

# The Construct and Concurrent Validity of the EQ-5D in a Renal Transplant Population

Irina Cleemput, PhD,<sup>1</sup> Katrien Kesteloot, PhD,<sup>1,2</sup> Philip Moons, MSc,<sup>1,2</sup> Yves Vanrenterghem, MD, PhD,<sup>2</sup> Johannes P. Van Hooff, MD, PhD,<sup>3</sup> Jean-Paul Squifflet, MD, PhD,<sup>4</sup> Sabina De Geest, PhD<sup>1</sup>

<sup>1</sup>Centre for Health Services and Nursing Research, Katholieke Universiteit Leuven, Leuven, Belgium; <sup>2</sup>Department of Nephrology, University Hospitals Leuven, Leuven, Belgium; <sup>3</sup>Department of Nephrology, Academic Hospital Maastricht, Maastricht, the Netherlands; <sup>4</sup>Department of Kidney and Pancreas Transplantation, Cliniques Universitaires Saint-Luc, Brussels, Belgium

## ABSTRACT

**Objectives:** Cost–utility analysis in renal transplant populations requires the use of a generic instrument for health status measurement that generates a single value for health. Such instruments should be widely applicable in diverse patient populations and their validity should be established. The aim of this study was to explore the validity of the EQ-5D in renal transplant patients.

**Methods:** The EQ-5D was compared with the Short-Form 36 Health Survey (SF-36), the modified transplant symptom occurrence and symptom distress (MTSOSD) scale, the short-form Beck Depression Inventory (BDI), and the State Trait Anxiety Inventory (STAI). Construct and concurrent validity were tested on cross-sectional data of 350 patients.

**Results:** Construct validity is good for some but not all EQ-5D dimensions, and the EQ-5D discriminates well among groups of patients with different health states according to the SF-36, MTSOSD scale, BDI, and STAI and between patients and the general public. Concurrent validity is good, as shown by the correspondence of EQ-5D and SF-36 results.

**Conclusion:** It is concluded that the EQ-5D is a valid instrument for the measurement of health status in renal transplant patients.

**Keywords:** concurrent validity, construct validity, EQ-5D, renal transplantation.

## Introduction

The increasing need to set priorities in health care has urged health professionals to develop instruments that reflect patients' subjective health status and use these in economic evaluations. Subjective health status is the personal perception of a patient's own health. This concept is closely related to the concept health-related quality of life—and in the literature both concepts are used interchangeably—although health-related quality of life can be interpreted more broadly, including all aspects of life that have a possible relationship with health. Because there is a large variability in the aspects of life that are considered important for an individual's health-related quality of life, health-related quality of life is more difficult to measure and it cannot be claimed that a generic instrument, with a limited

number of non-disease-specific dimensions, measures actual health-related quality of life.

One of the concerns in using generic questionnaires, as required for economic evaluations, is that they may fail to capture small but relevant changes in the health status of specific patient groups. Developers of health status measurement instruments necessarily make trade-offs between simplicity and comprehensiveness. Simplicity of the instrument is crucial for the feasibility of using the instrument in routine practice or in large clinical or economic studies. Moreover, the instrument should allow the translation of the health state description into one single index value to allow the calculation of the number of quality-adjusted life-years gained by an intervention compared to other interventions.

The EQ-5D is a generic five-dimensional instrument that is used worldwide for health status measurement [1,2]. Each health state description corresponds with an index value, derived from the general public [3]. The validity and reliability of the EQ-5D have been tested in the general public of different countries [4–7] as well as in several patient

*Address correspondence to:* Irina Cleemput, Belgian Health Care Knowledge Centre, Wetstraat 155, B-1040 Brussels, Belgium. E-mail: Irina.Cleemput@kenniscentrum.fgov.be

populations, such as rheumatoid arthritis patients [8] and patients with osteoarthritis of the knee [9], lower limb ischemia [10], sleep apnea [11], chronic obstructive pulmonary disease [12], and intestinal failure [13].

To be useful for comparisons across patient groups, the EQ-5D should be valid in all patient populations. The validity of an instrument is “the degree to which evidence and theory support the interpretation of the test scores entailed by the proposed uses of tests” [14]. The proposed use of the EQ-5D is as an outcome measure in economic evaluations. Hence, to be valid the EQ-5D must allow sensible comparisons among patient groups and/or patients with different characteristics. The problem with evaluating the validity of an instrument that measures subjective health status, such as the EQ-5D, is that there is no gold standard against which the instrument can be tested. A number of concepts that focus on different aspects of validity have been described in the literature. A recent classification of validity concepts explains which evidence is needed to perform a validity test: evidence on test content, response process, internal structure, relations to other variables, and consequences of testing (see Goodwin [15] for a full discussion).

As yet, there is no evidence on the validity of the EQ-5D in the renal transplant population. Such evidence is important, however, to allow decision makers to assess the usefulness of the EQ-5D for economic evaluations that are used to set priorities between different disease areas. Evidence on the validity of the EQ-5D in other disease areas does not guarantee validity in the renal transplant population. The aim of this study was therefore to assess the construct and concurrent validity of the EQ-5D in renal transplant patients.

## Methods

### *Patient Sample*

The analyses in this article are based on the data of a cross-sectional descriptive study that aimed to examine the subjective health status of patients following a tacrolimus-based immunosuppressive regimen after renal transplantation (RTX) [16]. The sample consisted of 361 renal transplant patients from 3 RTX centers in Belgium and the Netherlands. Patients were considered eligible for participation in the study if they were 18 years of age or older, literate, and Dutch or French speaking; had a post-transplant status of at least 6 months; and had not received a combined transplantation (e.g., kidney-pancreas) or a second renal transplantation. All

patients were asked for a written informed consent before the questionnaires were distributed. Eleven patients refused to participate in the study, resulting in a complete data set for 350 patients. At the moment of inclusion, patients had a median post-RTX status of 16.7 months ( $Q_1 = 7.9$ ;  $Q_3 = 38.6$ ).

### *Variables and Measurement*

The validity of the EQ-5D was tested by means of comparisons with already validated instruments for the measurement of different aspects of health status in the renal transplant population. Instruments include the Short-Form 36 Health Survey (SF-36) [17] as a generic profile measure for health status, the short-form Beck Depression Inventory (BDI) [18] for the measurement of depressive symptomatology, and the modified transplant symptom occurrence and symptom distress (MTSOSD) scale [19] for the measurement of symptom experience and symptom distress associated with side effects of immunosuppressive drugs. All 350 patients who agreed to participate in the study were given an SF-36, an EQ-5D, an MTSOSD scale, and a BDI at inclusion in the study.

**EQ-5D.** The EQ-5D is a simple two-page questionnaire consisting of a self-classifier and a visual analog scale (VAS) on which patients can value their current health state. The developers of the EQ-5D relied on existing instruments for the measurement of health status, such as the quality of well-being scale, the Nottingham Health Profile, the Sickness Impact Profile, and the Rosser Index [20].

The self-classifier includes five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with three levels of severity in each dimension (no problems, some problems, and severe problems). This classification allows for the description of 243 ( $3^5$ ) different health states. Each of these 243 health states corresponds with a single index measure between 0 (= dead) and 1 (= perfect health), derived from and reflecting the preferences of the general public [3]. The index values used for the analyses in this study were derived from a large Flemish EuroQol validation survey [21].

The reliability of the EQ-5D with VAS is acceptable, as shown by studies in different patient populations and the general public [8,9,12,22–26]. The reported intraclass correlation coefficients range from 0.7 to 0.93.

**Short-form 36 health survey.** The SF-36 is a generic instrument for the measurement of patients' current health status. The SF-36 is often referred as

an instrument that measures quality of life. Given the far more limited nature of health status compared to quality of life, however, preference is given to the term health status. The SF-36 contains 36 items that cover eight dimensions of health: physical functioning, role limitation due to physical problems, bodily pain, general health perception, vitality, social functioning, role limitation due to emotional problems, and mental health [17]. It results in a profile for health, with a value between 0 (worst health) and 100 (best health) for each dimension. Because the SF-36 does not yield one single index score for general health status, it has, so far, been inappropriate for economic evaluations. Recent attempts have been made, however, to derive a preference-based single index value from a revised six-dimensional classification of the SF-36 [27]. This approach is still under development.

The validity of the subscales of the SF-36 has been established against clinically defined groups in general [28] and renal transplant populations in particular [29,30]. The SF-36 has been used frequently in studies on RTX for the measurement of health status [29–38] as well as for the validation of a new disease-specific instrument [39]. Piehlmeier and colleagues [40] examined the validity, reliability, and responsiveness of the SF-36 in a renal transplant population and concluded that the SF-36 is a reliable (internal consistency), valid (high correlation with overall evaluations of the health-related quality of life), discriminant (detects differences in clinical status), and responsive (detects differences in patients before and after transplantation) instrument for evaluating subjective health status.

*Short-form Beck Depression Inventory.* The short-form BDI is an instrument containing 13 items that measures depressive symptomatology in patients [18,41]. In contrast to the original version of the BDI with 22 items [42], the short-form BDI does not include any physical symptoms of depression, thus preventing false-positive results in chronically ill patients such as renal transplant recipients. The 13 items are scored with three response categories [1–3]. The total score is obtained by summing the scores on the individual items. On the basis of their individual total score on the BDI, patients are classified into four categories: no depression (0–4), mild depression (5–7), moderate depression (8–15), or severe depression (16–39).

The BDI has been used frequently for the measurement of depression in renal transplant populations [43–46]. One study examined the validity of using the 22-item BDI in a renal transplant popula-

tion [47]. It was found that the symptoms that most discriminate between depressed and nondepressed renal transplant patients were the symptoms included in the short-form BDI (e.g., death wish, suicidal intention, guilt, dysphoric mood, and problems in concentration) [47]. Both the long- and the short-form BDI have good internal consistency reliability [48,49].

*Modified transplant symptom occurrence and symptom distress scale.* The MTSOSD scale is a 45-item questionnaire about a patient's perceived occurrence of symptoms related to the side effects of the immunosuppressive regimen after transplantation and the distress associated with these symptoms [19]. It includes symptoms such as increased hair growth, itching, impotence, fatigue, and trembling hands. The MTSOSD scale is a disease-specific instrument, developed especially for the transplant population. For each potential symptom, occurrence (cognitive component of symptom experience) is rated on a scale from 0 (never) to 4 (always), and distress (emotional component of symptom experience) is rated on a scale from 0 (not distressing) to 4 (extremely distressing). The psychometric properties of the MTSOSD scale are described elsewhere [19]. The authors conclude that the MTSOSD scale is a valid instrument for the measurement of symptom occurrence and symptom distress related to side effects of immunosuppressive therapy. Test-retest reliability was not assessed and internal consistency could not be tested because the conditions for this test were not satisfied.

### Hypotheses

Based on prior research in the renal transplant population, five hypotheses were developed for this study that address anticipated relationships between the EQ-5D and other relevant variables. The tested hypotheses and the instruments used are schematically presented in Table 1.

Given that the SF-36 and the EQ-5D both claim to measure (subjective) health status, the results obtained with both instruments should be comparable ( $H_4$ ). Results of the BDI and MTSOSD scale should be comparable with the responses on the EQ-5D dimensions anxiety/depression and pain/discomfort, respectively, because both claim to measure the same aspect of health ( $H_5$ ). Therefore, patients were grouped by their response on the relevant dimension of the EQ-5D, and scores on the BDI, MTSOSD scale, and SF-36 were compared among the patient groups. In case of comparisons with the SF-36, it is expected that there are larger

**Table 1** Hypotheses tested and instruments used in the validation study

		Instruments
<b>Construct validity</b>		
<i>Convergent validity</i>		
H <sub>1</sub>	An inverse relationship exists between age and level of reported problems on mobility, self-care, usual activities, and pain/discomfort on the EQ-5D [2]	EQ-5D, demographic questionnaire
H <sub>2</sub>	Women report more often anxiety/depression on the EQ-5D than men [50]	EQ-5D, demographic questionnaire
<i>Discriminant validity</i>		
H <sub>3</sub>	The EQ-5D is able to discriminate between renal transplant patients and—presumably more healthy—people from the general Flemish public	EQ-5D
<b>Concurrent validity</b>		
H <sub>4</sub>	Differences in mean SF-36 scores between patients, grouped according to their response on the EQ-5D, are largest for comparable dimensions of the SF-36 and the EQ-5D	EQ-5D, SF-36
H <sub>5</sub>	The classification of patients according to their response on the anxiety/depression dimension and the pain/discomfort dimension of the EQ-5D reflects the patterns obtained from the BDI and the MTSOSD scale, respectively	EQ-5D, BDI, MTSOSD scale

differences between the mean SF-36 scores of patients grouped according to their response on comparable domains of the EQ-5D (e.g., mobility<sub>EQ-5D</sub> and physical functioning<sub>SF-36</sub> or pain/discomfort<sub>EQ-5D</sub> and bodily pain<sub>SF-36</sub>) than between the mean SF-36 scores of patients grouped according to their response on noncomparable domains of the EQ-5D (e.g., mobility<sub>EQ-5D</sub> and role limitations due to emotional problems<sub>SF-36</sub>). This should reveal whether the EQ-5D dimension response distinguishes between groups of patients with higher or lower mean SF-36 scores. For the BDI, it is expected that the classification of the EQ-5D into “no depression,” “some depression,” and “severe depression” reflects the classification into no, mild, moderate, and severe depressive symptomatology of the BDI [41]. For the MTSOSD scale, it is expected that mean total scores differ significantly between groups of patients with different levels of pain/discomfort on the EQ-5D.

#### Statistical Analysis

Differences in percentages of patients reporting problems on the respective dimensions of the EQ-5D between men and women and among the different age groups were tested by means of a chi-square test statistic. Differences in EQ-5D index scores among age groups were tested by means of a Kruskal–Wallis test. For differences in EQ-5D index scores between men and women an independent sample Student’s *t* test could be used because of the large number of patients in both groups. Age groups were defined as 10-year intervals between the ages of 20 and 80.

Differences between percentage of patients reporting problems on the five EQ-5D dimensions and the percentage of respondents from the general public were tested by means of a Fisher exact test,

because no single patient reported extreme problems with mobility.

Comparisons of BDI and MTSOSD scale scores across patient groups were made by means of an analysis of variance (ANOVA). For comparisons of SF-36 scores across patient groups an independent sample Student’s *t* test was used; response categories were dichotomized because a very small proportion of the respondents reported severe problems on the EQ-5D dimensions. The nonparametric Spearman rank correlation coefficients were calculated between EQ-5D index scores and SF-36 domain scores because the EQ-5D index scores were not normally distributed.

Symptom occurrence and symptom distress were analyzed using riddit analysis [51,52]. Riddit analysis, as used in this study, calculates the probability that a respondent reports a level of symptom occurrence, or distress, that is higher than the symptom occurrence, or distress, reported by a randomly selected patient from the entire patient sample, given his or her profile of responses on all the symptoms. Riddits are hence probabilities that are calculated for each individual across all symptoms, both for symptom occurrence and for symptom distress. Individual riddit scores were compared with the riddits of all patients in the sample; i.e., the entire patient sample was used as the reference group with which individual scores were compared. For example, if a patient has higher symptom occurrence than the “average” patient from the entire sample, his or her riddit score will be higher than 0.5. The “average” patient has a riddit of 0.5: the probability that his or her level of symptom occurrence is higher than a randomly selected patient from the entire patient sample is 50%; the probability that it is lower is also 50%. Riddits give an indication of the profile of a patient. A patient who reports high levels of symptom

All statistical analyses were performed with SPSS statistical software version 9.0 (SPSS Inc., Chicago, IL, 1999). The level of significance was set at 5%.

### Demographic Characteristics

### Construct Validity

Significantly more women are anxious or depressed than men. Men and women do not show significant differences on any other dimension of the EQ-5D (Table 2).

Demographic characteristics of the patient group are significantly different from those of the general public; patients are significantly older ( $P = 0.024$ ) and have higher educational levels ( $P = 0.001$ ), and there are more men in the patient sample than in the general public sample ( $P < 0.0005$ ). A multivariate analysis, in which the EQ-5D index values are corrected for age, educational level, and sex, reveals stable results, with patients having a lower EQ-5D index score than people from the general public.

**Table 2** Percentage of respondents reporting problems (moderate or severe) on the five dimensions of the EQ-5D by age category and sex

	Age in years						Sex			
	20-29 (n = 19)	30-39 (n = 41)	40-49 (n = 78)	50-59 (n = 94)	60-69 (n = 79)	70-79 (n = 25)	P value	Male (n = 209)	Female (n = 141)	P value
Mobility	10.5	26.8	28.2	34.0	55.7	68.0	<0.01*	40.7	34.0	0.22*
Self-care	0	12.2	5.1	4.3	8.9	20.0	0.05*	6.2	8.5	0.53*
Usual activity	10.5	24.4	32.1	34.0	32.9	36.0	0.36*	27.7	36.2	0.10*
Pain/discomfort	42.1	48.7	61.6	57.4	54.4	40.0	0.33*	52.6	58.1	0.33*
Anxiety/depression	15.8	29.3	34.6	35.1	26.6	24.0	0.46*	23.9	40.5	0.01*
Mean EQ-5D index score (SD)	0.86 (0.15)	0.76 (0.25)	0.71 (0.22)	0.72 (0.21)	0.69 (0.20)	0.73 (0.17)	0.01†	0.75 (0.21)	0.70 (0.21)	0.21†

\*\*Chi-square test.

<sup>†</sup>Kruskal–Wallis test.<sup>#</sup>Independent sample Student's *t* test.



**Table 3** Percentage of respondents reporting no, some, and severe problems on each of the EQ-5D dimensions

	Study sample (n = 350)	General public (n = 967)	P value
Mobility			
No problems	62.0	83.1	<0.01*
Some problems	38.0	16.7	
Severe problems	0	0.2	
Self-care			
No problems	92.9	95.0	0.17*
Some problems	6.6	4.1	
Severe problems	0.6	0.9	
Usual activities			
No problems	68.9	82.5	<0.01*
Some problems	27.1	15.9	
Severe problems	4.0	1.6	
Pain/discomfort			
No problems	45.1	53.0	0.04*
Some problems	52.0	44.5	
Severe problems	2.9	2.6	
Anxiety/depression			
No problems	69.4	77.3	0.01*
Some problems	26.9	21.7	
Severe problems	3.7	1.0	
EQ-5D index, mean (SD)	0.73 (0.21)	0.81 (0.21)	<0.01†

\*Fisher exact test.

†Mann-Whitney U test.

### Concurrent Validity ( $H_4$ and $H_5$ )

For the concurrent validity testing, patients were grouped according to their responses on the EQ-5D to see whether this classification reflected the patterns obtained with the SF-36, BDI, STAI, and MTSOSD scale. Results for the SF-36 are presented in Table 4.

Patients reporting some or severe problems on the EQ-5D dimensions have significantly lower SF-36 scores for almost all SF-36 domains. A few exceptions are observed. The difference in SF-36

scores for mental health among patients reporting no, moderate, or severe problems on mobility is not significant ( $P = 0.162$ ). The same applies for the difference in scores on role limitation due to emotional problems and mental health between groups reporting no, moderate, or extreme problems on self-care ( $P = 0.160$  and  $P = 0.062$ ). Differences are largest between dimensions of the SF-36 and the EQ-5D covering comparable domains of health (e.g., mobility<sub>EQ-5D</sub> and physical functioning<sub>SF-36</sub>, mental health<sub>SF-36</sub> and anxiety/depression<sub>EQ-5D</sub>, bodily pain<sub>SF-36</sub> and pain/discomfort<sub>EQ-5D</sub>, physical role limitation<sub>SF-36</sub> and usual activities<sub>EQ-5D</sub>).

The range of possible scores on the different health domains is larger in the SF-36 than in the EQ-5D. In the EQ-5D the range of responses on the level of the dimensions is, by design, limited to three, whereas the domain scores in the SF-36 can be any discrete number between 0 and 100. This implies that the SF-36 has a higher potential to capture small changes in health status if these changes manifest themselves in one health domain than the EQ-5D.

Spearman rank correlation coefficients between the EQ-5D index scores and the SF-36 domain range from 0.41 (with role limitation due to emotional problems) to 0.68 (with bodily pain) (all  $P < 0.01$ ). The correlations with all other domains of the SF-36 are also significant on the 1% significance level. This means that more problems on any domain of the SF-36 are associated with lower EQ-5D index scores and vice versa.

For the BDI, it is found that the classification into “not,” “moderately,” or “severely” depressed resulting from the responses on the anxiety/depres-

**Table 4** SF-36 scores by EQ-5D response

	SF-36 domain								
EQ-5D	N	PF	SF	RP	RE	BP	MH	V	GHP
Mobility									
No problems	217	84.0	83.2	71.5	78.5	80.1	72.8	68.7	65.3
Any problems	133	55.3*	71.6*	45.9*	66.2*	63.3*	69.8	54.6*	53.7*
Self-care									
No problems	325	75.8	80.3	64.3	74.9	75.3	72.3	64.7	61.9
Any problems	23	38.0*	58.5*	29.0*	60.0	53.4*	63.7	45.2*	47.9*
Usual activities									
No problems	241	81.9	85.0	74.8	81.3	80.6	75.3	69.1	65.7
Any problems	95	53.6*	65.1*	33.0*	57.2*	58.4*	63.6*	50.5*	50.1*
Pain/discomfort									
No problems	158	79.6	87.1	76.6	86.7	90.4	78.2	72.3	68.2
Any problems	182	67.7*	71.9*	49.6*	63.2*	60.0*	66.3*	56.0*	54.9*
Anxiety/depression									
No problems	243	74.6	84.5	67.4	85.3	77.9	79.7	69.1	64.6
Any problems	94	69.6*	65.9*	49.0*	47.7*	64.3*	53.4*	50.3*	52.4*

\* $P < .01$  according to an independent sample t test.

Abbreviations: PF, physical functioning; SF, social functioning; RP, role limitations due to physical problems; RE, role limitations due to emotional problems; BP, bodily pain; MH, mental health; V, vitality; GHP, general health perception.

**Table 5** Percentage of patients reporting no, some, or extreme anxiety/depression, given their BDI classification and mean BDI score for respondents with different answers on the EQ-5D anxiety/depression dimension

EQ-5D: anxiety/depression	Classification according to the short-form BDI				Total	BDI score mean (SD)
	No depression	Mild depression	Moderate depression	Severe depression		
No problems	80.6	58.5	20.0	0	69.4	2.31 (2.59)
Some problems	18.2	39.6	65.0	25	26.9	5.59 (3.97)
Severe problems	1.2	1.9	15.0	75	3.7	12.54 (8.54)
Total	72.3	15.1	11.4	1.1	100	3.57 (4.06)

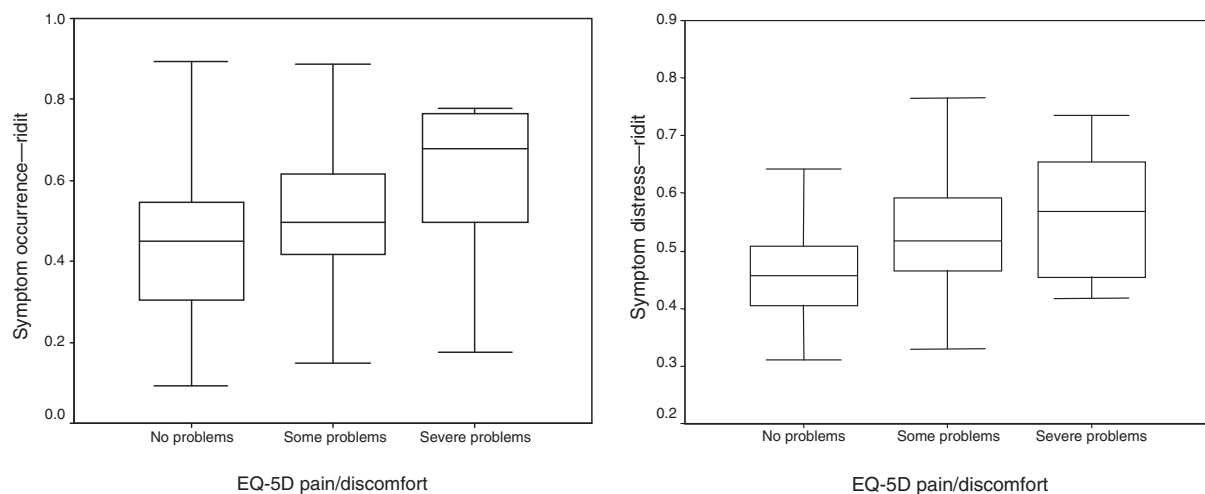
sion dimension on the EQ-5D reflects the classification of the BDI into “no,” “mild,” “moderate,” and “severe” depression [41] (Table 5). The majority of the patients without depression according to the BDI report no anxiety or depression on the EQ-5D (80.6%). The balance is changed the more depressive symptoms patients have according to the BDI. Still a majority of patients with mild depression report no anxiety or depression on the EQ-5D (58.5%), but patients with moderate depression are more likely to report some anxiety or depression on the EQ-5D. Seventy-five percent of the patients with severe depression according to the BDI also report extreme anxiety or depression on the EQ-5D (Table 5).

The concurrent validity of the EQ-5D with respect to depression is also illustrated by the significant difference between the mean BDI scores of the different groups of patients, defined by their response on the anxiety/depression dimension of the EQ-5D (Table 5). Patients with extreme anxiety/depression according to the EQ-5D have significantly higher BDI scores than

patients with moderate or no anxiety/depression according to the EQ-5D ( $P < 0.0005$ ). The same applies for patients with moderate anxiety/depression in comparison with patients with no anxiety/depression.

Mean ridits for symptom occurrence and symptom distress are significantly different among patients grouped by their level of pain/discomfort ( $P < 0.01$ ). Respondents reporting more pain/discomfort have significantly higher ridits (Fig. 1). This means that the probability that a respondent reports higher symptom experience and distress than a randomly selected patient from the entire sample increases as the level of problems reported on the EQ-5D pain/discomfort dimension increases.

The negative correlations between the EQ-5Q index scores and the MTSOSD scale scores ( $\rho = -0.386$  for symptom experience and  $\rho = -0.498$  for symptom distress) are highly significant ( $P < 0.0005$ ), meaning that the more symptoms patients experience and the more distressed they are about these symptoms, the lower their score for overall health.

**Figure 1** Box plots of ridit scores for the symptom occurrence and symptom distress subscales of the MTSOSD scale, according to response on EQ-5D pain/discomfort.

## Discussion

The purpose of this study was to investigate the construct and concurrent validity of the EQ-5D in renal transplant recipients. The construct validity of the instrument was tested by looking at the convergent and discriminant validity of the instrument.

The hypothesis of an inverse relationship between age and level of problems on mobility, self-care, usual activities, and pain/discomfort ( $H_1$ ) could not entirely be accepted. The expected relationship was seen only between age and mobility. An unexpected pattern in frequency of reported problems with self-care was observed. The percentage of patients reporting problems on the EQ-5D self-care dimension was high in the category 30 to 39 years of age, declines up to 50 to 59 years of age, and then increases again up to 70 to 79 years of age. Unexpected patterns were also observed in the usual activities and pain/discomfort dimensions. The results are in contrast to most existing research in general populations [2,53–56], in which an almost systematic increase in frequency of reported problems is observed on all dimensions, except for anxiety/depression.

There are three possible explanations for these findings. In chronic patient populations disease severity is a covariate that might count for the unexpected correlation between age and self-care limitations. Because there was no disease severity index available in this study, this is a hypothesis that should be tested in future research. The unusual pattern in usual activities may be due to the fact that younger patients may suffer from their disease less long than the older patients. Therefore, they may not have adapted certain coping strategies, such as changing lifestyle, habits, and ways to perform usual activities to the same extent as the older patients. Because older patients have adopted their usual activities to their current physical capacity, they might no longer feel equally impaired in performing usual activities. The same applies for the self-care dimension. The longer people suffer from a chronic disease, the more they find ways to get around this problem. The unusual pattern observed on the pain/discomfort dimension may be explained by the fact that older patients generally have less comorbidities at transplantation than younger patients and thus a better self-reported health status afterward. There is evidence that older patients with many comorbidities are less likely to be transplanted than younger patients [57]. Patients with much comorbidity, e.g., those with cardiovascular diseases who were not transplanted at a younger

age, tend to grow old on dialysis and were hence not included in our study of transplant patients.

The second hypothesis ( $H_2$ ) could not be rejected because women did show higher levels of anxiety/depression than men, analogous to findings from the general public in different countries [2,53–56,58–61].

For the discriminant validity test, percentages of patients reporting problems on the respective EQ-5D dimensions were compared to percentages of people in the general public reporting problems on the same dimensions. Renal transplant patients scored worse on all EQ-5D dimensions than the general Flemish public ( $H_3$ ), though not significantly on the self-care dimension. The latter observation may be due to the relatively restricted formulation of the self-care question in the EQ-5D (“I have no/some/severe problems with washing or dressing myself”), whereas the other dimensions are formulated in more general terms. The difference persisted after correction for difference in age, sex, and educational level, three aspects that differed significantly between both samples.

For almost all dimensions of the EQ-5D, the SF-36 scores for all SF-36 domains differ significantly among patients grouped by their response on the respective EQ-5D dimensions ( $H_4$ ). Differences in SF-36 scores between comparable dimensions of the EQ-5D and the SF-36 were larger than between noncomparable dimensions, pointing at good concurrent validity of the EQ-5D. The response categories of the EQ-5D are very limited in comparison to the response categories of the SF-36. This is an acknowledged problem, however, for which the proposed solution is to use a separate disease-specific instrument alongside the generic EQ-5D [3] if more information about subdomains of health is required. For use in economic evaluations, however, the observed concurrent validity of the EQ-5D is satisfactory, especially in view of the highly significant positive correlation between the EQ-5D index score and the SF-36 score for general health perception.

The classification into depressed and not depressed patients according to the EQ-5D was highly comparable with the classification of the BDI ( $H_5$ ). If no information about potential reasons for depressive symptomatology in renal transplant patients is required, as in economic evaluations, the use of the descriptive EQ-5D system may be sufficient.

A similar pattern was found for the MTSOSD scale scores. Patients with more problems on self-care, usual activities, pain/discomfort, and anxiety/



depression have higher scores for symptom occurrence and symptom distress.

## Conclusions

In conclusion, the results of this study show that the EQ-5D has acceptable convergent and discriminant validity and concurrent validity for the measurement of health status in a renal transplant population. The results of the EQ-5D in renal transplant patients are in line with existing evidence. Consequently, the use of the EQ-5D in the outcome measurement of interventions for economic evaluations will reveal results that are consistent with evidence, which is an important indication for its validity in renal transplant patients.

One of the merits of the EQ-5D is that it is a simple instrument that can easily be administered during routine clinical practice. This feature makes the instrument especially attractive from a research perspective. Given the relatively good indications of validity of the instrument in a renal transplant population, we recommend further use of this instrument in this patient group.

This study was supported by an unrestricted scientific grant from Fujisawa GmbH, Germany, and by a grant from the Research Council of the K.U. Leuven, Belgium. The research was carried out at Center for Health Services and Nursing Research.

## References

- 1 Johnson JA, Coons SJ, Ergo A, et al. Valuation of EuroQol (EQ-5D) health states in an adult US sample. *Pharmacoeconomics* 1998;13:421–33.
- 2 Kind P, Dolan P, Gudex C, et al. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998;316:736–41.
- 3 Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.
- 4 Brazier J, Jones N, Kind P. Testing the validity of the EuroQol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res* 1993;2:169–80.
- 5 van Agt HME, Essink-Bot ML, Krabbe PFM, et al. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. *Soc Sci Med* 1994;11:1537–44.
- 6 Essink-Bot ML, Krabbe PFM, Bonsel GJ, et al. An empirical comparison of four generic health status measures: the Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Med Care* 1997;35:522–37.
- 7 Johnson JA, Coons SJ. Comparison of the EQ-5D and SF-12 in an adult US sample. *Qual Life Res* 1998;7:155–66.
- 8 Hurst NP, Jobanputra P, Hunter M, et al. Validity of EuroQol—a generic health status instrument—in patients with rheumatoid arthritis. *Br J Rheumatol* 1994;33:655–62.
- 9 Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. *Rheumatology* 1999;38:807–13.
- 10 Chetter IC, Spark JI, Dolan P, et al. Quality of life in patients with lower limb ischaemia: suggestions for European Standardisation. *Eur J Endovasc Surg* 1997;13:597–604.
- 11 Jenkinson C, Stradling J, Petersen S. Comparison of three measures of quality of life outcome in the evaluation of continuous positive airways pressure therapy for sleep apnoea. *J Sleep Res* 1997;6:199–204.
- 12 Harper R, Brazier J, Waterhouse JC, et al. Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax* 1997;52:879–87.
- 13 Richards DM, Irving MH. Assessing the quality of life of patients with intestinal failure on home parenteral nutrition. *Gut* 1997;40:218–22.
- 14 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington (DC): American Educational Research Association, 1999.
- 15 Goodwin LD. Changing conceptions of measurement validity: an update on the new standards. *J Nurs Educ* 2002;41:100–6.
- 16 Moons P, Vanrenterghem Y, Van Hooff JP, et al. Health-related quality of life and symptom experience in tacrolimus-based regimens after renal transplantation: a multicentre study. *Transplant Int* 2002;16:653–64.
- 17 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). *Med Care* 1992;30:473–83.
- 18 Beck AT, Beck RW. Screening depressed patients in family practice: a rapid technique. *Postgrad Med* 1972;52:81–5.
- 19 Moons P, De Geest S, Versteven K, et al. Psychometric properties of the “Modified Transplant Symptom Occurrence and Symptom Distress Scale”. *J Nurs Meas* 2001;9:115–34.
- 20 The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
- 21 Cleemput I, Kind P, Kesteloot K. Re-scaling social preference data: implications for modeling. In: Kind P, Macaran S, eds., 19th Plenary Meeting of the EuroQol Group. York: University of York, 2002.
- 22 Dorman P, Slattery J, Farrell B, et al. Qualitative comparison of the reliability of health status

- assessments with the EuroQol and SF-6 questionnaires after stroke. *Stroke* 2000;29:63–8.
- 23 Badia X, Monserrat S, Roset M, et al. 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;8:303–10.
  - 24 Gudex C, Dolan P, Kind P, et al. Health state valuations from the general public using the visual analogue scale. *Qual Life Res* 1996;5:521–31.
  - 25 Ünal G, de Boer JB, Borsboom GJJM, et al. A psychometric comparison of health-related quality of life measures in chronic liver disease. *J Clin Epidemiol* 2001;54:587–96.
  - 26 Stavem K, Bjornaes H, Lossius MI. Properties of the 15D and EQ-5D utility measures in a community sample of people with epilepsy. *Epilepsy Res* 2001;44:179–89.
  - 27 Brazier J, Roberts J, Deverill H. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271–92.
  - 28 Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. *Qual Life Res* 1993;2:369–95.
  - 29 Fujisawa M, Ichikawa Y, Yoshiya K, et al. Assessment of health-related quality of life in renal transplant and hemodialysis patients using the SF-36 health survey. *Urology* 2000;56:201–6.
  - 30 Bremer BA, Wert KM, Durica AL, et al. Neuropsychological, physical, and psychosocial functioning of individuals with end-stage renal disease. *Ann Behav Med* 1998;19:348–52.
  - 31 Manu MA, Radulescu S, Harza M, et al. Quality of life assessed by SF-36 health survey in renal transplant patients. *Transplant Proc* 2001;33:1927–8.
  - 32 Gross CR, Limwattananon C, Matthees B, et al. Impact of transplantation on quality of life in patients with diabetes and renal dysfunction. *Transplantation* 2000;70:1736–46.
  - 33 Singer MA, Hopman MW, MacKenzie TA. Physical functioning and mental health in patients with chronic medical conditions. *Qual Life Res* 1999;8:687–91.
  - 34 Rebollo P, Ortega F, Baltar JM, et al. Health related quality of life (HRQOL) of kidney transplanted patients: variables that influence it. *Clin Transplant* 2000;14:199–207.
  - 35 Forsberg A, Lorenzon U, Nilsson F, et al. Pain and health related quality of life after heart, kidney and liver transplantation. *Clin Transplant* 1999;13:453–560.
  - 36 Matas AJ, McHugh L, Payne WD, et al. Long-term quality of life after kidney and simultaneous pancreas-kidney transplantation. *Clin Transplant* 1998;12:233–42.
  - 37 Shield CF, McGrath MM, Goss TF. Assessment of health-related quality of life in kidney transplant patients receiving tacrolimus (FK506)-based versus cyclosporine-based immunosuppression. *Transplantation* 1997;64:1738–43.
  - 38 Benedetti E, Matas AJ, Hakim N, et al. Renal transplantation for patients 60 years or older. A single-institution experience. *Ann Surg* 1994;220:445–58.
  - 39 Franke GH, Riemer J, Kohnle M, et al. Quality of life in end-stage renal disease patients after successful kidney transplantation: development of the ESRD symptom checklist—transplantation module. *Nephron* 1999;83:31–9.
  - 40 Piehlmeier W, Bullinger M, Kirchberger I, et al. Evaluation of the quality of life of patients with insulin-dependent diabetes mellitus before and after organ transplantation with the SF-36 health survey. *Eur J Surg* 1996;162:933–40.
  - 41 Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psych Rev* 1988;8:77–100.
  - 42 Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. *Arch General Psych* 1961;4:561–7.
  - 43 Christensen AJ, Ehlers SL, Raichle KA, et al. Predicting change in depression following renal transplantation: effect of patient coping preferences. *Health Psychol* 2000;19:348–53.
  - 44 Frazier PA, Davis-Ali SH, Dahl KE. Correlates of noncompliance among renal transplant recipients. *Clin Transplant* 1994;8:550–7.
  - 45 Shulman R, Price JD, Spinelli J. Biopsychosocial aspects of long-term survival on end-stage renal failure therapy. *Psychol Med* 1989;19:945–54.
  - 46 Schlebusch L, Pillay BJ, Louw J. Depression and self-report disclosure after live related donor and cadaver renal transplants. *S Afr Med J* 1989;75:490–3.
  - 47 Smith MD, Hong BA, Robson AM. Diagnosis of depression in patients with end-stage renal disease: comparative analysis. *Am J Med* 1985;79:160–6.
  - 48 Richter P, Werner J, Heerlien A, et al. On the validity of the Beck Depression Inventory: a review. *Psychopathology* 1998;31:160–8.
  - 49 Leahy JM. Validity and reliability of the Beck Depression Inventory-Short Form in a group of adult bereaved females. *J Clin Psychol* 1992;48:64–8.
  - 50 Kawachi I, Berkman LF. Social ties and mental health. *J Urban Health* 2001;78:458–67.
  - 51 Sermeus W, Delesie L. Redit analysis on ordinal data. *West J Nurs Res* 1996;18:351–9.
  - 52 Bross I. How to use redit analysis. *Biometrics* 1958;14:18–38.
  - 53 Burstrom K, Johannesson M, Diderichsen F. Health-related quality of life by disease and socioeconomic group in the general population in Sweden. *Health Policy* 2001;55:51–69.
  - 54 Burstrom K, Johannesson M, Diderichsen F. Swedish population health-related quality of life

- results using the EQ-5D. *Qual Life Res* 2001;10:621–35.
- 55 Anderson C, Wittenburg R. Health status of British adult population by the EuroQol instrument and other measures. EuroQol Plenary Meeting Barcelona 1995 1995;173–95.
- 56 Badia X, Schiaffino A, Alonso J, et al. Using the EuroQol EQ-5D in the Catalan general population: feasibility and construct validity. *Qual Life Res* 1998;7:311–22.
- 57 Krishnan M, Lok CE, Jassal SV. Epidemiology and demographic aspects of treated end-stage renal disease in the elderly. *Semin Dial* 2002;15: 79–83.
- 58 Devlin N, Hansen P, Herbison P. Variations in self-reported health status: results from a New Zealand survey. *N Z Med J* 2000;113:517–20.
- 59 Ikegami N, Ikeda S. Health status in the Japanese population: results from the Japanese EuroQol study. 16th Plenary Meeting of the EuroQol Group 1999;271–83.
- 60 Mayo NE, Goldberg M, Kind P. Performance of the EuroQol EQ-5D in a Canadian population. EuroQol Plenary Meeting Rotterdam 1997 1997;25–41.
- 61 Ohinmaa A, Sintonen H. Quality of life of Finnish population measured by EuroQol. EuroQol Plenary Meeting Barcelona 1995 1995;161–73.