

EQ-5D 건강상태에 대한 가치 평가 작업의 타당도와 신뢰도 평가: 컴퓨터 보조프로그램을 이용한 표준도박

울산대학교 의과대학 예방의학교실¹, 명지대학교 공과대학 컴퓨터공학과²

조민우¹, 권동섭², 김선하¹, 길선령¹, 이상일^{*1}

Validity and Reliability of Korean EQ-5D Valuation Study Using a Computer-assisted Standard Gamble Method

Min-Woo Jo¹, Dong-Seop Kwon², Seon-Ha Kim¹, Seol-Ryoung Kil¹, Sang-Il Lee^{*1}

¹Department of Preventive Medicine, University of Ulsan College of Medicine

²Department of Computer Software, Myongji University College of Engineering

Background: The Standard Gamble (SG) method is the gold standard for valuing health states as a utility. The objective of this study is to examine the validity and reliability of EQ-5D valuation task using a computer-assisted SG method in the general population of South Korea.

Methods: An internet-based computer program was developed to allow respondents to conveniently conduct a valuation of EQ-5D health states with SG. For the first survey, 29 persons were recruited randomly from the general population in Gyeongju city. Of these, 22 participated in the second survey. Direct valuation was elicited for 43 health states (42 EQ-5D health states and 'unconscious' state) during face-to-face interviews through an internet program with a visual analogue scale (VAS), time trade-off (TTO), and SG method. Validity was assessed using the number of respondents with 4 or more inconsistencies and the relationship between VAS, TTO, and SG values. Intra-class correlation coefficient (ICC) was calculated to evaluate the test-retest reliability.

Results: Of the 29 respondents in the first survey, 5 had 4 or more inconsistent responses compared with 2 in VAS and 9 in TTO. Pearson's correlation coefficient of SG-VAS and SG-TTO were 0.867 and 0.894, respectively ($P<0.001$). The relational formulas were $SG=1-(1-VAS)^{0.962}$ ($R^2=0.926$) and $SG=1-(1-TTO)^{0.932}$ ($R^2=0.868$). The medians of individual ICC of VAS, TTO, and SG were 0.919, 0.721, and 0.860, respectively.

Conclusions: We found that the internet-assisted SG method can be used as a valid and reliable tool for valuing EQ-5D health states in the general population of South Korea.

Korean J Health Promot 2010;10(3):105-112

Key Words: Validity, Reliability, EQ-5D, Standard gamble method

서 론

보건 의료비용이 급속히 증가함에 따라 보건 의료 중재의 경제적 효율성을 평가하는 경제성 평가에 대한 관심이 점

점 증가하고 있다. 보건 의료 분야에서 경제성 평가에는 비용-효과 분석(cost-effectiveness analysis, CEA), 비용-효용 분석(cost-utility analysis, CUA), 비용-편익 분석(cost-benefit analysis, CBA) 등이 있는데, 이 중 비용-효용 분석은 질보정수명(Quality-adjusted life years, QALYs) 등 환자의 선호도를 반영하면서도 공통으로 사용할 수 있는 지표를 결과 지표로 이용한다. QALYs는 삶의 양적인 측면을 나타내는 생존기간에 질적인 수준을 함께 반영한 지표이며 이 지표를 이용하여 비용-효용 분석을 수행하게 되면, 여러 질환이나 중재간의 비용-효용비를 비교할 수 있기 때문

■ Received : May 24, 2010 ■ Accepted : August 2, 2010

* Corresponding author : Sang-Il Lee

86, Asanbyeongwon-gil, Songpa-gu, Seoul 138-736, Korea

Tel: 82-2-3010-4284 Fax: 82-2-477-2898

E-mail: sleemd@amc.seoul.kr

■ This study was supported by a grant (2006-063) from the Asan Institute for Life Sciences, Seoul, Korea

에 유용하다는 장점이 있다.¹⁾

QALYs를 구하기 위해서는 건강 관련 삶의 질(Health related quality of life, HRQOL) 수준을 반영해 주는 선호도인 질 가중치(quality weight) 혹은 효용 가중치(utility weight)가 필요하다. 질 가중치를 구하는 방법은 크게 두 가지가 있는데, 하나는 직접법이고, 다른 하나는 간접법이다. 직접법은 어떤 건강상태에 대해 시각 아날로그 척도법(visual analogue scale, VAS), 시간교환법(time trade-off, TTO), 표준도박법(standard gamble, SG) 등의 가치평가 방법을 직접 적용하여 가중치를 구하는 방법이다. 간접법은 어떤 건강상태 분류 체계에 가치 평가 방법을 적용하여 각 상태별 가중치를 미리 구하고, 이를 개인이 평가한 체계별 상태에 적용하여 질 가중치를 부여하게 된다.^{2,3)}

EQ-5D는 EuroQOL 그룹이 개발한 HRQOL 평가도구로 질 가중치를 구하는 데 흔히 적용하는 건강상태 분류체계 중 하나이다.⁴⁾ EQ-5D는 운동능력, 자기관리, 일상 활동, 통증/불편감, 불안/우울의 5개 차원을 ‘문제 없음’, ‘다소 문제 있음’, 그리고 ‘심각한 문제가 있음’의 세 가지 수준으로 평가하도록 하고 있다. 매우 간단하게 HRQOL 수준을 평가할 수 있기 때문에 우리나라를 포함하여 많은 나라에서 일반인구집단이나 특정 질환을 가진 환자들을 대상으로 HRQOL을 구하기 위한 연구에 많이 이용하였다. 특히 질 가중치를 구하는 가치 평가 연구도 많이 수행하였다. 하지만 이 연구들의 대부분은 가치 평가 방법으로 VAS와 TTO를 이용하였다.⁵⁻⁷⁾ 대개 비용-효용 분석 수행을 위해 QALYs를 구할 때 이론적으로는 SG를 이용하여야 하지만 수행 과정에서 평가자인 보통 사람들이 확률에 대해 이해하기 어렵다는 문제가 있기 때문에 TTO를 이용하여 구한 값을 대안으로 이용하고 있다.⁸⁾ 하지만 TTO를 이용하여 구한 선호도는 가치(value)로 효용(utility)의 대리값일 뿐 효용이 아니다. 최근 몇몇의 연구에서 시각적 도움을 제공하는 컴퓨터 프로그램을 보조적으로 사용하여 SG를 이용한 가치평가 작업을 수행한 바 있고, 그 결과 SG를 이용하여 가치평가를 수행하는 작업이 받아들일만 하다고 보고한 바 있다.^{5,6)} 하지만 EQ-5D에서는 아직 SG를 적용할 수 있는지에 대한 연구가 진행된 바 없다.

이 연구의 목적은 한국의 일반집단을 대상으로 컴퓨터 보조도구를 활용하여 SG를 이용한 EQ-5D 건강상태에 대한 가치 평가작업의 타당도 및 신뢰도를 평가하는 것이다.

방 법

1. 연구대상

경주시에 사는 20세 이상의 성인 남녀를 설문 대상으로

하였다. 인명 전화번호부를 이용하여 무작위 추출하였고, 전화 통화를 통해 성별과 연령을 확인한 후 연구에 대한 설명을 한 후 연구 참여에 대한 동의를 구하였다. 20대부터 60대까지 10세 단위로 남녀 각각 3명씩 포함하도록 하였고, 전화에서 확인한 성별, 연령이 조건에 부합되는 경우 연구에 참여하도록 요청하였다. 연구 참여에 동의를 한 경우, 2명의 훈련받은 면접자가 직접 방문하여 가치 평가 작업을 진행하였다. 29명의 대상자가 1차 면접 조사에 응답하였고, 이중 22명이 2주 후에 시행된 2차 조사에 응답하였다. 이 연구는 2007년 1월부터 2월까지 수행되었다.

2. 가치 평가를 위한 컴퓨터 프로그램 개발

영국 MVH 연구와⁵⁾ 우리나라에서 수행한 가치 평가 연구를⁴⁾ 참조하여 조사자가 수행하는 종이 설문지 기반의 가치평가작업을 인터넷 기반 컴퓨터 프로그램으로 개발하였고, 언어는 한국어로 개발하였다. 컴퓨터 프로그램으로 구현할 때 다른 과정들은 모두 기존의 설문조사 방식과 같이 자가 건강상태 평가 그림 1(a)를 먼저 하고, 평가할 가상의 건강상태를 선택하도록 하였는데 그림 1(b), 이 때 이용한 건강상태는 영국 MVH 연구에서⁵⁾ 사용한 43개 건강상태였다. 1차 조사에서는 이 건강상태들을 무작위로 선택할 수 있도록 하였고, 재조사에서는 같은 건강상태를 선택할 수 있도록 건강상태 카드를 강제로 선택할 수 있게 하였다. 건강상태를 선택한 후 먼저 VAS를 이용하여 가치 평가하도록 하였고 그림 1(c), 이후 선택한 건강상태를 죽음보다 못한 건강상태와 그렇지 않은 건강상태로 나누고 선호도 순서로 순위를 매긴 후 그 순서에 따라 TTO를 이용하여 가치 평가 작업을 수행하도록 하였다 그림 1(d). SG에서도 TTO에서도 같이 선택한 가상상태의 가치평가를 수행하기 전에 죽음보다 나은 건강상태와 죽음보다 못한 건강상태로 나눈 후 선호도 순서로 순위를 매긴 후에 그 순서대로 가치평가를 수행하도록 하였다(그림 2).

3. 가치평가

먼저 모든 응답자는 자신의 건강상태를 EQ-5D로 평가하였다. 두 번째로 15개의 가상 EQ-5D 상태 즉, 중등도에 따라 ‘아주 경한 상태’ 2개의 건강상태를, ‘경한 상태’, ‘중등도 상태’ 및 ‘중한 상태’에서는 각각 3개의 건강상태를 무작위로 선택된 11개의 상태와 ‘11111’, ‘33333’, ‘무의식’, ‘사망’ 상태에 대해 VAS를 사용하여 0에서 100점까지 평가를 수행하였다. 세 번째로는 ‘11111’과 ‘사망’을 제외한 13개의 상태에 대해 ‘죽는 것보다 나은 건강상태’와 ‘죽는 것보다 못한 건강상태’로 선택하게 한 후 TTO를 이용하여

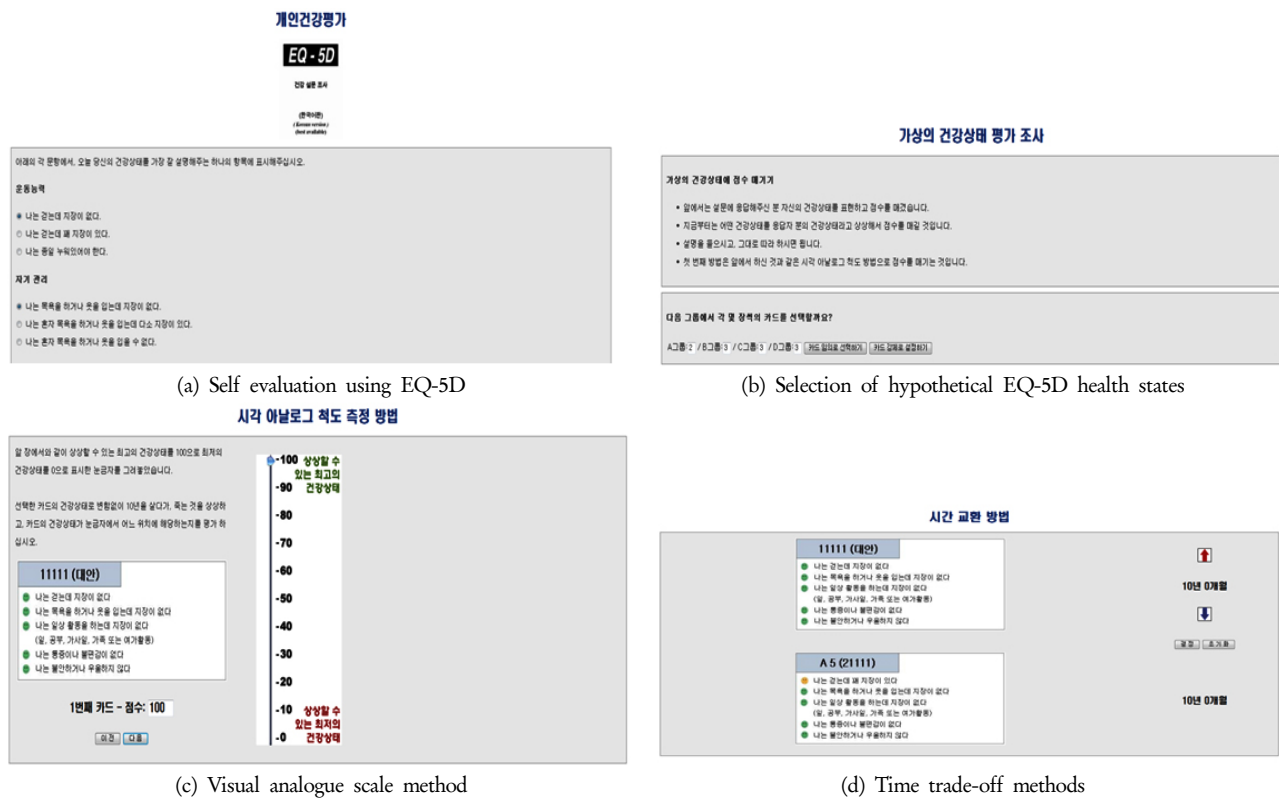


Figure 1. Computer program for valuation task: self evaluation, visual analogue scale, and time trade-off.

Table 1. Example of standard gamble method: valuation process for '12221' state

		Choice A: 11111 for 10 yr													
		100%	0%	90%	10%	80%	89%	81%	88%	82%	87%	87.90%	87.10%	87.80%	
Choice B: 12221 for 10 yr	A	A			A			A			A				Indifference point
		B			B	B	B			B	B	B			

가치를 평가하였다. ‘죽는 것보다 나은 건강상태’에서는 선택한 건강상태로 10년을 사는 것과 EQ-5D로 표현할 때 완전한 건강상태인 ‘11111’의 건강상태로 살지만 10년보다 짧은 t 년을 사는 것 중 하나를 선택하게 되는데 두 대안의 선호도에 차이가 없는 무차별점(indifference point)에서의 t 값을 구했다. 그리고 ‘죽는 것보다 못한 건강상태’라고 분류한 그 건강상태로(10- t)년을 살다가 나머지 t 년을 ‘11111’의 건강상태로 사는 선택에서 무차별점 t 값을 구했다. 무차별점을 구하기 위해 t 의 초기값을 10년을 먼저 제시하였고, 이를 수용하면 0년을 사는 것을 제시하였으며, 0년을 수용하지 않으면, 다시 9년, 1년, 8년 순으로 무차별점에 접근하는 이른바 ‘핑퐁’ 방식을 이용하여 구하도록 하였다. 만일 연단위에서 가려지지 않는 부분이 있다면 월 단위 수준에서 평가하도록 하였다. 마지막으로 표준도박법은 TTO와 같이 13개 가상 건강상태를 ‘죽는 것보다 나은 건강상태’와 ‘죽는 것보다 못한 건강상태’로 구분하게

한 후 시행하였다. ‘죽는 것보다 나은 건강상태’에서는 이 상태로 10년을 산다는 것을 가정하게 하고, 약을 먹으면 완전한 건강을 되찾을 확률도 있지만, 즉시 사망할 확률도 p_1 인 상황에서, 확률 p_1 를 변화시켜 가며 약을 복용할 것인가를 결정하도록 하였다. ‘죽는 것보다 못한 건강상태’에서는 약을 복용하지 않으면 사망하게 되는데, 복용하게 되면 완전한 건강을 되찾을 가능성도 있지만, 해당 상태 즉, ‘죽는 것보다 못한 건강상태’로 계속 살 가능성(p_2)도 있다. 이러한 경우 p_2 를 변화시켜 가면서 무차별점이 되는 확률 p_2 를 선택하도록 하였다. SG에서 무차별점에 접근하는 방식도 시간교환법과 같은 방식으로 하였는데, 초기에 완전한 건강을 찾을 확률로 100%를 제시하고 수용하면 0%를 제시하도록 하였다. 이후 90%, 10% 순으로 변화시키면서 선택하도록 하였고, 10% 수준의 변화는 1% 단위로, 1% 수준의 변화에서는 0.1% 단위로 확률을 변화시켜 가면서 무차별점을 구하도록 하였다. 예를 들어 ‘12221 (B)’ 상태에

표준 도박 방법

* 10년을 사는 것을 상상한 커드입니다.
* 이 중에서 그 상태에 10년을 사거나 죽는 것이 낫다고 생각되는 상태를 선택하여 주십시오.

번호	카드	죽는 것에 대한 공포
1	A02(1111)	□
2	A03(1112)	□
3	B04(2113)	□
4	B03(1131)	□
5	B06(2121)	□
6	C05(2222)	□
7	C03(3231)	□
8	C06(2132)	□
9	D05(2223)	□
10	D07(3223)	□
11	D10(3321)	□
12	D10(3333)	□

(a) Selection of worse than death states

표준 도박 방법

* 사는 것이 더 나은 상태라고 선택한 커드입니다.
* 이 커드도 좋은 상태부터 나쁜 상태로 정렬하십시오.

번호	카드	정렬 순서
1	A02(1111)	1
2	A03(1112)	2
3	B04(2113)	3
4	B03(1131)	4
5	B06(2121)	5
6	C05(2222)	6
7	C03(3231)	7
8	C06(2132)	8
9	D05(2223)	9
10	D07(3223)	10
11	D10(3321)	11
12	D10(3333)	12

(b) Ranking health states by preference

표준 도박 방법

A5 (21111)

● 나는 관능에 별 지장이 있다
● 나는 목숨을 잃거나 혼을 잃는다는 지장이 없다
● 나는 일상 생활을 하는데 지장이 없다
(일, 공부, 가사, 가족 또는 여가활동)
● 나는 통증이나 불편감이 없다
● 나는 불안하거나 우울하지 않다

약을 먹으면 완전한 건강을 찾을 수도 있지만,
약을 먹으면 죽음 확률이 0% 미만 확률로 낮아집니다?

0% 100%

(c) Standard gamble method: valuation start point

표준 도박 방법

A5 (21111)

● 나는 관능에 별 지장이 있다
● 나는 목숨을 잃거나 혼을 잃는다는 지장이 없다
● 나는 일상 생활을 하는데 지장이 없다
(일, 공부, 가사, 가족 또는 여가활동)
● 나는 통증이나 불편감이 없다
● 나는 불안하거나 우울하지 않다

약을 먹으면 완전한 건강을 찾을 수도 있지만,
약을 먹으면 죽음 확률이 10% 미만 확률로 낮아집니다?

0% 100%

(d) Standard gamble method: valuation process

Figure 2. Computer program for valuation task: standard gamble.

대해 표준도박법으로 가치평가 작업을 수행한 C씨의 사례를 살펴보자(표 1). B상태인 C씨가 어떤 중재를 받았을 때, '11111 (A)'로 10년을 살 확률이 100%일 때 C씨는 HRQOL 수준이 높은 상태로 같은 기간을 사는 경우(A)를 선택하였다. A상태로 살 확률이 0% 즉, 바로 사망하게 되는 경우에는, 다소간 HRQOL 수준이 낮지만 10년간 더 살게 되는 B를 선택하였다. 90%의 확률에서는 A를 선택했고, 10%에서는 B를 선택하였으며, 80%에서는 B를 선택하였다. 이러한 방식으로 진행한 결과, C씨는 A와 B 사이의 무차별점으로 87.8%를 선택하게 되었다.

4. 시각 아날로그 척도법, 시간교환법, 표준도박법의 점수 변환

어떤 상태 X에 대해서 VAS로 측정한 값(VAS_x)은 사망에 대한 VAS (VAS_{death})와 EQ-5D로 표현한 완전한 건강상태인 '11111' 상태의 VAS (VAS₁₁₁₁₁) 사이의 범위에서 해당 건강상태에 대한 값의 위치(VAS_{Tx})가 어디인가로 살펴보기 위해서 다음 식을 이용하여 변환하였다.

$$VAS_{Tx} = \frac{VAS_x - VAS_{death}}{VAS_{11111} - VAS_{death}}$$

VAS_{Tx}: 상태 x의 변환 후 VAS값

VAS₁₁₁₁₁: 상태 '11111'에 대해 VAS로 평가한 값

VAS_x: 상태 x에 대해 VAS로 가치 평가한 값

VAS_{death}: 죽음에 대해 VAS로 가치 평가한 값

X상태에 대한 TTO값 (TTO_x)은 최고의 건강상태에서는 1, 죽음은 0, 죽는 것보다 못한 상태는 음수인데 -1까지 값을 갖도록 다음의 두 식을 이용하여 변환하였다.

$$TTO_{Tx} = \frac{TTO_x}{10} \quad \dots \text{'죽는 것보다 나은 건강상태'}$$

$$TTO_{Tx} = \frac{(10 - TTO_x)}{10} - 1 = -\frac{TTO_x}{10} \quad \dots \text{'죽는 것보다 못한 건강상태'}$$

TTO_{Tx}: 상태 x의 변환 후 TTO 값

TTO_x: 상태 x에 대해 TTO로 가치 평가한 값

SG로 측정한 값은 '죽는 것보다 나은 상태'는 1-p₁로, '죽는 것보다 못한 상태'는 -p₂로 변환하여 질 가중치로 사용하였다.

5. 타당도와 신뢰도

SG로 가치 평가하는 작업에 대한 타당도 평가를 위해 수렴 타당도(convergent validity)를 평가하였는데, 이 연구에서는 SG를 이용하여 구한 EQ-5D 건강상태에 대한 질 가중치를 같은 상태에 대한 응답자의 VAS, TTO 값의 상관관계를 피어슨 상관계수로 평가하였고, 관련성을 회귀식으로 살펴보았다. 다른 타당도의 지표로서 자료의 질을 살펴보았다. EQ-5D로 표현된 건강상태 중 논리적으로 비

교할 수 있는 건강상태가 존재한다.⁹⁾ 예를 들어 ‘11112’는 ‘11122’보다 좋은 건강상태이므로, ‘11112’의 질 가중치 값은 ‘11122’의 값보다 같거나 커야 한다. 그 반대의 값을 가진다면 이것은 모순되는 것이다. Ohinmaa and Shntonen 이¹⁰⁾ 제시한 대로 4개 이상의 모순값을 가진 응답자의 수를 평가하였다.

그리고 두 번의 조사에 모두 응답한 사람을 대상으로, 개인별 급내상관계수(intra-class correlation coefficients, ICC)를 구하여 검사-재검사 신뢰도를 평가하였다. ICC를 구하기 위하여 다음의 공식을 이용하였다.

$$ICC = \frac{\sigma_{jw}}{\sigma_{jw} + \sigma_{jb}}$$

σ_{jw} : j번째 응답자에서 실 평가 도출 작업내 변이

σ_{jb} : j번째 응답자에서 실 평가 도출 작업간 변이

6. 소요 시간 및 난이도 평가

소요 시간은 EQ-5D를 이용한 자신의 HRQOL 수준 평가와 더불어 VAS, TTO, 그리고 SG를 이용한 건강상태별 가치평가 작업에 걸린 전체 시간을 살펴보았다. 더불어 각각의 과정에 대한 수행 난이도를 평가하였는데, 모든 과정을 다 종료하고 난 후 자가평가와 VAS, TTO, 그리고 SG로 평가한 가치 평가 작업에 대해서 ‘매우 쉬웠음’부터 ‘매우 어려웠음’까지 리커트 5점 척도로 평가하도록 요청하였다.

7. 통계분석

집단별 특성 비교에는 카이제곱검정을 이용하였고, 수렴타당도를 살펴보기 위해서는 상관분석을 수행하였다. 모든 통계분석은 SPSS 12.0을 이용하여 분석하였고, 유의

확률이 0.05 미만인 경우를 통계적으로 유의하다고 판단하였다.

결 과

1. 컴퓨터 프로그램 개발

가치 평가를 위해 종이 설문지를 이용한 방식과 같은 방식으로 인터넷 기반의 컴퓨터 프로그램을 개발하였고, SG의 경우 이해를 돕기 위해 시각적 보조도구를 추가로 개발하여 가치 평가 작업에 적용하였다(그림 1, <http://www.hrql.org/>).

2. 연구 참여자

시도된 전화 147건 중에 42명은 전화를 받지 않았고, 76명은 연구에 참여를 거부했다. 거부 사유로는 다른 일로 바빠다거나, 설문 조사에 대한 반감 등이 있었다. 결과적으로 총 29명이 1차 조사에 응답하였고, 2차 조사는 이중 7명이 참여하지 않았다. 1차 조사와 2차 조사의 응답자의 일반적 특성을 제시하였다(표 2). 연구 참여자의 교육수준은 대부분이 중학교 이상이었고, 조사간 응답자의 특성을 비교한 결과 통계적으로 유의한 차이를 보인 변수는 없었다.

3. 타당도와 신뢰도

1) 수렴 타당도

본 조사 자료를 이용하여 VAS로 측정된 상태의 평균값과 TTO로 측정된 상태의 평균값, 그리고 SG로 측정된 상태의 평균값을 각각 산점도로 살펴 상관관계를 살펴보았다(그림 3). TTO-SG, VAS-SG 간 피어슨 상관계수는 각각 0.894, 0.867로 모두 통계적으로 유의하였다(각각 $P < 0.001$). VAS와 TTO를 각각 (1-VAS)와 (1-TTO)로 한 뒤 이를

Table 2. General characteristics of study participants in 1st and 2nd surveys

General characteristics		1st survey N (%)	2nd survey N (%)	P-value [*]
Gender	Male	15 (51.7)	11 (50.0)	1.000
	Female	14 (48.3)	11 (50.0)	
Age-group	20~29	6 (20.7)	4 (18.2)	0.966
	30~39	6 (20.7)	5 (22.7)	
	40~49	5 (17.2)	5 (22.7)	
	50~59	6 (20.7)	3 (13.6)	
	60~69	6 (20.7)	5 (22.7)	
Education	Elementary school or less	2 (6.9)	2 (9.1)	0.762
	Middle or high school	12 (41.4)	11 (50.0)	
	University or above	15 (51.7)	9 (40.9)	

^{*} χ^2 test.

로그변환하여 SG와의 회귀모형을 개발한 결과 $SG=1-(1-VAS)^{0.962}$ ($R^2=0.926$), $SG=1-(1-TTO)^{0.932}$ ($R^2=0.868$)의 관계식을 보였다.

2) 측정방법별 모순값

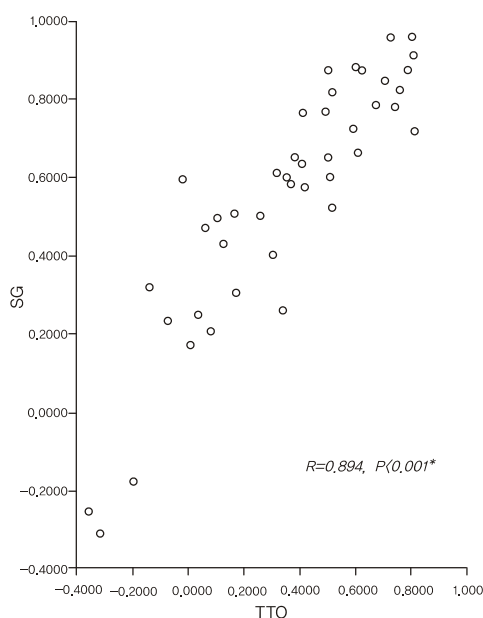
측정방법별로 모순값 개수를 조사한 결과, 1차 조사에서 모순값 개수가 4개 이상인 경우는 TTO 방법에서 9건 (31.0%), SG 방법에서 5건 (17.2%)으로 나타났고, 2차 조사에서는 그 건수가 다소 감소하였다(표 3).

3) 신뢰도

신뢰도 측정을 위해 조사-재조사한 자료에서 개인별로 ICC를 계산한 결과, VAS, TTO, 그리고 SG 방법에서 각각 평균(표준편차)이 0.919 (0.039), 0.701 (0.174), 0.805 (0.123)로 나타났다.

4. 소요 시간과 난이도 평가

전체 작업에 소요된 시간은 1차 조사에서는 평균 46분이 소요되었고, 2차 조사에서는 10분 정도 감소한 36분 가량이 소요되었다. 1차 조사에서 최소 소요 시간은 24분이 소요된 것으로 나타났고, 최장 1시간 11분 가량이 소요되었으며, 2차 조사에서는 최소 16분에서 최장 1시간 4분 가량이 소요된 것으로 나타났다. 그리고 가치평가 방법에 따른 난이도를 살펴본 결과, VAS, TTO, 그리고 SG에 대해 어렵다고 한 경우는 각각 8명(27.6%), 12명(41.4%), 13명(44.8%)이었고 매우 어렵다고 답한 경우는 각각 1명(3.4%), 3명(10.3%) 그리고 4명(13.8%)이었다(표 4).



*Pearson correlation test.

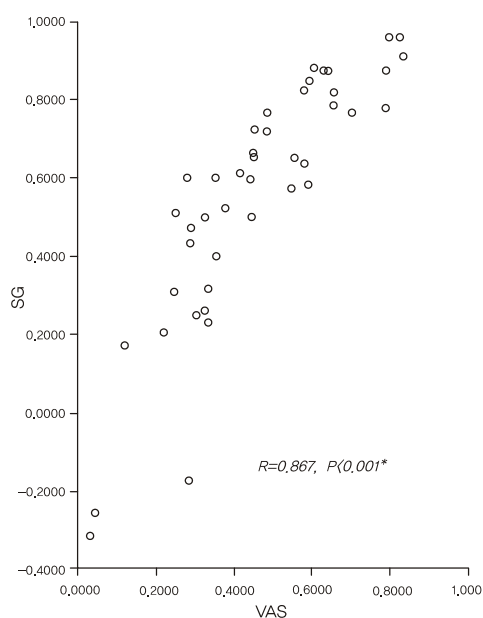


Figure 3. Correlation between time trade-off, visual analogue scale, and standard gamble.

Table 3. Number of respondents with 4 or more in UK MVH study, Korean study, and this study

Studies		The No. of inconsistencies	
		3 or less N (%)	4 or more N (%)
UK MVH study (1997)	Time trade-off	3,074 (91.7)	277 (8.3)
Jo and Lee (2007)	Time trade-off	488 (97.6)	12 (2.4)
1st survey	Visual analogue scale	27 (93.1)	2 (6.9)
	Time trade-off	20 (69.0)	9 (31.0)
2nd survey	Standard gamble	24 (82.8)	5 (17.2)
	Visual analogue scale	22 (100.0)	0 (0.0)
	Time trade-off	18 (81.8)	4 (22.2)
	Standard gamble	21 (95.5)	1 (4.5)

Table 4. Level of difficulty of valuation methods in 1st survey

	Very easy (%)	Easy (%)	Normal (%)	Difficult (%)	Very difficult (%)
Visual analogue scale	2 (6.9)	8 (27.6)	10 (34.5)	8 (27.6)	1 (3.4)
Time trade-off	1 (3.4)	2 (6.9)	11 (37.9)	12 (41.4)	3 (10.3)
Standard gamble	0 (0.0)	3 (10.3)	9 (31.3)	13 (44.8)	4 (13.8)

고 찰

이 연구는 우리나라에서 일반 인구집단을 대상으로 SG를 이용하여 건강상태별 가치 평가 작업을 수행한 탐색적 연구이다. 시각적 보조도구를 인터넷 프로그램의 형태로 개발하였고, 가치 평가 작업에 활용할 수 있도록 하였다. 가치 평가 결과를 타당도와 신뢰도 측면에서 살펴본 결과, SG를 이용한 가치 평가 작업은 타당하게 진행되었음을 확인할 수 있었고, 그 신뢰도도 비교적 높은 수준으로 나타났다.

이 연구는 가치 평가 작업을 통해 EQ-5D 전체 건강상태의 질 가중치를 구하는 것이 목적이 아니라 우리나라에서 EQ-5D 건강상태에 대해 SG를 이용하여 가치 평가작업을 수행할 수 있겠는가를 탐색적으로 살펴보기 위하여 타당도와 신뢰도를 평가할 목적으로 수행된 연구이다. 따라서 가치 평가 작업의 연구대상을 우리나라 전체 인구집단을 대표할 표본을 구성하지 않았고 일개 시지역에서 연령과 성별로 나누어 일정한 수를 포함시키도록 하였다. 그러므로 우리나라 전체 인구집단을 대표하지 못한다는 측면에서는 연구의 제한점이 있다고 할 수 있다.

가치평가 결과값을 타당도와 신뢰도 측면에서 평가하였다. 이 연구에서는 타당도와 관련하여 먼저 수렴 타당도를 보기 위해 VAS, TTO, SG를 같은 건강상태에 대해 평가하도록 한 후, 질 가중치를 변환한 상태에서 상관분석을 시행하였다. SG와 TTO는 0.894, SG와 VAS는 0.867로 높은 상관관계를 보였다. 우리나라에서 TTO로 수행했던 기존 연구에서의 수렴 타당도를 살펴보았던 피어슨 상관계수는 0.934였고, Badia 등¹¹⁾이 수행한 연구가 0.94 수준이었는데 그와 거의 유사한 값을 보였다. SG가 수렴타당도 측면에서 비교적 타당하게 수행된 것으로 판단할 수 있었다. 그리고 SG를 TTO와 VAS의 관계에 따른 식을 구성해본 결과, 지수로 표현된 회귀계수가 우리나라에서 기존에 수행한 연구 결과¹²⁾와 비슷하게, 그리고 외국의 연구 결과¹¹⁾보다는 높은 수준으로 나타났다. 아직은 관련 연구결과가 많지 않기 때문에 우리나라에서의 SG, TTO, 그리고 VAS의 관련식이 외국의 결과와 다른 이유에 대해서 살펴보는 데 제한이 있지만, 최소한 VAS로 구한 값을 이용하여 CUA를 수행하고자 하는 경우, 효용값으로 적절한 변환이 필요한데 외국의 자료를 그대로 쓰는 것은 문제가 있음을 제시해 준다고 할 수 있을 것이다.

모순값의 측면에서 살펴보았을 때, SG에서 모순값의 비율이 1, 2차 조사에서 17.2%, 4.5%를 보였다. Dolan 등¹³⁾은 SG 방법에서 일관성의 역치 수준을 85% 정도는 받아들일만 하다고 주장하였는데, 그에 비추어 볼 때 이 연구 결과는 비교적 높은 수준의 일관성을 보여준다고 할 수 있을 것이다. 그리고 신뢰도 평가에서 SG 방법은 TTO보다 높은 ICC를 나타냈다. SG는 질 평가를 위한 황금기준으로 여겨졌으나 TTO에 비해 실행의 어려움으로 인해 질 가중치 계산을 위해 덜 사용되어 왔다. 그러나 해외의 다른 연구에서¹⁴⁾ TTO의 신뢰도가 SG보다 높았던 것과는 달리 본 연구결과 SG가 건강상태 가치평가에 있어 TTO에 비해 더 높은 타당도와 신뢰도를 보여주었고, 수행난이도와 관련하여 SG에서 어렵거나 매우 어렵다는 응답자가 가장 많았으나 TTO와 거의 유사함을 확인할 수 있었다. 이러한 요인으로는 인터넷 기반으로 한 시각적인 효과를 주는 보조도구가 확률에 대한 의미를 평가자에게 잘 전달하는 데 도움을 준 것이 한 가지 이유가 될 수 있을 것이다.

몇 가지 제한점이 있음에도 불구하고, 이 연구는 여러 연구에서 많이 사용하고 있는 건강 관련 삶의 질 평가도구인 EQ-5D에 대해 인터넷 보조도구를 활용하여 SG를 적용한 질 가중치를 구하여 VAS와 TTO로 구한 결과와 비교해 봄으로써 SG 방법의 타당도와 신뢰도를 살펴본 첫 번째 연구라는 측면에서 의의가 있다. 그 결과, SG를 EQ-5D 가치 평가 작업에 적용하는 것은 가능한 것으로 판단할 수 있었다. 더불어 IT기술이 발달한 한국에서 무선 인터넷을 통해 컴퓨터를 설문조사에 보조적으로 활용하는 방법은 SG를 이용한 가치평가에 유용하게 활용할 수 있는 방법이라고 판단할 수 있었다. 이러한 방식이 종이 설문지를 활용한 후 입력작업을 거치면서 발생하는 자료관리의 오류를 줄일 수 있다는 이점도 존재할 것이다.

요 약

연구배경: 건강상태별 효용 수준을 살펴보기 위해서는 표준도박법을 적용하여 건강상태에 대한 가치 평가 작업을 수행해야 하나 우리나라에서는 아직 표준도박법을 이용하여 가치 평가 작업을 수행한 바 없다. 이 연구는 우리나라 일반 인구집단을 대상으로 EQ-5D 건강상태에 대해 인터넷 보조도구를 활용하여 표준도박법으로 가치 평가하는 작업의 타당도와 신뢰도를 평가하기 위하여 수행하였다.

방법: 먼저 EQ-5D 건강상태에 대해서 표준도박법으로 가치 평가하기 위하여 인터넷 보조프로그램을 개발한 후, 가치 평가 작업을 수행하였다. 가치 평가 작업을 수행할 연구참여자는 경주시 인명전화번호부에서 무작위로 선택하여 20대부터 60대 이상까지 10세 단위로 남녀별로 3명씩 모집하였다. 43개 건강상태에 대해 시각 아날로그 척도법, 시간교환법, 그리고 표준도박법을 이용하여 가치 평가 작업을 수행하였다. 검사-재검사 신뢰도를 살펴보기 위하여 첫 번째 조사 4주 후에 두 번째 조사를 진행하였고, 두 번째 조사에서도 같은 방법으로 연구를 진행하였다. 타당도는 각 가치 평가 방법에 따른 모순값의 수를 살펴보고, 더불어 각 방법으로 구한 가중치 간 상관관계를 살펴보았다. 그리고 신뢰도는 급내상관계수를 이용하여 살펴보았다.

결과: 가치평가작업을 위해 인터넷 보조프로그램을 개발하였다. 그리고 1차 조사에서는 29명의 참가자가 참여를 하였고, 2차 조사에는 이 중 22명의 참가자가 참여하였다. 1차 조사에서 시각 아날로그 척도법, 시간교환법, 그리고 표준도박법에서 4개 이상의 모순값을 보인 참가자는 각각 2명, 9명, 그리고 5명이었다. 더불어 표준도박법과 시각 아날로그 척도법, 표준도박법과 시간교환법 간의 피어슨 상관관계수는 0.867과 0.894였고, 유의확률은 모두 0.001 미만으로 통계적으로 유의하였다. 개인수준에서의 급내상관계수는 시각 아날로그 척도법, 시간교환법, 그리고 표준도박법에서 각각 0.919, 0.701, 그리고 0.805로 나타났다.

결론: 인터넷 보조프로그램을 개발하고 이를 이용하여 표준도박법으로 EQ-5D 건강상태에 대한 가치평가작업을 수행한 결과, 모순값과 수렴타당도 측면에서 살펴본 타당도는 받아들일만하다고 판단되었고, 조사-재조사 방법으로 살펴본 신뢰도는 높은 수준으로 나타나 인터넷 보조도구를 활용하여 표준도박법으로 EQ-5D 건강상태에 대해 가치평가작업을 수행하는 것이 타당하고 신뢰도가 있는 방법으로 나타났다.

중심단어: EQ-5D, 표준도박법, 타당도, 신뢰도, 삶의 질

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Studdart GJ. Methods for the economic evaluation of health care programmes. 3rd ed. New York;Oxford University Press; 2005. p. 8-25.
2. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;5(1):1-30.
3. Froberg DG and Kane RL. Methodology for measuring health-state preferences-II: scaling methods. *J Clin Epidemiol* 1989;42(5):459-71.
4. EuroQol Group. EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;16(3):199-208.
5. Dolan P. Modeling valuation for EuroQol health states. *Med Care* 1997;35(11):1095-108.
6. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43(3):203-20.
7. Jo MW, Yun SC, Lee SI. Estimating quality weights for EQ-5D health states with the time trade-off method in South Korea. *Value Health* 2008;11(7):1186-9.
8. Torrance GW, Thomas WH, Sackett DL. A utilities maximization model for evaluation of health care programs. *Health Serv Res* 1972;7(2):118-33.
9. Brooks R, Rabin R, Charro F, eds. The measurement and valuation of health status using EQ-5d: a European perspective. Dordrecht: Kluwer Academic Publisher;2003. p.143-55.
10. Ohinmaa A, Shntonen H. Inconsistencies and modelling of the Finnish EuroQol(EQ-5D) preference value. *EuroQol plenary meeting discussion papers*; 1-2 October 1998; Hannover, Germany. Hannover:Uni-Verlag Witte;1999. p.57-74.
11. Badia X, Monserrat S, Roset M, Herdman M. Feasibility, validity, and test-retest reliability of scaling methods for health states: the visual analogue scale and the time trade-off. *Qual Life Res* 1999;8(4):303-10.
12. Jo MW, Lee SI. Validity and reliability of Korean EQ-5D valuation study using the time-trade off method. *Korean J Health Promot Dis Prev* 2007;7(2):96-103.
13. Dolan P, Gudex C, Kind P, Williams A. Valuing health statues: a comparison of methods. *J Health Econ* 1996;15(2):209-31.
14. Stewart ST, Lenert L, Bhatnagar V, Kaplan RM. Utilities for prostate cancer health states in men aged 60 and older. *Med Care* 2005;43(4): 347-55.