

■ 원 저

한국형 건강관련 삶의 질 측정도구의 개발

— 신뢰도 및 타당도 검증 —

김수영, 심재용*, 원장원**, 선우성***, 박훈기****, 이정권****

한림대학교 가정의학교실, 연세의대 가정의학교실*, 경희의료원
가정의학과**, 울산의대 가정의학교실***, 한양대학교 가정의학교실****

— 요약 —

연구 배경 : 한국형 삶의 질 측정도구(Korean health related Quality of Life Scale, KQOLS)의 영역 설정, 증상채집, 문항화, 범주화, 문항 줄이기, 사전조사에 대한 것은 다른 논문으로 출판된 바 있다. KQOLS가 측정도구로서 가지는 가치를 검증하기 위해서는 신뢰도, 타당도 그리고 반응성에 대한 검증이 필요하다.

방법 : 전문가 회의를 거쳐 KQOLS의 최종문항을 46개로 확정하였고, 응답항목을 결정하였다. 1998년 12월부터 1999년 2월까지 6개 병원 가정의학과에 내원한 환자 중 2주 이상 지속되는 증상을 가진 환자군과 3개 종합검진센터를 방문한 사람 중 특별한 증상이 없고 검진결과가 정상인 정상군을 대상으로 내적일치도법에 의한 신뢰도, 문항분석을 실시하였고 타당도를 분석하였다. 타당도는 주요인 방법으로 고유근 1 이상으로 요인 수를 정한 후 varimax 회전 후의 요인 적재값으로 구한 요인분석과 환자군과 정상군의 각 영역별 점수를 비교하는 임상타당도로 검증하였다. 측정도구의 반응성은 환자군에 대한 적절한 의학적 처치 후 호전군과 비호전군에 대해서 영역별 점수의 변화로 판정하였다. 또한 일개대학병원 사무직원으로 특별한 증상을 호소하지 않는 사람을 대상으로 2주간 검사-재검사법을 실시하였다.

결과 : 측정도구의 내적 일치도는 모든 영역에서 0.7을 넘었고 2주 검사-재검사법에 의한 신뢰도 계수는 건강상태 인식영역을 제외하고는 통계적으로 유의한 상관성을 보였다. 요인분석결과는 긍정적 건강 측정 문항과 일부영역을 제외하고는 원래 의도한 요인으로 묶였다. 삶의 질 각 영역별 점수는 영적 건강과 사회적 기능을 제외한 모든 영역에서 정상군이 환자군보다 유의하게 점수가 높았다. 환자군 중 호전군은 건강상태인식과 영적 건강을 제외한 모든 영역에서 치료전후에 유의한 차이를 보였고 불변군의 경우 모든 영역에서 유의한 차이를 보이지 않았다.

결론 : KQOLS는 신뢰도와 타당도가 높고 변화에 민감하고 특이한 측정도구이다. 다만 건강인식문항 변경, 긍정적 건강 측정 문항 응답항목 수정과 같은 개정이 필요하다. (가정의학회지 2000 ; 21 : 382~394)

중심단어 : 삶의 질, 측정도구, 신뢰도, 타당도, 반응도

연구비 수혜원천 : 이 연구는 대한민국 보건복지부 보건의료기술 연구 개발사업(#HMP-97-M-5-0056) 연구비의 지원으로 이루어졌음.

서 론

삶의 질은 정의하기 힘든 개념으로 '안녕', '기능상태', '건강상태', '생활만족' 등의 여러 가지 의미로 쓰인다. 보건 의료 분야에서의 삶의 질은 각 개인이 주관적으로 평가한 건강 상태를 의미하며 이를 특별히 건강관련 삶의 질 (Health Related Quality of Life, HRQoL) 이라고 부른다¹⁾.

건강관련 삶의 질은 환자 스스로의 경험에 관심을 둔다는 점에서 주관적 요소가 강하며 그러한 이유로 각 문화권이나 국가에 따라서 상당히 다른 양상을 보일 수 있다^{2,3)}. 또한 삶의 질 측정의 주요 대상인 '건강', '상병감', '고통', '질병' 같은 낱말들의 의미는 사람들의 기대와 가치에 따라 차이가 나기 때문에 문화에 따라서 의미가 달라지는 경우가 많다. 그러한 점에서 그 나라 문화나 정서, 언어습관 등을 반영한 삶의 질 측정도구로 그 나라 삶의 질을 측정하는 것의 의미는 클 수밖에 없다⁴⁾.

한국형 삶의 질 측정척도 (Korean quality of life scale, KQOLS) 는 한국인의 언어 표현 및 문화를 반영할 수 있는 것으로 영역 설정, 증상채집, 문항화, 범주화, 문항 줄이기, 사전조사 등의 내용은 다른 논문으로 출판된 바 있다⁵⁾.

삶의 질 측정도구는 체중계나 혈압계 같은 다른 물리적 측정도구와 마찬가지로 신뢰도, 타당도 같은 심리측정학적 속성 (psychometric properties) 이 검증되어야 하고 변화에 민감하게 반응하는 지를 검증하는 반응성도 입증되어야 한다. 그리고 여러 연구를 거쳐 측정도구가 충분히 신뢰할만하고 타당성이 있는 것으로 나타나면 물리적 측정도구의 표준 (standard) 에 해당하는 규준 (norm) 을 만들어 점수를 나이나 성별에 맞게 해석할 수 있도록 해야 한다⁶⁾.

이 연구는 이러한 과정 중 신뢰도와 타당도 검증, 문항분석 그리고 반응성에 대한 기초적 검증에 대한 것이다. 측정도구의 신뢰도는 동일한 대상에 대하여 같거나 비슷한 측정도구를 사용하여 반복 측정할 경우 동일하거나 비슷한 결과를 얻을 수 있는 정도로 정의할 수 있고 측정의 비체계적 오차와 관련된 개념이다⁷⁾. 신뢰도는 주로 측정도구의 시간에 따른 안정성을 측정하는 검사-재검사법과 설문지 문항이 하나의 구성개념을 측정하는지를 검증하는 내적일치도법

을 주로 이용한다⁸⁾.

타당성은 측정하고자한 특성을 얼마나 진실에 가깝도록 측정했는가 하는 정도를 말하며 측정의 체계적 오차와 관련된 개념이다. 타당도는 황금기준 유무에 따라서 준거타당도나 구성개념 타당도로 검증해야 하는데 삶의 질처럼 황금기준이 없는 경우에는 구성개념 타당도로 타당도를 검증해야 한다⁹⁾. 구성개념타당도는 비슷한 개념을 측정한 측정도구와는 상관성이 높고 (수렴타당도, convergent validity), 다른 개념을 측정한 측정도구와는 상관성이 적은지 (판별타당도, discriminative validity) 를 평가해서 결정하며 수렴타당도와 판별타당도는 다속성다측정법 (multi-trait multi-method matrix) 으로 동시에 검증하는 것이 일반적 과정이다. 구성개념타당도는 요인분석이나 임상타당도로도 검증할 수 있다⁹⁾.

측정도구 분석에서 문항분석은 문항이 어느 정도 해당 영역을 대표하는지를 측정하는 것으로 문항-총점 상관관계와 문항-문항 상관계수로 많이 나타난다. 문항분석은 문항 줄이거나 오류가 있는 문항을 선별하는 데 많이 이용할 뿐만 아니라 각 문항의 평균과 표준편차를 나타내는 문항치 (item value) 는 가중치 여부를 판정하는 데 이용되기도 한다¹⁰⁾.

반응성은 측정하려고 하는 영역에 변화가 생겼을 때 측정도구로 그 변화를 찾아내는 정도를 말한다. 이러한 반응성은 임상시험에 표준설문지가 많이 이용되면서 나타난 개념으로 측정방법이 소개된 것이 최근이어서 외국에서 개발한 설문지의 경우에도 이러한 반응성이 검증된 경우는 많지 않다¹⁰⁾. 이러한 반응성 정도를 나타내는 수치에 대해서는 여러 가지 방법이 제안되고 있는데, 분자는 점수의 변화로 한다는 데 대해서는 대체로 의견이 일치하지만 분모는 측정도구의 표준편차, 표준오차 혹은 안정된 사람들의 표준편차 등 여러 가지로 표현할 수 있다. 이러한 이유로 민감도 (sensitivity) 나 특이도 (specificity) 라는 간편한 방법으로 반응도를 표시하기도 한다¹¹⁾.

이 논문의 목적은 이미 발표된 한국형 삶의 질 설문지의 신뢰도를 검사-재검사법과 내적일치도법으로 검증하고 타당도를 임상타당도와 요인분석으로 검증하고 반응도를 검증하는 것이다.

방 법

1. 문항 개정

한국형 건강관련 삶의 질 측정도구(KQOLS)⁹⁾를 전문가 회의에서 건강상태 변화 영역의 2문항을 합쳐서 '아주 나빠졌다', '조금 나빠졌다', '변하지 않았다', '조금 좋아졌다', '아주 좋아졌다'의 5 Likert 척도의 1문항으로 수정하였고, 신체 기능 영역의 '걷기가 힘들다'를 '100 미터 정도 걷기가 힘들다', '1000 미터 정도 걷기가 힘들다'의 2문항으로, '계단 오르기가 힘들다'를 '계단으로 한층 오르기가 힘들다', '계단으로 3-4층 오르기가 힘들다'의 2문항으로 각각 힘든 정도의 차이가 있도록 나누어서 46문항을 확정하였다.

2. 응답항목 및 점수화 방법 결정

응답 항목은 건강 인식 영역과 영적 건강 영역의 경우 '전혀 그렇지 않다', '거의 그렇지 않은 편이다', '보통 그런 편이다', '거의 그런 편이다', '확실하게 그렇다'의 동의를 묻는 척도를 사용하였고, 신체적 기능 영역은 '전혀 그렇지 않다', '조금 그렇다', '보통이다', '상당히 그렇다', '매우 대단히 그렇다'의 현재의 정도를 묻는 척도를 사용하였으며, 정신적 건강 영역은 '전혀 그렇지 않다', '드물게 그렇다', '가끔 그렇다', '자주 그렇다', '언제나 그렇다'의 과거 4주간 느낀 빈도를 묻는 척도를 사용하였고, 사회적 기능 영역, 통증 영역, 활력 영역, 역할 영역은 '이런 일이 전혀 없었다', '이런 일을 드물게 경험했다', '이런 일을 가끔 경험했다', '이런 일을 자주 경험했다', '언제나 그랬다'의 과거 4주간 경험의 빈도를 묻는 척도를 사용하였다. 건강관련 삶의 질 점수는 각 영역별로 최저 0점에서 최고 100점으로 변형하여 계산하였으며 가중치는 두지 않았다.

3. 대상군 선정

KQOLS의 측정심리학적 검증은 두 개의 대상군에 대하여 이루어 졌다. 한 대상군은 내적 일치도, 요인 분석, 임상타당도, 반응도를 보기 위한 것(대상군 A)이고 다른 하나는 검사-재검사법을 통한 신뢰도 검증을 위한 대상군(대상군 B)이다.

대상군 A에서 환자군은 1998년 12월-1999년 2월 사이에 6개 병원에 내원한 환자 중 뚜렷한 증상을 가지고 있거나 검사상 객관적인 이상 소견을 보이고, 적어도 2주 이상 지속되는 질병(당뇨병, 고혈압, 갑

상선기능항진증, 갑상선기능저하증, 결핵, 빈혈, 소화성궤양 질환, 천식, 심부전, 두통, 류마치스 관절염, 퇴행성관절염, 요통 등의 근골격계 질환 등)을 가지고 있는 환자였으며, 정상군은 3개 종합검진센터에 건강검진을 목적으로 방문한 사람으로서 특별한 신체 증상이 없고 검진결과가 정상인 사람이었다.

대상군 B는 일개 대학병원 사무직원으로 특정한 증상을 호소하지 않는 30명이었다.

4. 설문지의 신뢰도

대상군 B에게 KQOLS를 완성하게 한 후 2 주 후 다시 설문을 완성하게 하여 검사-재검사법에 의한 신뢰도를 검증하였다. 또한 대상군 A에게 조사한 자료를 바탕으로 내적 일치도 검증을 위해 Cronbach's alpha 값을 산출하였다.

5. 문항분석

KQOLS의 문항분석을 위하여 각 문항의 평균점수와 표준편차 그리고 문항-영역총점 상관계수를 구하였다.

6. 요인분석

대상군 A에게서 얻은 자료를 바탕으로 삶의 질 영역에 대한 확증 요인분석(confirmatory factor analysis)을 시행하였다. 요인분석은 주요인 방법(principal component method)으로 하였고, 고유근이 1이상인 경우로 요인 수를 결정하였으며, varimax 회전 후의 요인 적재값(factor loading)으로 최종 요인을 결정하였다.

7. 임상타당도

임상 타당도를 보기 위해 환자군과 정상군의 삶의 질 점수를 비교하였으며, 다중 회귀분석을 통하여 인구사회학적 변수를 보정한 후에도 정상군과 환자군의 '건강관련 삶의 질' 점수가 차이가 있는지를 알아보았다.

8. 설문지의 반응성

환자군에 대해서 선정당시 1차 설문 조사를 하고, 의학적 치료를 받고 4주(2주-6주) 뒤 재진 방문시에 2차 설문을 하였다. 또 이때 담당 의사가 환자의 상

Table 1. Sociodemographic distribution of group A and group†

		Group A		Group B
		Patients (%) (n=151)	Normal (%) (n=116)	(n=30)
Sex	Male	75 (49.7) *	43 (37.1)	14 (46.7)
	Female	76 (50.3)	73 (62.9)	16 (53.3)
Age	Under 30	8 (5.3) *	24 (20.7)	11 (36.7)
	30-39	31 (20.5)	41 (35.3)	9 (30.0)
	40-49	46 (30.5)	25 (21.6)	7 (23.3)
	50-59	40 (26.5)	23 (19.8)	3 (10.0)
	Over 60	26 (17.2)	3 (2.6)	
Income (10,000 Won/month)	Under 100	42 (27.8)	21 (18.1)	8 (26.7)
	100-199	51 (33.7)	39 (33.6)	15 (50.0)
	200-299	28 (18.5)	19 (16.4)	4 (13.3)
	300-399	15 (10.0)	17 (14.7)	3 (10.0)
	Over 400	15 (10.0)	20 (17.2)	
Marital status	Not married	16 (11.0)	27 (24.3)	11 (36.7)
	Married	114 (78.6)	78 (70.3)	18 (60.0)
	Other	15 (10.4)	6 (5.4)	1 (3.3)
Education (year)	Under 9	41 (29.5) *	22 (19.6)	1 (3.3)
	9-12년	59 (42.5)	36 (32.2)	14 (46.7)
	Over 12	39 (28.0)	54 (48.2)	15 (50.0)
Occupation	None	19 (13.6) *	8 (7.3)	
	Housekeeper	55 (39.6)	30 (27.5)	
	Professional	25 (18.0)	19 (17.4)	
	Manual laborer	10 (7.2)	5 (4.6)	
	Sales	19 (13.7)	25 (22.9)	
	Administrative	11 (7.9)	22 (20.2)	
Residence	Large City	114 (78.1)	87 (78.4)	
	Small City	18 (12.3)	15 (13.5)	
	Town/village	14 (9.6)	9 (8.1)	

* P < 0.05 (by chi-square test)

† Sum of patients and normal control in each sociodemographic factor is not same as total number due to missing data.

Table 2. Reliability of KQOLS.

Domains of Quality of life	Group A		Group B		
	Number of Items	Cronbach's alpha	First test	Seonc test	상관계수
Health status change	1	.	35.8±17.0	34.2±18.0	0.58*
Health status perception	1	.	80.8±25.2	79.2±18.7	0.22
Spiritual health	2	0.83	41.2±29.6	39.2±28.6	0.82*
Physical function	12	0.94	87.2±15.9	84.7±18.2	0.83*
Psychological health	12	0.88	87.2±15.9	84.7±18.2	0.86*
Role limitation	10	0.87	80.8±15.5	80.3±18.2	0.78*
Social function	3	0.77	90.0±15.4	87.2±16.5	0.69*
Pain	2	0.80	84.2±18.0	84.2±17.3	0.72*
Vitality	3	0.77	56.1±19.3	53.1±21.2	0.60*

* P < 0.05 (by spearman correlation test)

태에 대한 소견과 호전여부를 판단하여 기록하였다. 측정도구의 코딩과 점수 계산 및 분석은 모든 설문지 수거가 완료된 이후 (즉, 환자의 두 번째 측정 완료 후) 이루어졌으므로 환자의 삶의 질 점수가 의사의 판단에 영향을 미치지 않았다. 의사에 판단에 의해 호전군과 비호전군으로 구분하여 호전군에서 각 영역 별 점수가 어떻게 변화하는지를 보아 민감도 (sensitivity)를 보았고 비호전군에서의 점수변화를 바탕으로 특이도 (specificity)를 살펴보았다.

9. 분석방법

평균치의 비교는 student t-test를 시행하였으며, 짝지은 평균치의 비교는 paired T test를 시행하였다. 내적 일치도는 Cronbach's α coefficient로 표시하였고 검사-재검사법은 스피어만 상관계수를 이용하였다. 사회인구학적 변수를 보정하기 위해 다중 로짓 분석을 이용하여 단위 점수 증가에 따른 측정도구의 정상군과 환자군 판별능력을 조사하였다.

결 과

1. 대상군의 인구사회학적 특성

대상군 A의 경우 환자군은 151명, 정상군은 116명 이었다. 정상군은 환자군에 비해 더 젊고, 여자가 많 으며, 교육정도가 높았다. 대상군 B의 경우 남녀 비율은 비슷했고 40대 이하가 대부분이었다 (Table 1).

2 신뢰도

대상군 A의 환자군과 정상군의 설문자료를 이용하여 구한 측정도구의 내적 일치도는 신체적 기능 영역 에서 0.94로 가장 높고, 활력과 사회적 기능 영역에서 0.77로 가장 낮아서 모든 영역에서 0.7을 넘었다. 대상군 B에게 2주 검사-재검사법에 의한 상관계수를 구하였을 때 0.8이상으로 나온 것은 영적 건강, 신체 건강, 정신건강이었고, 0.7 정도로 나온 것은 역할제한, 통증, 사회적 건강이었다. 활력, 건강 상태 변화 는 0.6정도의 상관 값을 가졌고 건강상태 인식의 경우 상관계수가 없었다 (Table 2).

3. 문항분석

KQOLS의 각 문항에 대해서 문항점수의 평균 및 표준편차 값과 문항-영역총점 상관관계를 구하였다. 문항점수의 경우 각 영역간 비슷한 평균점수와 표준 편차를 보였으나 긍정적 건강 측정 문항 (EP1, EP2,

Table 3. Mean score and item-total correlation of each item of KQOLS.

Item†	Mean±SD	Item-Total correlation*
B1	87.9±23.7	0.68
B2	77.7±28.7	0.69
B3	85.6±24.3	0.67
B4	72.4±28.4	0.69
B5	75.8±28.1	0.76
B6	71.5±29.7	0.70
B7	72.5±27.5	0.76
B8	74.0±27.9	0.68
B9	77.9±26.7	0.73
B10	88.7±22.2	0.71
B11	91.3±19.6	0.70
B12	88.8±21.6	0.75
C1	49.0±32.8	0.71
C2	47.2±33.8	0.71
E1	73.5±25.1	0.62
E2	68.8±25.2	0.63
E3	68.8±25.2	0.65
E4	65.0±27.1	0.60
E5	54.9±28.9	0.69
E6	67.2±27.1	0.66
E7	67.9±25.0	0.53
E8	68.4±25.4	0.69
E9	68.2±27.1	0.70
EP1	44.6±27.5	0.37
EP2	44.3±27.5	0.35
EP3	43.7±26.9	0.45
P1	68.3±30.5	0.66
P2	56.4±30.1	0.66
R1	74.8±29.1	0.52
R2	83.6±23.6	0.58
R3	82.5±27.1	0.69
R4	84.8±25.1	0.72
R5	86.2±26.1	0.66
R6	80.4±28.2	0.74
R7	77.2±24.5	0.65
R8	68.4±26.0	0.59
R9	85.3±22.9	0.59
RP1	31.5±30.4	0.11
S1	82.6±25.3	0.61
S2	75.2±25.3	0.64
S3	82.8±25.8	0.56
V1	47.7±28.1	0.48
V2	58.5±29.4	0.47
VP1	39.1±30.0	0.01
H1	34.5±22.2	
H2	71.4±26.7	

* P < 0.05 in all items except RP1 and VP1

† Refer to appendix about each item

EP3, VP1, RP1)의 경우, 평균점수와 표준편차 점수는 해당영역의 부정적 건강 측정 문항과 다른 본분을 보였다. 문항-영역총점 상관관계에 있어서도 대부분의 영역점수와 0.4 이상의 상관관계를 보여주었으나 긍정적 건강 측정 문항은 문항-영역 총점 상관관계가 낮았고 특히 VP1, RP1은 상관관계가 없었다 (Table 3).

4. 타당도

1) 요인분석

요인분석에서 결정된 요인의 수는 9개이었으며 이에 대한 varimax 회전 후 요인 부하량을 산출한 결과 표 4와 같은 결과를 보였다. 제 1 요인은 신체건강에 대한 문항으로 구성되었으며, 제 2요인은 '정신적 건강영역'과 '활력', 제 3 요인은 '역할제한', 제 4요인은 '사회적 건강'과 '건강상태인식'과 '건강상태 변화', 제5요인은 '통증', 제 6요인은 '긍정적 정신건강 영역', 제 7요인은 영적 건강, 제 8요인은 '긍정적 역할제한'과 '긍정적 활력'으로 제9요인에는 해당 문항이 없었다. 영역고유값(eigenvalues)은 31.9로 전체 46개 항목의 69.4%를 설명하였다. (표 4).

2) 임상타당도

삶의 질 각 영역별 점수는 영적 건강과 사회적 기능을 제외한 모든 영역에서 정상군이 환자군보다 유의하게 점수가 높았다(표 5).

다중 회귀분석으로 환자군과 정상군간 차이가 있는 성별, 연령, 교육정도, 직업, 수입 등 인구사회학적 변수를 보정한 후 각 건강영역이 10점 단위 증가함에 따라 정상군이 될 대응비를 산출하였다. 영적 건강과 사회적 기능은 통계적으로 유의하지 않았고 신체적 기능은 약한(p=0.09) 관련성을 보였으며 그 밖의 건강영역은 모두 유의하였다. 즉 단위점수(10점) 상승에 따라 환자군이 될 교차비가 1 이하로서 삶의 질 측정 점수가 유의하게 정상군과 환자군을 판별할 수 있었다(표 6).

5. 반응도

4주 후 임상경과에 대해 추적 조사된 64명 가운데 호전군은 51명, 비호전군은 14명이었다. 호전군의 경우 건강상태인식과 영적 건강을 제외한 모든 영역에

Table 4. Factor analysis of KQOLS.*

Item	factor 1	2	3	4	5	6	7	8	9
B1 †	<u>0.77 ‡</u>	0.15	0.12	0.04	-0.05	-0.05	0.04	0.09	0.01
B2	<u>0.77</u>	0.18	0.17	0.08	0.04	0.09	0.07	-0.06	0.05
B3	<u>0.72</u>	0.11	0.15	0.08	0.09	0.12	0.17	-0.06	-0.19
B4	<u>0.76</u>	0.20	0.26	0.07	-0.03	0.15	-0.02	-0.01	-0.20
B5	<u>0.78</u>	0.06	0.11	0.03	-0.08	0.09	0.06	-0.03	-0.18
B6	<u>0.70</u>	0.11	0.05	0.05	0.07	0.11	0.01	0.08	-0.04
B7	<u>0.76</u>	-0.34	0.05	0.11	0.19	0.05	-0.04	0.02	-0.02
B8	<u>0.69</u>	0.16	0.20	0.01	0.12	-0.06	-0.04	-0.02	0.13
B9	<u>0.76</u>	0.10	0.14	0.12	0.30	-0.04	-0.07	0.04	0.08
B10	<u>0.69</u>	0.07	0.19	0.22	0.20	-0.09	0.02	0.11	0.37
B11	<u>0.70</u>	0.08	0.09	0.27	0.27	-0.12	0.09	0.09	0.31
B12 †	<u>0.72</u>	0.11	0.15	0.18	0.11	-0.14	0.02	0.07	0.30
C1	<u>0.03</u>	0.07	-0.01	0.09	-0.04	0.04	0.86	0.14	0.07
C2	-0.004	0.03	0.01	-0.05	-0.02	0.12	0.90	0.04	-0.07
E1	<u>0.12</u>	0.57	0.21	0.06	0.16	0.18	-0.12	-0.01	0.45
E2	<u>0.13</u>	0.53	0.20	0.03	0.25	0.35	0.04	-0.13	0.40
E3	<u>0.30</u>	0.65	0.15	0.14	0.15	0.09	0.02	0.03	-0.07
E4	<u>0.15</u>	0.70	0.03	0.13	0.20	0.12	0.06	0.03	-0.22
E5	<u>0.14</u>	0.67	0.11	0.01	0.14	0.17	0.08	-0.05	0.31
E6	<u>0.08</u>	0.77	0.13	0.17	-0.06	0.10	0.07	0.04	-0.14
E7	<u>0.06</u>	0.79	0.14	0.11	-0.01	-0.10	0.12	0.02	-0.01
E8	<u>0.13</u>	0.77	0.08	0.14	0.07	0.11	-0.09	0.04	0.03
E9	<u>0.11</u>	0.76	0.07	0.06	0.14	0.13	-0.02	0.05	0.24
EP1	<u>0.01</u>	0.16	-0.09	0.07	0.01	0.77	0.15	0.07	0.05
EP2	<u>0.09</u>	0.15	0.02	0.17	-0.01	0.59	0.01	0.44	-0.10
EP3	<u>0.03</u>	0.24	-0.05	0.06	0.09	0.76	0.14	0.20	0.02
P1	<u>0.34</u>	0.13	0.14	0.16	0.70	-0.01	-0.05	0.06	0.07
P2	<u>0.42</u>	0.28	0.17	0.17	0.62	0.07	-0.01	0.07	0.02
R1	<u>0.38</u>	0.22	0.45	0.05	0.07	-0.27	0.19	0.19	-0.03
R2	<u>0.39</u>	0.34	0.50	0.07	-0.02	-0.23	0.20	0.20	-0.13
R3	<u>0.19</u>	0.13	0.81	-0.08	0.17	-0.13	0.09	0.09	0.04
R4	<u>0.20</u>	0.17	0.81	0.09	0.05	0.01	0.04	0.04	0.08
R5	<u>0.08</u>	0.04	0.81	0.11	0.08	0.08	-0.08	-0.08	0.10
R6	<u>0.26</u>	0.07	0.78	0.21	0.04	0.03	0.02	0.02	0.06
R7	<u>0.18</u>	0.31	0.56	0.38	0.16	0.02	0.05	0.05	-0.11
R8	<u>0.20</u>	0.26	0.40	0.50	0.38	0.12	-0.08	-0.08	-0.17
R9	<u>0.17</u>	0.16	0.47	0.53	0.08	0.06	-0.03	-0.03	-0.06
RP1	<u>0.005</u>	0.10	0.02	0.04	0.16	0.17	0.07	0.74	-0.01
S1	<u>0.20</u>	0.15	0.36	0.58	-0.11	0.21	-0.05	0.02	0.02
S2	<u>0.34</u>	0.15	0.30	0.64	0.15	0.05	0.11	0.02	0.10
S3	<u>0.40</u>	0.29	0.26	0.47	0.06	0.04	0.03	0.06	0.15
V1	<u>0.15</u>	0.52	0.14	0.03	0.51	-0.05	-0.05	0.17	0.02
V2	<u>0.15</u>	-0.46	0.25	0.21	0.58	-0.05	-0.05	0.11	0.10
VP1	<u>0.09</u>	-0.02	0.02	0.02	0.01	0.09	0.09	0.82	0.15
H1	<u>0.04</u>	0.12	-0.11	0.50	0.29	0.15	0.07	0.17	-0.03
H2	<u>0.18</u>	0.32	-0.02	0.34	0.16	-0.22	0.10	-0.03	0.14
Eigenval-ue	7.97	5.92	4.57	2.59	2.38	2.26	1.78	1.74	1.24

* principal components analysis with varimax rotation

† factor loadings (Items having highest value are underlined)

‡ refer to appendix about each item

Table 5. Score of patient and normal control by health domains of KOQOLS.

Domains of Quality of life	Patients (mean±S.D)	Normal (mean±S.D.)
Health status change**	30.0±22.8	37.1±17.9
Health status perception***	67.5±28.3	78.9±23.1
Spiritual health	47.2±30.8	50.1±33.3
Physical function**	77.1±22.6	83.3±15.3
Psychological health**	57.9±17.8	64.1±17.6
Role limitation**	72.3±18.6	79.3±15.7
Social function	77.8±22.9	80.8±24.3
Pain***	55.1±28.1	70.4±24.1
Vitality**	43.8±20.8	52.0±19.9

** P < 0.01 *** P < 0.001 (by student t-test)

Table 6. Odds ratios of patients group by increase of 10 points in each health domain of KQOLS.

Domains of Quality of life	Odds ratio (95% Confidence Interval) *
Health status change**	0.86 (0.75 - 0.98)
Health status perception***	0.84 (0.75 - 0.94)
Spiritual health	1.03 (0.94 - 1.13)
Physical function**	0.87 (0.74 - 1.02)
Psychological health**	0.80 (0.68 - 0.94)
Role limitation**	0.75 (0.63 - 0.89)
Social function	0.94 (0.83 - 1.06)
Pain***	0.79 (0.70 - 0.88)
Vitality**	0.79 (0.68 - 0.91)

*by multiple logistic regression

서 치료 전후에 유의한 차이를 보여 설문이 민감하였고, 불변군의 경우 모든 건강 영역에서 유의한 차이를 보이지 않아 설문이 특이하였다. 약화군은 숫자가 적어(n=2) 제외하였다(표 7).

고 찰

이 논문의 목적은 앞서 출판된 한국형 삶의 질 측정도구(KQOLS)를 측정에 편리하도록 개정하고 이에 대한 신뢰도와 타당도를 검증하여 측정도구로서의 가치를 점검해보는 것이다.

가장 먼저 시도한 것은 문항을 변경하는 것이었다. 건강상태 변화에 대한 두 문항 '나의 건강상태가 나빠졌다고 생각한다'와 '나의 건강상태가 좋다고 생각한다'는 사실상 한 문항으로 응답할 수 있는 내용으로 생각하여 하나의 문항으로 줄여서 문항의 중복성을 줄였다. 또한 신체적 건강 영역 중 '걷기가 힘들다'와 '계단 오르기가 힘들다'라는 문항은 각각 '100미터 정도', '1,000미터 정도'와 '한층', '3-4층'으로 구분하여 두 문항씩으로 늘렸는데 이는 신체적 건강을 측정하는 문항 중 걷기와 계단 오르기 문항을 세분하기 위해서 이었다.

Table 7. Change of mean score of quality of life by treatment outcome.

Domains of Quality of life	Improved group (n=51)		Not improved group (n=16)	
	Mean	Standard error	Mean	Standard error
Health status change	13*	3	4	4
Health status perception	-3	4	-4	6
Spiritual health	0	4	7	6
Physical function**	4*	2	-2	2
Psychological health**	8*	2	3	3
Role limitation**	4*	2	1	4
Social function	7*	3	1	3
Pain***	12*	3	-2	4
Vitality**	10*	3	2	4

* $p < 0.05$ (by paired T test)

KQOLS의 신뢰도는 두 가지 방법으로 평가하였다. 첫째 방법은 내적일치도법이었다. 내적일치도법은 설문지의 문항이 하나의 구성개념을 측정하는지를 검증하는 것으로 모든 가능한 반복법을 시행한 신뢰도의 평균값과 비슷하다. KQOLS의 내적일치도법에 신뢰도 평가 결과는 Cronbach's alpha값이 한 문항으로 이루어져 있어 평가가 곤란한 건강상태인식과 건강상태 변화를 제외하고 0.77-0.94의 분포를 보여서 내적 일관성이 강함을 알 수 있었다. 일반적으로 신뢰도 계수가 0.7이상이면 임상시험과 같은 집단을 선별할 정도의 유용한 도구로 볼 수 있으므로^{9,10}, KQOLS는 신뢰도가 높은 측정도구라는 사실을 알 수 있다. 2주 검사-재검사법으로 신뢰도를 분석한 결과는 0.22-0.86의 분포를 보였다. 건강상태인식의 상관계수가 가장 낮아서 0.22였고 활력과 건강상태 변화가 0.60정도의 값을 보여주어서 중등도의 상관관계를 보였다. 앞에서 언급한 신뢰도계수 수준으로 볼 때 건강상태 변화는 신뢰도에 문제가 있다고 할 수 있다. 일반적으로 신뢰도는 문항수가 적고 문항이 애매할 때 낮아지게 된다⁷. 건강상태 인식 영역은 '건강이 좋지 않아서 내 자신이 원망스럽다'의 한 문항으로 이루어져 있는데 이 문항의 경우 문항수가 적을 뿐만 아니라 '내 자신이 원망스럽다'와 같은 모호한 표현이 있어 신뢰도가 낮게 나온 것으로 추정할 수 있다. 따라서 이 문항의 경우 개정의 필요성이 있다고 볼 수 있다. 통계적으로 의미 있는 상관성을 보였지만 건강상태변화와

활력이 0.7이하의 검사-재검사 신뢰도 계수를 보인 것도 같은 맥락으로 설명이 가능하다. 건강상태의 경우 문항수가 적은 문항이 있고 활력의 경우 문항 중 긍정적 건강 측정 문항의 문항-영역 총점 상관계수가 0.01로 의미가 없게 나온 것이 신뢰도를 떨어뜨리는데 기여를 한 것으로 추정된다. 왜냐하면 문항분석의 문항-총점 상관계수는 개념상 신뢰도와 밀접한 관련성이 있기 때문이다¹⁰.

타당도는 요인분석과 임상타당도의 두 가지로 평가하였다. KQOLS가 측정하려고 하는 삶의 질은 황금기준이 없기 때문에 타당도는 구성개념타당도를 이용하여 검증할 수밖에 없다. 구성개념타당도는 일반적으로 다속성다측정법, 요인분석, 임상타당도의 세 가지 방법으로 평가한다¹¹. 그러나 아직 우리 나라에서 삶의 질 측정도구로 검증되어 쓰이는 설문지가 거의 없으므로 다속성다측정법으로 타당도를 검증하기에는 무리가 따른다.

요인분석은 개념상 내용타당도와 구성개념타당도 모두에 해당한다¹⁰. 즉 각 문항이 어떠한 요인으로 구별되는지를 살핀다는 의미에서 내용타당도의 내용을 가지고 있고, 같은 것을 측정하는 요인에 결집하고 그렇지 않은 것을 측정하는 요인에는 상관성이 낮아 짐을 관찰함으로써 구성개념타당도의 개념과 유사하다는 것이다⁹. KQOLS에 대한 요인 분석 결과 대체로 연구자들이 애초 구성해 놓은 데로 구분이 되었지만 일부 문항에 대해서는 요인이 해체되는 결과가 나타

났다. 요인분석결과에서 가장 두드러진 것은 긍정적 건강 측정 문항들이 부정적 건강 측정 문항과는 다른 요인으로 묶인다는 것이다. 이러한 결과는 몇 가지로 이유를 생각해 볼 수 있다. 가장 가능성이 높은 것은 긍정적 건강 측정 문항이 부정적 건강 측정 문항과 같은 응답항목으로 구분되어 일부 사람에서 혼동을 일으켰을 가능성이 있다. 다시 말해 긍정적 건강 측정 문항과 부정적 건강 측정 문항 모두에 같은 응답항목을 적용함으로써 일부 사람에게 일시적인 혼동이 와서 잘못된 응답이 나왔을 가능성을 생각해 볼 수 있다. 그럼에도 정신적 건강 영역의 긍정적 문항의 경우 하나의 요인으로 묶이는 것은 이들 문항들이 독립적인 하나의 영역 예를 들어 생활만족감 같은 독립적인 영역을 측정하는 것으로 생각해 볼 수 있다. 이를 규명하기 위해서는 긍정적 건강 측정 문항의 응답항목을 따로 구분하여 다시 측정해 보면 가능한 원인이 무엇인지 좀 더 분명하게 구별할 수 있을 것이다. 또한 가지 생각해 볼 수 있는 것은 이 연구가 요인분석에 적합한가 하는 것이다. 일반적으로 요인분석을 이용하기 위해서는 응답항목이 중간 척도이어야 하고 응답은 정규분포에 가까워야 한다¹²⁾. 또한 대체로 분석하는 문항수보다 5배 (어떤 학자는 20배)에 해당하는 응답자가 있어야 하는 조건도 있다³⁾. 그러나 이 연구에서 사용한 응답항목은 5개로 중간 척도로 보기에 큰 무리가 없고 대상자수도 250명 정도로 큰 무리는 없다고 할 수 있다.

임상타당도는 임상적으로 알려진 집단을 도구가 구별할 수 있는지를 보는 것이다. 이는 환자와 정상인 같이 이미 알려진 집단간에 점수의 차이가 통계적으로 차이가 나는 지를 검증한다. 그렇지만 이 방법을 사용하기 위해서는 주의해야 하는 점이 있다. 실제로 정상인 가운데는 진단되지 않은 환자가 있어서 타당도가 과소평가 될 가능성이 있다. 이러한 이유로 실험적으로 예를 들어 충격을 경험하게 하고 불안을 측정하는 것과 같은 상황을 유발시키는 방법이 추천되기도 하는데 이러한 일이 항상 가능하지 않다는 문제점이 있다. 결국 임상타당도는 이미 알려진 집단 (known group)이 얼마나 적절한가에 성패가 달려 있다고 할 수 있다¹⁴⁾. 이 연구에서는 뚜렷한 증상, 객관적 검사 이상, 2주 이상 지속이라는 세 가지 조건으로 환자군을 모았고 이들 환자군과 정상군을 비교한

결과 단변수 분석에서는 영적 건강 영역과 사회적 건강영역에서 인구사회학적 변수로 보정한 이후에는 영적 건강 영역과 사회적 건강영역 그리고 신체적 건강에서 차이가 없었다. 영적 건강 영역의 경우 질병 유무에 따라서 크게 변화하지 않는다는 사실이 알려져 있기 때문에¹⁴⁾ 이러한 결과가 나왔을 것으로 추정할 수 있다. 사회적 건강이나 신체적 건강 영역은 질환의 영향으로 점수의 차이가 있을 것으로 예상할 수 있는데 이 번 연구에서는 통계적으로 유의한 차이는 없었다. 이러한 결과는 몇 가지로 설명할 수 있다. 첫째는 천장효과(ceiling effect)이다. 천장효과란 더 높은 점수를 충분히 판별하지 못하는 것으로 천장효과가 있는 측정도구는 가벼운 질병에 민감하지 못하다. 그런 이유로 이번 연구의 환자군 질병이 심각하지 않아 신체적 건강이나 사회적 건강을 덜 훼손하였기 때문에 그런 결과가 나왔을 가능성이 있다. 둘째는 대상수의 문제로 이 번 연구 대상수가 적어서 충분한 검정력으로 차이를 판별하지 못했을 가능성이 있다. 또한 각 영역의 문항들이 실제적인 건강상태를 대표할 수 없을 가능성도 있다. 이러한 가능성 중 어떠한 설명이 가장 올바른지를 판별하기 위해서는 대상수를 늘리고, 응답자의 천장효과를 구해보는 것과 같은 작업이 필요하다.

이 연구에서 반응도는 임상적으로 호전된 군에서 KQOLS의 각 영역의 점수가 호전되었는지로 판단하는 민감도와, 호전되지 않은 군에서 각 영역 점수도 변화하지 않는지로 판단하는 특이도로 판정하였다. 반응도는 여러 가지 수치로 나타낼 수 있지만 이런 방법은 간편하게 표현하는 장점 때문에 몇몇 설문지의 검증에 이용된 바 있다¹⁵⁾. 이 연구의 결과는 영적 건강영역과 건강상태 인식 영역을 제외하고는 모두 민감도와 특이도가 증명되었다. 영적 건강의 경우 질병 유무에 따라서 쉽게 변하지 않는 것이 원인으로 작용했을 것으로 보이며 건강상태인식의 경우 낮은 신뢰도 때문에 이러한 결과가 나왔을 것으로 추정할 수 있다. 이러한 결과로 KQOLS는 변화에 민감한 측정도구라고 할 수는 있지만 반응도에 있어서는 더 많은 사람을 대상으로 입증되어야 할 것이다.

이러한 결과를 바탕으로 KQOLS는 충분히 신뢰할 만하고 타당성이 있으며 변화에 민감한 도구라는 결론을 유출해 낼 수 있다. 그러나 다음과 같은 세밀한

개정작업이 요구된다고 할 수 있다.

첫째, 건강상태 인식에 대한 새로운 문항이 필요하다. 건강상태 인식영역의 문항은 검사-재검사 신뢰도가 낮고 민감도에도 문제가 있다는 결과로 미루어 문항내용을 수정하여 재검증하는 것이 필요할 것으로 보인다.

둘째, 긍정적 건강 측정 문항에 대한 재검증이 필요하다. 긍정적 건강 측정 문항은 원래 정했던 요인으로 묶이지 않을 뿐만 아니라 문항-영역총점상관관계도 낮게 나타났는데, 이에 대해서는 응답항목을 따로 두는 등의 방법으로 불일치의 원인을 찾아야 할 것이다.

셋째, 반응도에 대한 추가적 검증이 필요하다. KQOLS의 경우 변화에 민감한 측정도구라는 사실이 입증되었지만 더 많은 대상 수를 바탕으로 더욱 더 다양한 반응도 측정치를 제시하는 연구가 필요할 것이다.

참 고 문 헌

1. 한림의대가정의학교실편. 삶의 질 측정의 이론과 실제. 초판, 고려의학, 서울;1999.
2. 김수영. 삶의 질 측정도구의 번역. 가정의학회지 1998;19(11);1034-42.
3. Bowling A. Measuring health A review of quality of life measurement scales. 1st ed. Milton ; Open University press, 1991.
4. Hunt SM. Cross-cultural issues in the use of quality of life measures in randomized controlled trials In Staquet MJ (ed). Quality of life assessment in clinical trials. 1st ed. Oxford; Oxford university press, 1998;51-67.
5. 심재용, 이정권, 김수영, 원장원, 선우성, 박훈기, 신호철. 한국형 건강관련 삶의 질 측정도구의 개발: 증상체집에서 사전조사까지. 가정의학회지 1999;19(12):1197-208.
6. Juniper EF, Huyatt GH, Jaeschke R. How to develop and validate a new health-related quality of life instrument. In Quality of Life and Pharmacoeconomics in Clinical Trials. 2nd ed. Philadelphia: Lippincott-Raven, 1996;49-56.
7. 채서일. 사회과학조사 방법론. 수문사 1st ed. 1988.
8. 김수영. 가정의학회지에 게재된 논문 중 설문지와 관련된 논문 분석. 가정의학회지 1996;17(9):748.
9. Murphy KR, Davidshofer CO. Psychological testing. 1st ed 1988 Prentice-Hall, New Jersey.
10. McDowell. Measuring health. 1st ed. New York; Oxford university press 1996.
11. Siu AL, Ouslander JG, Osterweil D. Change in self-reported functioning in older persons entering a residential care facility. J Clin Epidemiol 1993;46:1093-101.
12. Staquet MJ. Quality of life assessment in clinical trials. 1st ed. Oxford; Oxford university press, 1998.
13. Boyle GJ. Self-report measures of depression: some psychometric considerations. Br J Clin Psychol 1985;24:45-9.
14. Spilker B. Introduction. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1996. p. 1-10.
15. Rosen RC, Riley A, Wagner G, Osterloh IH, Kirkpatrick J, Mishra A. The International index of erectile function (IIEF) :a multidimensional scale for assessment of erectile dysfunction. Urology 1997;49:822-30

— Abstract —

The development of Korean health related quality of life scale(KQUOLS)

— Testing reliability and validity—

Kim Soo-Young, M.D., Shim Jae-Yong, M.D.^{*}, Won Jang-Won, M.D.^{**},
Sun Woo-Sung, M.D.^{***}, Park Hoon-Ki, M.D.^{****}, Lee Jung-kwon, M.D.^{****}

Department of Family Medicine, Hallym University College of Medicine,
Department of Family Medicine, Yonsei University College of Medicine^{*},
Department of Family Medicine, Hanyang University College of Medicine^{**},
Department of Family Medicine, Kyung Hee University College of Medicine^{***},
Department of Family Medicine, Ulsan University College of Medicine^{****},

Background : Korean health related Quality of Life Scale(KQOLS) was developed to measure the quality of life reflecting Korean's own language expression and culture. In order to confide in the results of KQOLS, it is necessary to establish verification of the psychometric properties(reliability, validity, responsiveness).

Methods : KQOLS's final 46 items and item option responses were established by expert panels. Reliability was tested by 2 weeks test-retest method and internal consistency method. Validity was tested by factor analysis and clinical validity. Sensitivity and specificity for detecting treatment-related changes were also tested.

Results : A high degree of internal consistency was observed for each of domains(Cronbach's alpha value of 0.77 or higher). Two weeks test-retest reliability correlation coefficients scores were highly significant except health perception domain. A principal components analysis identified 9 factors with eight values greater than 1.0 which were approximately the same as intended domains. Normal control group had statistically significant higher scores than patients group except in spiritual health and social function domains. Improved group of patients had statistically significant higher scores than non-improved group of patients after treatment except in spiritual health and health status perception domains.

Conclusion : Korean health related Quality of Life Scale(KQOLS) is both a valid and a reliable instrument, but fine refinements such as items revision and item response options modification will be needed. (J Korean Acad Fam Med 2000;21 : 382~394)

Key words : quality of life, validity and reliability, responsiveness, Korea.

Appendix: KQOLS의 문항

H1 1년 전보다 건강상태가 어떻다고 생각하십니까?

H2 건강이 좋지 않아서 내 자신이 원망스럽다

C1 나의 신념 혹은 신앙으로 삶의 어려움을 극복할 수 있다

C2 신념 혹은 신앙이 내 삶에 의미가 있다

B1 100 미터 정도 걷기가 힘들다

B2 1000 미터 정도 걷기가 힘들다

B3 계단으로 한 층 오르기가 힘들다

B4 계단으로 3-4 층 오르기가 힘들다

B5 가볍게 뛰기가 힘들다

B6 평소에 들던 무거운 물건을 들기가 힘이 든다

B7 평소에 하던 운동을 하기가 힘이 든다

B8 차를 타고 장거리 여행하기가 힘들다

B9 버스나 지하철을 타고 출근하기가 힘이 든다

B10 힘이 들어서 외출할 수 없다

B11 혼자서 목욕하기 힘이 든다

B12 힘이 들어서 가고 싶은 곳을 마음대로 갈 수가 없다

E1 소극적이고 자신이 없었다

E2 재미가 없고 의욕이 나지 않았다

E3 공연히 걱정이 생겼다

EP1 즐겁고 행복하였다

E4 작은 일에도 신경이 날카로워지고 긴장이 되었다

E5 혼자라는 느낌이 들고 외로웠다

EP2 정신적으로 안정되고 여유 있는 편이었다

E6 공연히 불안하고 초조하였다

E7 주변 사람들에게 작은 일에도 화를 내고 짜증스러웠다

E8 속이 상하고 실망하였다

E9 우울한 느낌이 들었다

EP3 기분이 좋고 만족스러웠다

R1 다른 사람의 일을 도와주지 못했다

R2 나의 일을 스스로 하지 못하고 다른 사람의 도움을 청해야 했다

R3 평소에 하던 집안 청소를 하지 못했다

R4 평소에 하던 쇼핑이나 장보기를 하지 못했다

R5 식사 준비나 설거지를 하지 못했다

R6 빨래를 하거나 아이들을 돌보아주기가 힘들었다

R7 나에게 주어진 일을 하지 못하고 뒤로 미루었다

R8 평소에 비해 일(업무) 능률이 떨어졌다

R9 직장 동료들과 협조해서 할 일을 하지 못했다

RP1 평소에 비해 일을 하는 것이 쉬웠다

S1 친척들이나 친구들과의 모임에 나가지 못했다

S2 평소에 하던 취미 생활, 여가 활동을 하기가 힘들었다

S3 집안의 대소사에 참가하기가 어려웠다

P1 통증 때문에 고생하였다

P2 몸이 여기 저기 아팠다

V1 쉽게 피곤함을 느꼈다

V2 기운이 없고 무기력하였다

VP1 쉽게 지치지 않았다