77.8	0.80
알레르기성 결막염의 진단력 유무	0.49 (0.06, 0.92)	98.0	0.70
지난 1년간 알레르기성 결막염의 치료력 유무	0.80 (0.41, 1.19)	99.5	0.89

만이 응답하게끔 되어있다. 따라서 응답자가 적어서 표본크기가 10이하가 되었고, 따라서 이 문항들에 대한 카파값과 타당도 계수는 다소 통계적 검정력이 낮았다(표 1).

2) 알레르기성 비염 문항

알레르기성 비염에 관한 6개의 문항을 평가했을 때 일치율은 0.13~0.70의 범위였다. ‘지난 1년간 알레르기성 비염증상 유무’를 묻는 문항은 우연보다 더 드문 일치의 수준이었다. 그 외 문항들은 모두 보통 수준에서 좋은 수준 사이의 일치를 나타냈다. 알레르기성 비염에 관한 5개의 문항에서 타당도 계수는 모두 0.7이상으로 높은 수준이었다(표 2).

3) 아토피성 피부염 문항

아토피성 피부염에 관한 3개의 문항은 일치율이 0.58~0.83의 범위였다. ‘아토피성 피부염의 진단력 유무’와 ‘지난 1년간 아토피성 피부염의 치료력 유무’를 묻는 문항은 상당한 수준의 일치를 나타냈고, ‘아토피성 피부염의 과거력 유무’를 묻는 문항은 보통수준의 일치를 보였다. 타당도 계수도 모두 0.7 이상으로 높은 수준이었다(표 3).

‘지난 1년간 아토피성 피부염으로 인한 여러 가지 장애의 유무’를 묻는 문항들은 아토피성 피부염의 과거력이 있는 학생들만 답하게 되어있고, 따라서 응답자가 5명 이하가 되어 표본크기가 너무 작은 관계로 카파값을 산출할 수 없었다.

4) 알레르기성 결막염 문항

알레르기성 결막염에 관한 5개의 문항을 분석했을 때 일치율은 0.00~0.80의 범위였다. ‘알레르기성 결막염의 과거력 유무’와 ‘알레르기성 결막염의 진단력 유무’를 묻는 문항은 보통 수준의 일치를 나타냈고, ‘알레르기성 결막염에 이환된 눈 부위’와 ‘지난 1년간 알레르기성 결막염의 치료력 유무’를 묻는 문항은 좋은 수준의 일치를 나타냈다. 알레르기성 결막염에 관한 4개의 문항에서 타당도 계수는 모두 0.7 이상으로

높은 수준이었다.

‘지난 1년간 알레르기성 결막염 증상 유무’를 묻는 문항은 검사시 응답이 ‘예’ 하나 밖에 없으므로 자료가 2×1표의 형태이고 퍼센트 일치율은 89.5%로 적절한 수준의 일치가 관찰되었다(표 4).

5) 식품 알레르기 문항

식품 알레르기에 관한 4개의 문항들은 일치율이 0.66~0.82의 범위였다. ‘지난 1년간 식품 알레르기 증상 유무’를 묻는 문항은 상당한 수준의 일치를 나타냈고, 그 외 문항들은 좋은 수준의 일치를 보였다. 4개 문항에 대한 타당도 계수는 모두 0.8 이상으로 높은 수준이었다(표 5).

6) 약물 알레르기 문항

약물 알레르기에 관한 3개의 문항중 ‘약물 알레르기의 과거력 유무’를 묻는 문항은 일치율이 보통 수준이었고, ‘약물 알레르기의 진단력 유무’를 묻는 문항은 우연에 의한 일치와 같았다. ‘약물 알레르기의 과거력 유무’를 묻는 문항에 대한 타당도 계수는 0.70으로 높은 수준이었다.

‘지난 1년간 약물 알레르기의 치료력 유무’를 묻는 문항은 검사시 응답이 ‘아니오’ 하나 밖에 없기 때문에 자료는 1×2표의 형태이고 퍼센트 일치율은 99.5%로 좋은 일치를 나타냈다(표 6).

2. 비디오 설문법의 신뢰도 및 타당도

비디오 설문법의 신뢰도 및 타당도 비디오 설문법의 14개 문항들은 ‘천명증상’, ‘밤중에 기침증상’, 그리고 ‘심한 천식발작 증상’의 크게 3가지 증상에 관한 문항들로 분류할 수가 있으며, 그 일치율은 0.00~0.66 범위였다.

천명에 관한 문항들에서 ‘한달에 1회 이상 운동중 천명증상 유무’를 묻는 문항만 우연에 의한 일치를 보였고, 전반적으로 보통 수준의 일치를 나타냈다. 상대적으로 비교해 볼 때 ‘밤중에 천명증상’을 묻는 문항

표 5. ISAAC설문지 식품 알레르기 문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
식품 알레르기의 과거력 유무	0.76 (0.59, 0.92)	96.0	0.87
지난 1년간 식품 알레르기 증상 유무	0.82 (0.50, 1.15)	91.7	0.91
식품 알레르기의 진단력 유무	0.70 (0.43, 0.98)	98.0	0.84
지난 1년간 식품 알레르기의 치료력 유무	0.66 (0.22, 1.10)	99.0	0.81

표 6. ISAAC설문지 약물 알레르기 문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
약물 알레르기의 과거력 유무	0.49 (0.06, 0.92)	98.0	0.70
약물 알레르기의 진단력 유무	0.40 (-0.15, 0.94)	98.5	-
지난 1년간 약물 알레르기의 치료력 유무	0.00 -	99.5	-

표 7. 비디오 설문법 문항에 대한 신뢰도와 타당도

설문지 문항	카파값(95% 신뢰구간)	퍼센트 일치율(%)	타당도 계수
천명증상의 과거력 유무	0.50 (0.27, 0.72)	92.7	0.70
지난 1년간 천명증상 유무	0.34 (0.03, 0.65)	91.9	0.58
한달에 1회 이상 천명증상 유무	0.48 (0.05, 0.92)	96.7	0.70
운동중 천명증상의 과거력 유무	0.51 (0.32, 0.70)	90.6	0.71
지난 1년간 운동중 천명증상 유무	0.57 (0.31, 0.83)	93.9	0.76
한달에 1회 이상 운동중 천명증상 유무	-0.03 (-0.05, 0.00)	93.0	-
밤중에 천명으로 인한 수면장애의 과거력 유무	0.59 (0.23, 0.96)	98.0	0.77
지난 1년간 밤중에 천명으로 인한 수면장애 유무	0.66 (0.04, 1.28)	99.2	0.81
밤중에 기침증상의 과거력 유무	0.28 (0.11, 0.46)	84.0	0.53
지난 1년간 밤중에 기침증상 유무	0.34 (0.09, 0.60)	88.5	0.59
한달에 1회 이상 밤중에 기침증상 유무	0.00 -	94.5	-
심한 천식발작의 과거력 유무	0.35 (-0.02, 0.72)	96.5	-
지난 1년간 심한 천식발작의 유무	0.32 (-0.16, 0.80)	96.6	-
한달에 1회 이상 심한 천식발작 유무	0.66 (0.04, 1.28)	99.2	0.81

표 8. ISAAC설문지와 비디오 설문법의 공통된 문항에 대한 신뢰도와 타당도 비교

설문지 문항	ISAAC설문지				비디오 설문법			
	카파값	95% 신뢰구간	퍼센트 일치율(%)	타당도 계수	카파값	95% 신뢰구간	퍼센트 일치율(%)	타당도 계수
천명증상의 과거력 유무	0.36	(0.16, 0.55)	86.0	0.60	0.50	(0.27, 0.72)	92.7	0.70
지난 1년간 천명증상 유무	0.73	(0.24, 1.21)	88.9	0.85	0.34	(0.03, 0.65)	91.9	0.58
지난 1년간 운동중 천명증상 유무	0.24	(0.04, 0.44)	87.0	0.49	0.57	(0.31, 0.83)	93.9	0.76
지난 1년간 심한 천식발작 유무	0.33	(-0.23, 0.90)	66.7	0.58	0.32	(-0.16, 0.80)	96.6	0.57
지난 1년간 밤중에 천명으로 인한 수면장애 유무	1.00	(1.00, 1.00)	100.0	1.00	0.66	(0.04, 1.28)	99.2	0.81
지난 1년간 밤중에 기침증상 유무	0.15	(-0.06, 0.35)	88.9	0.39	0.34	(0.09, 0.60)	88.5	0.59

들은 그 일치율이 더 높은 것을 알 수 있었다. 천명 문항들의 타당도 계수는 '지난 1년간 천명증상 유무'를 묻는 문항만 0.58이었고, 그외는 모두 0.7 이상으로 높은 편이었다.

기침증상에 관한 문항들에서 '한달에 1회 이상 밤중에 기침증상 유무'를 묻는 문항은 검사시 응답이 '아니오' 하나 밖에 없어 자료는 1×2표의 형태를 취하고 퍼센트 일치율은 94.6%로 좋은 일치를 나타냈다. 그 외 '밤중에 기침증상'에 관한 문항들은 미약한 일치가 관찰되었다. '밤중에 기침증상'을 묻는 문항들의 타당도 계수는 0.53~0.59의 범위였다.

심한 천식발작 증상에 관한 문항들에서 '한달에 1회 이상 심한 천식발작 유무'를 묻는 문항의 경우만 좋은 수준의 일치가 관찰되었고, 나머지는 우연에 의한 일치와 같았다. '한달에 1회 이상 심한 천식발작 유무'를 묻는 문항의 타당도 계수는 높은 수준이었다(표 7).

3. ISAAC설문지와 비디오 설문법의 비교

천식에 관한 ISAAC설문지와 비디오 설문법의 공통된 6가지 문항에 대해서 그 일치율과 타당도 계수를 비교해 보았다. 일치율의 비교시 '천명증상의 과거력 유무', '지난 1년간 운동중 천명증상 유무', 그리고 '지난 1년간 밤중에 기침증상 유무'를 묻는 3가지 문항들은 비디오 설문법이 더 높았고 나머지 3가지 문항들은 ISAAC설문지가 더 높았다. 타당도 계수의 비교도 일치율 비교와 마찬가지로였다(표 8).

IV. 고 찰

인구 집단을 대상으로 한 천식 및 알레르기 질환의 유병률과 그 중증도의 조사는 여러 가지 제한점으로 인하여 수행하기가 힘든 것이 사실이다. 왜냐하면 이는 숙련된 전문의의 검진, 메타콜린 반응검사나 고장성 식염수 피내검사 등의 임상검사, 기존의 신뢰 가능한 자료원 등 많은 인력과 재원이 필요하기 때문이다.

따라서 한정된 인력과 재원을 활용하여 인구 집단에서의 평가를 하기 위해서는 그 적용과 일반화의 용이함, 그리고 비용측면의 장점 등에서 설문지를 사용하는 것이 가장 적합하다 할 수 있다.

그러나 이러한 설문지를 역학조사에 사용하려면 그 타당도와 신뢰도가 먼저 검증되어야만 한다. 즉 한국판 ISAAC설문지가 국내의 소아인구집단을 대상으로 한 역학조사와 국제간 자료비교에 적합하기를 먼저 평가해야 한다. 이 경우 타당도를 평가한 후에 신뢰도를 평가하는 것이 원칙적인 순서이며, 이 때 타당도는 메타콜린 유발검사나 전문의의 진단 등의 황금기준과 설문지 결과를 비교하여 평가하는 것이 일반적이다. 그러나 이전에 국내에서 행해진 역학조사들에서 이러한 타당도에 관한 정보를 얻는 것은 용이하지 않았다. 따라서 본 연구에서는 그 순서를 달리하여 설문지의 검사-재검사를 통하여 신뢰도를 먼저 산출하고, 이를 바탕으로 타당도를 추정하는 방법을 사용하였다. 이러한 타당도 산출과정은 이미 논리적으로 충분히 검증된 바가 있다(Allen과 Yen, 1979).

검사방법의 신뢰도를 평가하는 데는 여러 가지 분석방법을 적용할 수가 있다. 개개인의 수준에서 일치율을 볼 수도 있고, 군집수준에서의 일치율을 볼 수도 있다. 그러나 군집수준의 일치율을 평가하는 것은 결국 개개인의 일치율에 영향을 받게 되므로 실제로는 개개인의 일치율을 주로 이용하게 된다(주영수 등, 1997). 개개인의 일치율을 보는 방법에는 측정하고자 하는 변수가 범주형(명목, 혹은 순위)일 경우 Cohen's kappa나 weighted kappa를 이용하는 것이 흔히 추천된다. 결국 이러한 카파값들은 두 검사법간의 일종의 상관성을 평가하는 지표이면서 우연에 의한 일치를 보정해 준다는 데 그 장점이 있다.

그러나 천식과 같이 일반인구 집단에서 유병률이 낮은 질환의 경우 설문지 응답에서 응답자의 빈도가 검사에서 '아니오', 그리고 재검사에서 또한 '아니오'라고 답하는 칸으로 집중이 되는 경향이 있다. 이럴 경우 카파값은 병원 등의 천식 유병률이 높은 환경에서 조사한 값에 비해서 다소 과소추정이 될 수 있다

(Spitznagel과 Helzer, 1985). 따라서 본 연구에서도 카파값이 이러한 질환의 유병률에 민감하다는 사실을 염두에 두고 이러한 카파값의 단점을 보완하기 위하여 퍼센트 일치율을 같이 산출하여 일치정도를 해석하였다.

자료의 분포가 편중되어 1×2표, 2×1표의 형태인 경우 관찰된 일치분율(Po)과 기대일치분율(Pe)은 같아져 카파값의 점 추정치는 '0'이 되고, 95%신뢰구간은 계산할 수 없게 된다. 또한 자료의 분포가 1×1표의 형태인 경우 빈도는 한 칸으로만 집중하게 되고, 이때 카파값은 구할 수가 없지만(undefined) 퍼센트 일치율은 100%이므로 완전한 일치로 해석하였다(Brennan과 Hay, 1992).

천식 및 알레르기성 질환별로 신뢰도와 타당도를 비교 평가해 보았을 때, 천식과 약물 알레르기에 관한 문항들이 상대적으로 신뢰도와 타당도가 더 낮았고, 아토피성 피부염, 식품 알레르기에 관한 문항들은 더 높은 것이 관찰되었다. 그리고 비디오 설문법과 ISAAC설문지는 예상과는 달리 그 신뢰도나 타당도가 거의 같은 것으로 나타났다.

ISAAC설문지에서 신뢰도나 타당도가 낮게 나타났던 문항들을 살펴보면 이는 주로 조사대상자들이 질문의 내용이나 의도를 정확하게 이해하지 못한 데서 기인한 것이라는 것을 알 수 있다. 이는 크게 세가지로 분류해 볼 수가 있다. 그 중 첫 번째는 설문지에서 쓰인 일부 용어들이 일반인이 아닌 전문가 집단에서 주로 사용하고 이해하는 것들이라는 점이다. 이의 예로는 '태어나서 지금까지 천식으로 진단받은 적이 있습니까?'라는 문항과 '감기 또는 폐렴이 아니면서 가래가 나오지 않는 마른 기침을 한밤중이나 새벽에 한 적이 있습니까?' 등의 문항을 들 수 있다. 특히 '천식'이나 '폐렴'이라는 용어를 이해하지 못하여 질문하는 학생들이 상당히 많았다는 것으로 이러한 사실을 증명할 수가 있었고, '기침'과 같은 비특이적으로 흔한 증상은 어느 수준의 무엇을 가리키는지를 대부분 응답자들이 잘 이해하지 못했다. 두 번째로 질환의 증상을 나타내는 전문적 의학용어를 풀어서 기술할 때 그

전달하고자 하는 바를 완전하게 표현하지 못했다는 것이다. 이의 예로는 천명을 '숨쉴 때 가슴에서 나는 쉼쉼거리는 소리나 휘파람 소리'로 표현한 것이다. 이러한 표현을 학생들은 운동시에 숨찬 것과 혼동하는 경향이 있었다. 세 번째로 최근이라는 시간개념을 나타내기 위해 '지난 12개월간'이라는 표현을 사용했는데, 이를 학생들은 정확하게 지난 12개월내에 증상이 있었는 지를 기억해 내고자 하여 의도하는 바와 달리 혼란을 초래하였다.

천식에 관한 ISAAC설문지의 문항들 중에서 특히 '운동중 천명음'이나 '밤중에 기침증상', '심한 천식 증상'에 관한 문항들은 다른 문항들에 비해서 신뢰도와 타당도가 더 낮게 측정되었다. 이는 뉴질랜드 천식 연구 그룹에서 나온 결과와 일치했다(Shaw 등, 1995). 비디오 설문법의 경우 시청각 방법을 동원하여 설문지 상에서 표현하기 힘든 증상들에 대한 이해를 돕는 것을 그 목적으로 하여 개발되었으나, 본 연구에서는 예상외로 ISAAC설문지보다 나은 점을 발견할 수 없었다. 이는 기존에 행해졌던 외국의 연구들과는 일치하지 않는 결과이며 아마도 비디오 장면의 천식환자들이 유럽계나 서남 아시아계 인종으로 구성되어 문화적 이질성으로 인해서 증상표현이 다소 이해하기 어렵지 않았나 추측할 수 있었다. 실제로 Wellington Asthma Research Group에서는 인종간의 대표성을 높이기 위한 새로운 버전의 비디오 설문법을 개발중이었다(Shaw 등, 1995).

이상의 연구결과 ISAAC설문지의 경우 알레르기성 비염, 아토피성 피부염, 알레르기성 결막염, 식품 알레르기에 관한 문항들은 별다른 수정없이 사용이 가능한 것으로 보인다. 그러나 천식에 관한 문항들은 '천명'에 대한 표현이나 '기침', '폐렴'에 대한 설명이 조사 대상자들의 관점에서 기술되거나 추가되어야 하고, 운동중 천식발작에 대한 표현이 더욱 정확하게 기술되어야 한다고 판단된다. 약물 알레르기에 관한 문항들은 약물 알레르기 증상에 관한 설명이 포함되어 있지 않았으므로 조사 대상자들이 잘 이해할 수 있는 설명이 추가되어야 한다. 비디오 설문법의 경우 비디오

장면들을 한국판으로 다시 구성한다면 인종적, 그리고 문화적 이질감을 없애고 시청각 효과를 적극 활용할 수 있으리라 생각한다. 이상의 문항들에 대한 수정을 가한 후 보완된 설문지로 다시 타당도와 신뢰도를 검증하여 차후의 역학조사에 활용하는 것이 앞으로의 과제가 될 것이다.

본 연구에서는 설문조사의 대상이 중학생으로 국한되어 초등학생용 설문지의 신뢰도와 타당도는 평가해보지 못했다. 따라서 다음 기회에 초등학생 학부모를 대상으로 한 신뢰도와 타당도의 검증이 이루어진다면 본래 ISAAC설문지의 대상인 두 연령층에 대한 완벽한 평가가 이루어지리라 생각한다.

분석단계에서 응답자 수가 적어 통계적 검정력이 낮아진 문항들은 연구대상수를 늘려서 다시 평가해본다면 신뢰할 만한 결과를 얻을 수 있을 것이다. 그리고, ISAAC설문지와 비디오 설문법의 비교시, 그 차이에 대한 유의성 검정을 하지 못했는데 이는 철저한 문헌고찰을 했음에도 불구하고 기존에 개발되어진 통계적 방법을 발견할 수 없었기 때문이다.

한국판 ISAAC설문지의 신뢰도 평가는 원래 ‘한국 어린이, 청소년의 알레르기 질환에 관한 전국적 역학조사 연구’가 시작되기 전에 이루어져야 했다. 그러나 지금이라도 그 설문지의 신뢰도와 타당도가 평가된 것은 다행한 일로 생각한다. 이러한 신뢰도와 타당도의 바탕위에서 기존에 수행되었던, 그리고 앞으로 수행될 역학조사 결과들이 가치를 가질 수 있기 때문이다.

V. 요약 및 결론

세계 각국이 연계하여 진행하고 있는 ‘국제 소아천식 및 알레르기질환 연구’에 동참하여 국내에서 전국적 역학조사를 목적으로 개발한 한국판 ISAAC설문지와 비디오 설문법의 신뢰도와 타당도를 평가하였다. 만 13~14세 연령층을 대상으로 설문법의 검사-재검사를 통하여 일치율 지표인 카파값 또는 가중치를 둔 카파값을 산출하여 신뢰도를 평가하였고, 이를 이용하

여 타당도 계수를 추정하였다.

ISAAC설문지의 경우 알레르기성 비염, 아토피성 피부염, 알레르기성 결막염, 식품 알레르기에 관한 문항들은 신뢰도와 타당도가 전반적으로 높아서 별다른 수정없이 사용이 가능한 것으로 보인다. 그러나 천식이나 약물 알레르기에 관한 문항들은 그 용어와 증상에 대한 설명이 조사대상자들의 입장에서 더욱 이해하기 쉽도록 기술되어야 할 것이다.

비디오 설문법의 경우 그 개발목적 측면에서 ISAAC설문지에 비해 신뢰도와 타당도가 높을 것으로 기대하였으나 결과는 ISAAC설문지와 비슷한 수준이었다. 이는 비디오 장면들을 한국판으로 다시 구성한다면 인종이나 문화적 이질감을 없애고 시청각 효과를 적극 활용하여 기존의 ISAAC설문지의 부족한 부분을 보완할 수 있을 것으로 기대한다.

한국판 ISAAC설문지와 비디오 설문법은 일부 문항들을 수정, 보완한다면 국내 소아천식 및 알레르기질환의 역학조사시 도구로써 활용가능할 것이다. 그리고 알레르기 질환의 유병률이나 중증도의 국제간 비교에 있어서 비교성을 획득할 수 있는 방법이 될 것으로 기대한다.

참고문헌

- 권호장, 조수현, 임현술. 신경행동학적 증상 설문지의 타당도 및 신뢰도에 관한 연구. 대한산업의학회지 1995; 7(1):21-27
- 박병주, 김대성, 구혜원, 배종면. 노인 생활습관 설문서의 신뢰도 및 타당도 평가 연구. 예방 의학회지 1998; 31(1):49-58
- 이상일, 손병관, 이준성, 이하백, 고영율, 김규언. 한국 어린이, 청소년의 알레르기 질환에 관한 전국적 역학조사. 1995년도 보건의료기술연구개발사업 보고서. 1995
- 주영수, 김대성, 강종원, 성주현, 강대회, 조수현, 백도명. 대기오염과 관련된 건강영향을 평가하기 위한 설문 개발. 예방의학회지 1997; 30(4):852-69
- 하미나. 역학적 연구에서 인간공학적 폭로변수의 평가를 위한 설문 개발. 서울대학교 보건대 학원 석사 학위 논문. 1996

- Allen MJ, Yen WM. Introduction to Measurement Theory, pp.1-117. Brooks/Cole, Monterey. 1979
- Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology New York, Oxford University Press, 1992, p81-84
- Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, Mitchell EA, Pearce N, Sibbald B, Stewart AW, Strachan D, Weiland SK, Williams HC. International study of asthma and allergies in childhood (ISAAC): rationale and methods. *Eur Respir J* 1995; 8:483-491
- Brennan PF, Hays BJ. The Kappa Statistic for Establishing Interrater Reliability in the Secondary Analysis of Qualitative Clinical Data. *Research in Nursing & Health* 1992; 15:153-158
- Burney PGJ, Laitinen LA, Perdrizet S. Validity and repeatability of the IUATLD(1984) Bronchial Symptoms Questionnaire: an international comparison. *Eur Resp J* 1989; 2:940-945
- Crane J, O'Donnell TV, Prior IAM, Waite DA. Symptoms of asthma, methacholine airway responsiveness and atopy in migrant Tokelauan Children. *NZ Med J* 1989; 102:26-38
- Jekel JF, Elmore JG, Katz DL. Epidemiology, Biostatistics, and Preventive Medicine. pp96-97, W.B. Saunders Company, Philadelphia, 1996
- Jenkins MA, Clarke JR, Carlin JB, Roberston CF, Hopper JL, Dalton MF, Holst DP, Choi K, Giles GG. Validation of questionnaire and bronchial hyperresponsiveness against respiratory physician assessment in the diagnosis of asthma. *Int J Epidemiol* 1996; 25(3):609-616
- Maclure M, Willett WC. Misinterpretation and Misuse of the Kappa Statistic. *American Journal of Epidemiology* 1987; 126(2):161-169
- Osterman JW, Ledoux E, Sloan M, Ernst P. Comparison of French and English versions of the American Thoracic Society Respiratory Questionnaire in a bilingual working population. *Int J Epidemiol* 1991; 20:138-143
- Ponsonby AL, Couper D, Dwyer T, Carmichael A, Wood-Baker R. Exercise-induced bronchial hyperresponsiveness and parental ISAAC questionnaire responses. *Eur Respir J* 1996; 9: 1356-1362
- Shaw R, Woodman K, Ayson M, Dibdin S, Winkelmann R, Crane J, Beasley R, Pearce N. Measuring the Prevalence of Bronchial Hyper-Responsiveness in Children. *International Journal of Epidemiology* 1995; 24(3):597-602
- Spitznagel EL, Helzer JE. A proposed solution to the base rate problem in the kappa statistic. *Arch Gen Psychiatry* 1985; 42:725-729
- Topf M. Three Estimates of Interrater Reliability for Nominal Data. *Nursing Research* 1986; 35(4):253-255