

# 여성 복합성요실금 환자에서 한국어판 Incontinence Quality of Life 설문도구의 타당도 및 신뢰도에 관한 정신측정학적 특성

서울대학교 의과대학 비뇨기과학교실

오승준 · 손환철 · 김수용

## Psychometric Properties of the Korean Version of the Incontinence Quality of Life Instrument in Women with Stress Urinary Incontinence

Seung-June Oh, Hwancheol Son, Soo Woong Kim

From the Department of Urology, Seoul National University College of Medicine, Seoul, Korea

**Purpose:** To evaluate the psychometric properties of the Korean version of the Incontinence-Quality of Life (I-QOL) instrument in Korean women with symptoms of stress urinary incontinence.

**Material and Methods:** Prospective study involving women with stress urinary incontinence, who eventually underwent anti-incontinence surgery (n=89) was undertaken. Convergent and discriminant validity were assessed. Reliability was assessed by calculating Cronbach's alpha coefficient, and stability was assessed as well. Sensitivity to clinical change before and after operation was also evaluated.

**Results:** The Korean version of I-QOL proved to be not difficult to understand and acceptable to the patients. Significant association was found between the I-QoL scores and clinical parameters including overall symptom severity. The domain score were significantly different between control group and patient groups. Convergent validity of the I-QOL also proved adequate. Cronbach's alpha coefficients (>0.78) indicated reasonable internal consistency. Results from test-retest showed reproducible. I-QOL domains were generally responsive to clinical efficacy variables; the I-QOL also showed statistically significant sensitivity to changes in patients' perceptions of bladder condition in all domains.

**Conclusion:** The Korean version of the I-QOL is valid and reliable for clinical use in Korean patients with stress urinary incontinence. (J Korean Continence Soc 2009;13:51-60)

**Key Words:** urination, stress urinary incontinence, questionnaire, translations

접수일자: 2009년 4월 17일, 수정일자: 2009년 5월 8일, 채택일자: 2009년 5월 11일

교신저자: 오승준, 서울대학교 의과대학 비뇨기과학교실

서울시 종로구 연건동 28번지 ☎ 110-744

Tel: 02-2072-2406, Fax: 02-742-4665, E-mail: sjo@snu.ac.kr

\* 이 연구는 Eli Lilly and Company (과제번호: 06-2002-029-0)의 지원에 의하여 이루어진 것임.

## 서 론

배뇨증상을 가진 환자들에 있어서 환자들의 주관적인 배뇨 증상을 객관화하고 정량화하여, 배뇨증상이 환자들의 삶의 질에 미치는 영향을 평가하는 일은 환자들의 치료방침을 결정하는데 대단히 중요하다. 이러한 배경하에서 이때까지 많은 배뇨증상 측정 도구들이 개발되어 활용되고 있다 (1-6).

Incontinence Quality of Life (I-QOL)은 미국에서 Wagner 등 (5)에 의해 개발된 영문판 요실금 측정도구로서 각 문항들은 간단하고 짧으며 응답자들이 이해하기 쉽게 되어 있다. 이 설문지는 초등학교 5학년 수준이면 이해할 수 있도록 만들어진 22개의 질문 항목으로 구성된 자가 기입형 설문지로서 남녀 요실금 환자들의 삶의 질을 측정하는데 있어서 재현성과 타당도가 있는 것으로 알려져 있다. I-QOL 설문지는 프랑스어, 스페인어, 스웨덴어, 독일어 등으로 번역되어 정신훈정확적 타당성이 검증된 바 있고 이들 언어에서 3개의 소범주 (subscale) 점수와 총점은 임상증상을 잘 반영하며 내적 일치도와 재현성이 높다고 증명되어 있다 (7). 이 외에도 비록 타당도를 거치지는 못하였으나 영국 영어, 노르웨이어, 핀란드어, 덴마크어, 이탈리아어, 네덜란드어 등 11개국 언어로 직역되어 임상에 활용되고 있다 (2). 최근 WHO는 이때까지 출판된 문헌 근거를 검토하여 본 결과 I-QOL이 충분한 타당성과 신빙성이 있으며 변화에 민감하게 반응하는 기준에 부합하여 요실금에 대한 ‘highly recommended’되는 Grade A등급의 설문지로 인정하였다 (8). I-QOL은 이미 원저자로부터 번역에 관한 승인을 받고 국내에서 오 등 (7)에 의하여 한국어로 언어타당도 검증 과정을 거친 바 있다.

본 연구는 국내 배뇨장애 환자들의 임상진료에 있어 그 증상 정도를 객관화하는데 도움을 주고 국내 연구자들이 이를 임상 연구에 활용할 수 있도록 하기 위하여 이전에 한국어로 번역한 I-QOL 배뇨증상 설문측정도구 (9)를 여성복합성요실금 환자들을 대상으로 신뢰도와 타당도를 평가하고자 하였다.

## 대상 및 방법

### 가. 설문도구 및 부가 자료조사지

번역 대상 I-QOL 설문지의 각 항목 답변은 증상의 심한 정도에 따라 1 ("Extremely")에서 5 ("Not at all")중 하나를 선택하여 응답하도록 이루어져 있다. 설문도구 내용은 Avoidance and Limiting Behavior (ALB), Psychosocial Impacts (PS), Social Embarrassment (SE) 등 3개의 소범주들로 구성되어 총 22개 문항으로 이루어져 있다. ALB 소범주 항목은 1, 2, 3, 4, 10, 11, 13, 20번 등 8개 문항이고 PS 소범주 항목은 5, 6, 7, 9, 15, 16, 17, 21, 22번 등 9개 문항, SE 소범주 항목은 8, 12, 14, 18, 19 등 5개 질문 항목으로 이루어져 있다. 영역별 점수산정 방법은 원저자가 제공해 준 지침을 따랐다 (10).

본 연구를 수행하는데 필요한 부가자료조사지로는 임상 진단을 내리기 위한 배뇨증상 정보 조사지, 국제적으로 공인된 한글판 일반 건강 상태 평가 설문지인 SF-36, 설문지 작성과 관련된 질문지 및 인적사항 조사지 등을 이용하였다. 이 중에서 전반적인 증상의 심한 정도를 묻는 질문은 ‘귀하의 요실금이나 소변문제는 어느 정도 심각하다고 생각하십니까?’로 하였고 응답으로는 ‘약간 심각, 상당히 심각, 매우 심각’ 중에서 선택하도록 하였다. 이 연구에 이용한 한글판 SF-36 설문도구는 I-QOL의 타당도와 신빙도를 검증하는데 이용하였으며 QualityMetric사 (Lincoln, RI, U.S.A.)로부터 사용 라이선스를 받았고 해당 점수 산정지침에 따랐다. 그리고 객관적인 지표로서 패드검사 등을 이용하였다 (11).

### 나. 연구대상 환자

신뢰도, 타당도 및 반응도 조사를 위해 모집된 환자군의 포함기준으로는 복합성요실금을 호소하는 여성으로 하되, 안정적인 임상 상태를 보이고 한글의 해독이 가능하여 동의서를 작성할 수 있는 자, 항요실금 수술을 받게 될 환자로 하였다. 제외기준으로는 복합성요실금이나 복합성요실금에 해당되지 않는 경우, 방광암, 당뇨병성 신경병증, 선천적 비뇨기 질환, 중추신경계 및 말초신경 기능에 이상이 있는자, 요로감염, 18

세 이하, 글을 읽을 수 없거나 쓸 수 없는 경우, 중대한 과거병력이 있거나 패드검사 등을 시행할 수 없는 심한 신체적 장애 등으로 하였다. 복압성요실금의 정의는 가장 최근 규정된 Standardized Terminology of the International Continence Society (12)에 의거하였다. 요실금의 세부구분에 대한 형태 판단은 초진시 환자의 병력, 요역학검사 소견에 근거한 연구자의 판단에 일차적으로 근거하였다. 검사-재검사는 환자군에서 실시하였고 치료전에 2주 간격을 두고 동일한 설문지를 작성하게 하였다. 이기간 동안의 점수변화를 Guyatt responsiveness index에 산정하였다.

대조군의 포함기준으로는 급만성 질환자, 이전에 배뇨증상으로 치료를 받지 아니한 자, 주관적인 배뇨증상이 없는 자로 하였고 대조군의 제외기준으로는 복압성요실금이나 요로감염 등으로 인하여 치료를 받은 적이 있는 자로 하였다. 결론적으로 대조군은 배뇨증상을 동반하지 않은 무증상 신결석 환자를 대상으로 선정하였다.

연구 대상자수 산출은 적절한, 수술군과 대조군 사이의 standardized difference는 0.8로 가정하면 (two-sided type I error은 0.05이고 power는 0.9) 각군의 sample size는 33으로 계산되는데 환자의 순응도를 고려하여 40명으로 판단하였다. 그러나 본 연구에서는 통계 power를 높이기 위해 더 충분한 sample size를 두는 것으로 산정하여 80명 이상을 두는 것으로 산정하였다.

**다. 정신측정학적 검사**  
(psychometric performance)

2003년 4월부터 2004년 3월까지 약 12개월간 서울대학교병원 비뇨기과 및 보라매병원 비뇨기과 외래 환자로부터 크게 다음과 같은 3가지의 조사를 통하여 전향적으로 수술전후 자료를 수집하였다.

**가) 타당도 조사**

타당도 조사는 크게 구별타당도 (discriminant validity)와 수렴타당도 (convergent validity)로 구성하였다. 구별타당도 조사는 최초 I-QOL 점수와 최초 주관적 및 객관적 요실금 증상의 심한 정도와 비교하는 것이었다. 주관적 및 객관적 요실금 증상의 심한 정도 지표로

는 지난 1년간 요실금 치료를 위하여 병원을 찾은 횟수, 지난 1달간 요실금횟수, 패드검사에서 측정된 요실금의 양, 환자본인이 느낀 요실금의 심한 정도 범주 (경한정도, 중등도, 심한정도) 등이다. 이러한 요실금의 심한 정도와 최초 I-QOL 점수 사이의 상관관계를 조사하였다. 또한 이 I-QOL 점수를 대조군의 점수와 비교하였다. 수렴타당도는 한글판 SF-36설문지를 이용하여 이 설문지 점수와 I-QOL 점수 간에 비교하는 것으로 하였다. I-QOL 총점수 및 영역별 점수를 SF-36의 영역 (domain)별 점수와 소범주 점수와 비교하였다.

**나) 신뢰도 (reliability) 조사**

이 조사는 크게 내부 일치도 (internal consistency) 조사와 안정성 (stability, test-retest analysis) 조사로 이루어 졌다. 내부일치도 조사는 개개 질문항목과 영역별 점수 (또는 총점수) 간의 연관 정도를 평가하는 것이었다. 안정성검사는 시간간격을 두고 동일한 질문항목에 대하여 반응의 차이가 있는가를 평가하는 것으로서 처음 설문지를 완성한 후 동일한 환자들 대상으로 동일한 설문지를 2주 후에 재조사하였다. 2주의 시간은 처음 환자가 설문지를 완성한 후 그 응답내용을 기억하기 힘든 시간이라고 가정한 것이다.

**다) 반응도 또는 변화 민감도 (responsiveness or sensitivity to change) 조사**

증상변화는 미미하나 중요한 변화가 있었을 때 I-QOL 설문지가 이를 민감하게 반영해 줄 수 있는지를 평가하기 위한 것이 반응도 조사의 목적이다. 환자군의 모든 환자들에서 항요실금 수술을 시행하였다. 수술 전 2주째와 수술 후 3개월째 I-QOL 설문지 응답을 얻고 두 시점 사이의 점수변화가 호전되는 방향으로 유의하게 변화하였는지를 평가하였다. 또한 점수변화를 개월당 요실금 횟수와 항요실금 수술 후 환자의 전반적인 증상인지 정도 등 2가지의 독립적인 측정 항목들과 비교하였다. 또한 변화의 크기를 수치화하기 위해 effect size, standardized response mean, Guyatt responsiveness index를 계산하였다. Effect size는 (평균 변화점수) / (술전 점수의 표준편차)로 구하였고 (13), standardized response mean은 (평균 변화점수) / (변화의 표준편차)로 구하였다 (14). Guyatt responsiveness index

는 (수술군의 평균 변화점수)/ (대조군의 변화 점수의 표준편차)로서 구하는 것인데 (15), 본 연구에서는 환자군의 test-retest에서 얻은 결과를 대조군의 자료로 이용하였다. 이 모든 수치들은 절대값이 0.2이면 약한 반응을, 0.5은 중등도의 반응을, 0.8이상이면 큰 반응을 보여준다고 간주된다 (16).

**라) 자료 수집 및 통계분석 방법**

본 연구계획은 서울대학교병원 임상연구윤리위원회의 심의를 거쳤으며 조사 전 환자의 동의를 얻었다. 본 연구에서의 설문지 기입은 원칙적으로 환자가 자가 기입으로 완성하도록 하였다.

통계분석은 모수적 방법을 기본으로 하였으나 자료수가 작거나 또는 분포의 형태가 종 모양과 매우 상이한 경우는 비모수적 방법을 보조적으로 사용하였다. 통계적 유의성은 양측검정을 기준으로 유의수준 5%를 사용하였다. 구별 타당도를 통계적으로 검정하기 위하여 I-QOL 총 점수 및 소변주별 총 점수를 사용하여 치료전 환자군의 I-QOL 점수와 전반적인 배뇨증상의 심한 정도 범주 사이의 상관관계를 조사하는데는 Pearson 상관계수를, 환자군의 I-QOL 점수를 대조군의 I-QOL 점수와 비교하는 조사에는 independent two sample t-test를 이용하였다. 수렴 타당도를 조사하기 위한 I-QOL 총 점수 및 소변주별 총 점수와 SF-36의 영역별 점수와의 상관관계 조사에는 Pearson 상관계수를 이용하여 평가하였다. 내부일치도는 소변주별 Cronbach 알파 계수를 이용하여 평가하였다. 안정성 조사에는 Chi square 검증과 Sperman 상관계수를 구하였다. 반응도 또는 변화 민감도에 대한 분석에서 술전과 술후 I-QOL 의 영역별 점수 변화를 paired t-test를 통하여 비교하고 effect size, standardized response mean, Guyatt responsiveness index를 구하였다. 그리고 전반적인 증상 변화와 요실금의 횡수 변화와 각 영역별 점수 사이의 상관관계를 구하였다.

**결 과**

**가. Readability and acceptability**

환자들의 인적사항은 표 1와 같다. 응답이 불완전하

**Table 1.** Demographics of the patients (n=89)

Age (y.o. ± SD)	55.8±9.7
Previous urinary tract surgery	
Yes	3 (3.4%)
No	86 (96.6%)
Occupation	
Full-time	21 (23.6%)
Part-time	3 (3.4%)
None or retired	2 (2.2%)
Housewife	63 (70.8%)
Education	
None	1 (1.1%)
Elementary school	20 (22.5%)
Middle school	29 (32.6%)
High school	28 (31.5%)
College or above	11 (12.3%)
Level of Income (Korean Won)	
Less than 500,000	4 (4.5%)
510,000-1,000,000	12 (13.5%)
1,010,000-1,500,000	16 (18.0%)
1,510,000-2,000,000	15 (16.9%)
2,000,000 or above	42 (47.2%)
Marital status	
Married	78 (87.6%)
Spouse died or separated	9 (10.1%)
Divorced	2(2.2%)
Menopausal status	
Menstruating	26(29.2%)
Menopause	63(70.8%)
Incontinence	
Duration	Median 60.0 (range 9-360) months
Pad test	Median 16.9 (range 10-150) grams
Patient-perceived overall symptom severity	
Mild	23 (25.8%)
Moderate	41(46.1%)
Severe	25(28.1%)

거나 누락이 있었던 경우를 제외한 총 89명의 자료가 분석되었다. I-QOL 설문지를 작성하는데 걸린 시간은 평균 5.6 (±2.0, SD)분이었다. 증상정도 질문지의 질문내용이 명확하고 이해하기 쉽게 되어 있느냐는 질문에는 그렇다는 응답자가 81명 (91.1%)였다. 증상조사 질문내용이 대상자의 배뇨상황과 관련있고 의미있는 내용이라는 질문에는 그렇다는 응답이 82명 (92.1%)였

**Table 2.** Correlation between I-QOL domains and other subjective clinical parameters (n=89)

I-QOL domains	Spearman's correlation coefficient (p-value)			
	Duration of incontinence	No. of leak per month	Overall symptom severity	Pad amount
ALB	-0.1 (p=0.2)	-0.1 (p=0.1)	-0.40 (p=0.001)	-0.21 (p=0.06)
PS	-0.1 (p=0.3)	-0.23 (p=0.03)	-0.56 (p=0.001)	-0.20 (p=0.08)
SE	-0.02 (p=0.8)	-0.32 (p=0.002)	-0.50 (p=0.001)	-0.21 (p=0.07)
Sum	-0.1 (p=0.3)	-0.25 (p=0.02)	-0.54 (p=0.001)	-0.23 (p=0.04)

Avoidance and limiting Behaviors (ALB), Psychological Impact (PS), Social Embarrassment (SE)

**Table 3.** I-QOL scores between the patients group and the control group

I-QOL domains	Patient group (n=89)	Control group (n=50)	*Difference (p-value)
ALB	47.5±23.0	91.2±9.7	p<0.0001
PS	52.2±26.0	95.1±8.5	p<0.0001
SE	32.8±23.4	92.3±12.4	p<0.0001
Sum	46.1±22.1	93.0±9.0	p<0.0001

Avoidance and limiting Behaviors (ALB), Psychological Impact (PS), Social Embarrassment (SE)

\*Independent two sample t-test

다. 질문지의 전체 형식이 작성하기 쉽게 되어 있느냐는 질문에 그렇다는 응답은 85명 (95.5%)이었다.

#### 나. 타당도 검증

##### 1) 구별 타당도 (discriminant validity)

##### 1) 구별 타당도 조사

##### (1) 치료전 환자군 I-QOL 점수와 환자 본인이 느낀 전반적인 배뇨증상의 심한 정도와 상관분석

치료전 환자군 I-QOL 점수와 환자 본인이 느낀 전반적인 배뇨증상의 심한 정도와의 상관분석에서 모든 I-QOL 영역에서 0.4 이상의 매우 유의한 음의 상관관계가 관측되었다 (p<0.05). 패드양은 I-QOL 총점사이에 약한 음의 상관관계가 관찰되었다. 지난 1달간의 요실금횟수와 의 상관관계에서는 PS, SE영역에서 약한 음의 상관관계가 관찰되었으나 요실금이 있어 온 기간과 병원을 찾은 횟수와는 유의한 상관관계를 발견할 수 없었다 (p>0.05) (Table 2).

##### (2) 환자군의 I-QOL 점수와 대조군의 I-QOL 점수 비교

대조군과의 환자군에서의 I-QOL 점수 비교 결과 모든 영역에서 양군간에는 유의한 차이가 있었다 (p<0.05) (Table 3).

##### 2) 수렴 타당도 (convergent validity)

환자군에서 SF-36 점수와 I-QOL 점수간의 상관분석 결과 I-QOL의 PS영역은 SF-36의 MH, PF, RE, SF영역과 유의한 상관관계가 있었다 (p<0.05). I-QOL의 SE영역은 SF-36의 MH영역과 유의한 상관관계가 있었다. I-QOL의 총점은 SF-36의 PF, SF 영역들과 유의한 상관관계가 있었다 (p<0.05) (Table 4).

#### 다. 신뢰도 검증

##### 1) 내부 일치도 (internal consistency)

각 영역별 Cronbach 계수는 0.78~0.89였으며 I-QOL 총점은 0.79를 보여 내적일치도가 충분히 수용할 정도였다. PS영역에서 가장 높은 내적 일치도를 보였다 (Table 5).

**Table 4.** Correlation between I-QOL domains and SF-36 domains (n=89)

I-QOL domains	SF-36 domains							
	BP	GHP	MH	PF	RE	RP	SF	V
ALB	-0.02 (p=0.8)	-0.002 (p=0.9)	0.04 (p=0.6)	0.20 (p=0.06)	0.09 (p=0.3)	0.10 (p=0.3)	0.15 (p=0.1)	-0.008 (p=0.9)
PS	0.16 (p=0.1)	0.02 (p=0.8)	0.26 (p=0.01)	0.26 (p=0.01)	0.24 (p=0.02)	0.20 (p=0.06)	0.33 (p=0.002)	0.13 (p=0.1)
SE	0.14 (p=0.1)	0.10 (p=0.3)	0.26 (p=0.01)	0.17 (p=0.1)	0.14 (p=0.1)	0.10 (p=0.3)	0.17 (p=0.1)	0.06 (p=0.5)
Sum	0.10 (p=0.3)	0.03 (p=0.7)	0.20 (p=0.05)	0.24 (p=0.02)	0.18 (p=0.07)	0.15 (p=0.1)	0.26 (p=0.01)	0.07 (p=0.4)

\*Pearson's Correlation Coefficient (p-value)

I-QOL items: Avoidance and limiting Behaviors (ALB), Psychological Impact (PS), Social Embarrassment (SE)

SF-36 items: Physical functioning (PF), Role physical (RP), Body pain (BP), General health perceptions (GHP), Vitality (V), Social functioning (SF), Role emotional (RE), Mental health (MH)

**Table 5.** Internal consistency (n=89)

I-QOL domains	Cronbach's alpha
ALB	0.84
PS	0.89
SE	0.78
Sum	0.79

Avoidance and Limiting Behaviors (ALB), Psychological Impact (PS), Social Embarrassment (SE)

**2) 안정성 (test-retest analysis)**

2주간격을 두고 실시한 각 항목마다 test와 retest의 응답을 비교 분석한 결과 9번과 18번 문항에서는 유의한 차이가 관찰되었다. 10, 11, 12, 15, 16번 문항에서도 차이가 있었으나 통계적으로 유의하지는 않았다. Spearman 상관계수는 0.46~0.82 사이로 모든 항목간에 통계적으로 유의한 상관관계를 보였다 (p<0.05) (Table 6).

**다. 반응도 또는 민감도 조사**

**1) 치료전후의 I-QOL 점수 변화량 분석**

수술전에 비하여 모든 영역에서 수술후 I-QOL점수가 유의하게 증가하였다 (p<0.05). 그 변화는 양의 변화로서 증상이 호전되는 방향으로 변화하였다. Standardized response mean은 1.1~1.5, effect size는 1.2~2.0, Guyatt responsiveness index는 2.2~3.0으로 관찰되었다 (Table 7). 따라서 술 전과 술 후 조사한 한글

판 I-QOL 설문지는 모든 영역에서 의미있는 점수 변화를 보여주는 것으로 판단되었다.

**2) 수술전후 I-QOL 점수 변화량과의 주관적 및 객관적 지표 변화와의 관련성**

수술전후 전반적 증상의 심한정도와 월 요실금횟수 등의 지표의 변화와 I-QOL점수의 총점 또는 영역별 점수 사이에 유의한 상관관계가 관찰되었다 (p<0.05). 수술전후의 패드무게 변화와 I-QOL의 총점 또는 ALB 영역 점수 사이에는 약한 상관관계를 관찰할 수 있었다 (p<0.05). Spearman 상관계수는 0.31에서 0.58를 보였다. 술 전과 술 후에 조사한 요실금의 횟수간 차이와 I-QOL 영역별 점수변화 사이에는 모든 영역에서 의미있는 상관관계가 있었으며 (p<0.05), Spearman 상관계수는 0.27에서 0.38였다 (Table 7). 따라서 한글판 I-QOL는 환자의 술 전과 술 후 증상의 변화를 잘 반영하는 것으로 판단되었다.

**고 찰**

특정 질병으로 발생하는 증상들을 객관화하기 위한 목적으로 특정 언어로 개발된 설문 도구는 우선 해당 질환을 가진 환자들에 적용하여 보고 설문 도구가 의도한 목적으로 임상이나 연구에 이용함에 있어서 유용한지 그 신뢰도와 타당도를 검증하게 된다. 이렇게 검증된 설문도구를 언어와 문화배경이 다른 인구집단을 대상으로 적용할 때는 설문도구를 번역한 후 임상이나 연구에 적용하기 전에 해당 인구집단에서도 동일하게

**Table 6.** The difference and correlation between test-retest items (n=37)

I-Qol item	Difference (p-value*)	Spearman correlation coefficient (p-value)
1	0.4	0.49 (p=0.001)
2	0.5	0.57 (p=0.0001)
3	1.0	0.63 (p=0.0001)
4	0.4	0.50 (p=0.0001)
5	0.1	0.71 (p=0.0001)
6	0.6	0.68 (p=0.0001)
7	0.3	0.74 (p=0.0001)
8	0.7	0.82 (p=0.0001)
9	0.02	0.60 (p=0.0001)
10	0.06	0.65 (p=0.0001)
11	0.06	0.68 (p=0.0001)
12	0.08	0.46 (p=0.001)
13	0.3	0.63 (p=0.0001)
14	0.4	0.51 (p=0.0001)
15	0.05	0.70 (p=0.0001)
16	0.07	0.71 (p=0.0001)
17	0.4	0.75 (p=0.0001)
18	0.04	0.52 (p=0.0001)
19	0.3	0.66 (p=0.0001)
20	0.7	0.72 (p=0.0001)
21	0.1	0.68 (p=0.0001)
22	0.2	0.63 (p=0.0001)

\*Chi-square test

**Table 7.** Perceived change from baseline to end of treatment in medical group and surgical group (n=89)

Domain	Baseline mean score±SD	Post-treatment score±SD	Paired t-test (p-value)	SRM	ES	GRI	Spearman correlation coefficient (p-value)		Pad amount
							Overall symptom severity	No. of leakage	
ALB	47.5±23.0	82.7±19.4	p<0.0001	1.3	1.5	2.4	-0.39 (p=0.0001)	-0.27 (p=0.01)	-0.22 (p=0.05)
PS	52.2±26.0	85.0±20.4	p<0.0001	1.1	1.2	2.2	-0.47 (p=0.0001)	-0.31 (p=0.004)	-0.22 (p=0.5)
SE	32.8±23.4	78.4±24.7	p<0.0001	1.5	2.0	3.0	-0.43 (p=0.0001)	-0.38 (p=0.001)	-0.14 (p=0.2)
Sum	46.1±22.1	82.7±19.8	p<0.0001	1.3	1.6	2.9	-0.47 (p=0.0001)	-0.34 (p=0.001)	-0.23 (p=0.04)

Avoidance and limiting Behaviors (ALB), Psychological Impact (PS), Social Embarrassment (SE)  
SRM, Standardized response mean ES, Effect size; GRI, Guyatt responsiveness index

유용한 도구로 쓰일수 있는지를 검증하는 과정이 반드시 필요하다. 번역과정만 거친후 검증과정을 거치지 않고 임상에 적용하는 경우도 있다. 하지만 해당언어로 번역한 설문도구가 증상이 있는 환자와 증상이 없는 일반인을 실제 구별 할 수 있는가, 치료전후로 설문지가 증상의 변화를 객관적으로 나타낼 수 있는가, 동

일한 정도의 증상을 재현성있게 표현할 수 있는가, 번역된 문장들이 비슷한 내용을 묻는 질문에 응답이 증상의 정도를 일관되게 표현해줄 수 있는가 등의 질문에 대한 비판에 매우 취약한 상태로 남게된다. 따라서 번역후에 환자를 대상으로 실제 번역된 설문지가 유용한가를 평가하는 작업은 완성단계에 있어서 반드시 짚

고 넘어갈 문제이다. 또한 배뇨관련 설문 도구가 특정 질환에서 유용성이 입증되었다 하더라도 질환의 본질이 달라질 경우는 해당 질환을 가진 환자들을 대상으로 다시 동일하거나 유사한 검증과정을 거쳐야 한다. 예를 든다면 복잡성요실금 환자들에 대하여 유용성이 확립된 설문지가 과활동성방광을 가진 환자들에 대하여 동일한 유용성이 있다고 볼 수는 없다.

외국어로 먼저 제작되어 현재까지 한글판으로 번역되고 한국인에 대하여 신뢰도와 타당도를 검증한 요실금 관련 설문지는 아직까지 없으나 BFLUTS, KHQ 등 (17) 몇가지 설문지들에 대하여 이러한 검증 과정이 진행되었다. 본 연구는 이전에 한국어로 번역한 한글판 I-QOL 설문도구가 복잡성요실금을 가진 한국인 인구집단에서도 유용하게 쓰일 수 있는지를 검증하는 연구이다. 본 연구 결과 본 한글판 I-QOL 설문도구는 타당도와 신뢰도가 있다고 판단되고 또한 본 번역 설문지는 번역후 타당도와 신뢰도를 검증한 BFLUTS (unpublished data), KHQ (17) 등 다른 설문지와 비교하여 보아 대등하거나 상당히 높은 신뢰도와 타당도를 보여주고 있다.

본 연구결과 한글판 I-QOL 설문 도구에 응답한 대부분의 환자들이 질문 내용이 환자들의 배뇨상황과 관련있고 의미있다고 응답하였다. 또한 설문지를 작성하는데 걸린 시간은 5분 정도이며 대부분의 환자들이 질문 내용이 명확하고 이해하기 쉽게 되어 있고 작성하기 쉽게 구성되어 있으며 요실금에 대한 질문 내용이 자신들의 상황에 적절하다고 응답하여 실제 외래 진료 환경에서 진료 전 환자의 증상을 파악하는데 매우 신속하고 기입하기 쉬운 우수한 도구라는 것을 확인할 수 있었다.

본 연구에서는 I-QOL 설문지로 산정한 점수를 요실금 유병기간, 개월당 요실금 횟수, 요실금 증상의 전반적 심한 정도, 패드검사상 요실금의 양 등 다른 주관적 또는 객관적 지표들과 비교하여 보았을 때, 증상의 전반적 심한 정도 지표가 I-QOL 설문 점수와 가장 일치도가 높게 나타났다. 그러나 요실금 유병기간, 개월당 요실금의 횟수 등의 지표에서는 일부 소범주에서만 상관관계가 발견되었으며 패드검사에서는 요실금의 양과는 미약한 상관관계만이 발견되었다. 반응도조사에서도 치료전후 증상점수는 큰 차이가 있으나 증

상의 변화정도를 요실금의 횟수나 패드검사가 정확히 반영을 못하는 것으로 판단되었다.

동일한 질병을 평가하는 지표들 사이에서도 주관적인 지표와 객관적인 지표간의 평가 결과사이에는 일치하지 않는다는 결과는 다른 연구자들의 결과에서도 이미 보고된 바 있다. 아마도 동일한 정도의 객관적인 현상을 환자마다 받아들이는 정도가 달라 개개인마다 주관적으로 인지하는데 차이가 발생하는 것에 기인하지 않는가 한다. 반면, I-QOL점수는 증상의 전반적 심한 정도 지표 등과 같이 주관적인 것을 뭉뚱그려 하나의 정도로 표현한 지표와는 매우 큰 상관관계를 보여 I-QOL 도구는 주관적인 면을 평가하는 도구로서 매우 유용함을 보여주었다고 할 수 있다.

I-QOL 설문 도구는 일반적인 건강 상태를 표현하는 SF-36 설문도구 점수와는 세부영역에서도 잘 일치하지 않았다. 그 이유로는 이 표준설문지는 전반적인 건강상태를 묻는 질문들로 구성이 되어 있다는 점과 한국인들에서는 요실금 자체를 병으로 인식을 하지 않는 문화적 차이 등이 있지 않을까 생각한다.

신뢰성은 일관성을 의미하는 것인데 이 조사는 같은 대상에서 동일한 항목에 대해 측정을 반복했을 때 같은 반응이 나왔는가 여부를 측정하는 것으로서 한 검사에서 얻어진 점수를 얼마만큼 믿을 수 있느냐 하는 정도를 말한다. 이것을 측정하는 방법은 측정한 것을 다시 측정했을 때 동일한 결과가 나오는가 (test-retest), 비슷한 질문을 두 번했을 때 비슷한 답이 나오는가 (internal consistency)를 측정하는 것으로 이루어진다. 2주 전후로 동일한 설문을 주고 실시한 test-retest 안정성조사에서 시기를 두고 동일한 질문에 대한 응답이 일치하지 않는 항목이 일부 발견되었다. 이 이유는 2주라는 기간 사이에 증상의 변화가 발생하였을 가능성이 일부 기여한 것으로 판단된다. 내적일치도의 통계적 측정은 보통 Cronbach 알파 계수를 구하는데 통상 그 값이 0.6이상이면 신뢰성이 있다고 판단한다. 본 연구의 환자군에서 분석 결과 전반적으로 내적일치도가 매우 우수하다고 판단되었다.

본 연구의 반응도 또는 민감도 조사에서 치료전후의 I-QOL 각 영역별 점수는 술 전과 술 후에 의미있는 차이를 보여주었다. 또한 변화의 양을 표시하는 방법으로는 effect size, standardized response mean (14),



Guyatt responsiveness index (16) 등이 동원되는데 Guyatt responsiveness index의 장점으로는 비반응군 (위 약 대조군)의 변이를 고려할 수 있으므로 반응도 평가에 더 우수한 방법으로 인정받고 있다. 이 모두 본 연구에서는 1.1에서 2.9까지를 보여 I-QOL의 각 영역별 점수는 반응도를 구분하는 매우 우수한 결과를 보여 주었다.

본 연구 결과 한글판 I-QOL는 한국 여성의 복압성 요실금에서 매우 타당도가 높으며 신뢰도가 높고 재현성이 있으며 치료의 반응을 잘 반영하는 정신측정학적 특성을 보여주는 우수한 설문도구로 인정된다. 이 외에도 국제적으로 널리 통용되고 있는 좀 더 다양한 배뇨증상 측정 도구들이 지속적으로 한글화 및 타당도 검증과정을 통하여 국제 규격에 맞는 연구 도구들을 확보하고 정비하는 일은 계속되어야 할 것으로 생각한다. 본 한국어판 I-QOL 설문지가 임상진료에 유용하게 쓰이기를 기대하며 이 도구가 국제적 임상연구에 있어서 귀중한 밑거름이 되기를 기대한다.

## 결론

한글판 I-QOL는 한국 여성의 복압성요실금에서 대조군과 환자군은 구별에 유용하여 타당도가 매우 높고 또한 재현성이 있으며 신뢰도가 매우 높고 치료의 반응을 잘 반영하는 정신측정학적 특성을 보여주는 우수한 설문도구라고 판단된다.

## ACKNOWLEDGEMENT

서울대학교병원의 임승화, 송재은, 송지영 간호사 그리고 김귀식씨가 본 연구의 자료 전산화에 많은 협조를 하였다. 본 연구의 결과는 Eli Lilly사 (Indianapolis, IN, U.S.A.)에 통보되었으며 한글판 설문지 I-QOL의 판권은 Eli Lilly사에 있으므로 임상 사용 전에 반드시 사전 허락을 득하여야 한다.

## REFERENCES

1) Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for

women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program in Women (CPW) Research Group. Qual Life Res 1994;3:291-306

2) Sandvik H, Hunskaar S, Seim A, Hermstad R, Vanvik A, Bratt H. Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. J Epidemiol Community Health 1993;47:497-9

3) Black N, Griffiths J, Pope C. Development of a symptom severity index and a symptom impact index for stress incontinence in women. NeuroUrol Urodyn 1996; 15:630-40

4) Kelleher CJ, Cardozo LD, KhullarV, Salvatore S. A new questionnaire to assess the quality of life of urinary incontinent women. Br J Obstet Gynaecol 1997; 104:1374-9

5) Wagner TH, Patrick DL, Bavendam TG, Martin ML, Buesching DP. Quality of life of persons with urinary incontinence: development of a new measure. Urology 1996;47:67-71

6) Wyman JF, Harkins SW, Choi SC, Taylor JR, Fantl JA. Psychosocial impact of urinary incontinence in women. Obstet Gynecol 1987;70:378-81

7) Patrick DL, Martin ML, Bushnell DM, Yalcin I, Wagner TH, Buesching DP. Quality of life of women with urinary incontinence: further development of the incontinence quality of life instrument (I-QOL) Urology 1999;53:71-6

8) Bosch JDR, Jackson MGS, Naughton M, Valiquette SRL, Batista JE, Avery K. Symptom and quality of life assessment. In: Abrams P, Cardozo L, Khoury S, Wein AJ, editors. Incontinence. Paris: Health Publication Ltd 2005;519-84

9) 오승준, 박형근, 임승화, 홍성규, Martin ML, Ting BL, et al. 요실금자가기입설문지 한국어판 Incontinence Quality of Life (I-QOL)의 언어 타당도. 대한배뇨장애 및 요실금학회지 2002;6:10-23

10) Patrick DL, Martin M, Bushnell DM. The I-QOL, user's manual and scoring diskette. Health Research Associates, Inc. Seattle, WA, U.S.A.

11) Paick JS, Ku JH, Shin JW, Park K, Son H, Oh SJ, et al. Significance of pad test loss for the evaluation of women with urinary incontinence. NeuroUrol Urodyn 2005;24:39-43

12) Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P,

- Ulmsten U, et al. The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn* 2002;21:167-78
- 13) Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A, Mowat A. Importance of sensitivity to change as a criterion for selecting health status measures. *Qual Health Care* 1992;1:89-93
- 14) Gwady-Sridhar F, Guyatt GH, Arnold JM, Massel D, Brown J, Nadeau L, et al. Instruments to measure acceptability of information and acquisition of knowledge in patients with heart failure. *Eur J Heart Fail* 2003; 5:783-91
- 15) Pengel LH, Refshauge KM, Maher CG. Responsiveness of pain, disability, and physical impairment outcomes in patients with low back pain. *Spine* 2004;29:879-83
- 16) Wanders AJ, Gorman JD, Davis JC, Landewe RB, van der Heijde DM. Responsiveness and discriminative capacity of the assessments in ankylosing spondylitis disease-controlling antirheumatic therapy core set and other outcome measures in a trial of etanercept in ankylosing spondylitis. *Arthritis Rheum* 2004;51:1-8
- 17) 오승준, 주명수, 김홍식, 김준철, 이정구, 윤종민 외. 여성복압성요실금 환자에서 한국어판 King's Health Questionnaire 설문지의 타당도 및 신뢰도에 관한 정신 측정학적 특성. *대한배뇨장애 및 요실금학회지* 2005;9: 115-23