ORIGINAL ARTICLE

A Malaysian Study on the Reliability and Validity of the Health-Related Quality of Life Questionnaire (HRQOL-20) in Urological Patients

K F Quek, M.Phil*, W Y Low, PhD*, A H Razack, FRCS, C S Loh, FRCS (Urol)***, *Health Research Development Unit, **Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, ***Sunway Medical Centre and Gleneagles Intan Medical Centre**

Summary

Main objective of this study is to validate the Health-Related Quality of Life (HRQOL-20) in the Malaysian population. Reliability and internal consistency were evaluated using the test-retest method and Cronbach's alpha. Responsiveness was expressed as the effect size. Internal consistency was excellent (Cronbach's alpha value = 0.68 to 0.87). Test-retest correlation coefficients and intraclass correlation coefficient were significant (ICC=0.58 and 0.91) as well as the high degree of sensitivity and specificity. The HRQOL-20 is a reliable, valid and sensitive to clinical changes in the Malaysian urological population.

Key Words: Health-Related Quality of Life, Cronbach's alpha, Intraclass correlation coefficient, Internal consistency, Test-retest reliability, Validity

Introduction

Quality of life (QoL) has been described as a sense of well-being, contentment, fulfillment, flourishing or degree of satisfaction with present life circumstances. It is also between the hopes and expectations of an individual and the present experience at a particular stage of time. The improvement of quality of life can be noted either by deficient life dimension or adjustment of the patients' expectations. Measurement of quality of life includes psychological status, physical status, social interaction and economic status[1]. Medical disorder such as Benign Prostatic Hyperplasia (BPH) has been found to affect one's quality of life. Symptomatic BPH is a highly prevalent disease. BPH is one of the commonest diseases occurring among elderly men which may affect the patient's quality of life[2]. It is well known that lower urinary tract symptoms (LUTS) can interfere with one's daily activities. Symptoms such as frequency, nocturia, urgency, urge incontinence and dribbling can affect patient's quality of life[3]. Among the domains of the quality of life reported to be most affected were the physical, mental, social which consists of sleep, anxiety, worry about the disease, mobility, leisure, daily activities, sexual activities and satisfaction with sexual relationship[4].

This article was accepted: 4 July 2001

Like many Quality of Life instruments such as the MOS 36-item Short Form Health Survey (SF-36)[5], the Health-Related Quality of Life (HRQOL-20) which was developed by Lukacs 1995 is a multidimensional tool and has been widely used in the studies of quality of life in the field of Urology. The 20-item of HRQOL-20 is easy to administer, simple, short and only requires 10 - 15 minutes to be completed[3].

The present study was carried out at University Hospital, Kuala Lumpur and was designed to assess the reliability and validity of the Health-Related Quality of Life in a Malaysian urological population.

Materials and Methods

Study sample

The patients were selected using the inclusion and exclusion criteria. The inclusion criteria for the patients with LUTS in the surgical and medical group consists of patients in stable clinical condition, literate and able to give informed consent. The exclusion criteria were patients treated surgically and medically for lower urinary tract symptoms prior to this study, less than 40 years old, illiterate and patients with any chronic and acute diseases. For the control group, these include patients who were free from all major chronic and acute diseases and those with renal stone but with minimal severity. This study protocol was approved by the Ethics' Committee, University Hospital Kuala Lumpur.

The HRQOL-20 was assessed in patients with LUTS undergoing transurethral resection of the prostate (TURP), patients receiving medical treatment and the control patients. Validity and reliability were assessed in patients with LUTS treated medically (N=108) and in the control group with renal stones (N=50). Sensitivity of change was assessed in a group of patients admitted for TURP (N=79). Both the LUTS and renal stones were diagnosed based on clinical criteria such as medical history, physical examinations and radiological investigations where applicable.

Questionnaire

The HRQOL-20 is a 20 items clinician visual analogue scale and each item is rated from 0 (Good) to 10 (Worst). The HRQOL-20 has four domains of physical, mental, social and global assessment. The HRQOL-20 questionnaire consists of 20 questions that were scored according to a length of 10cm Visual Analogue Scale (VAS). Patients were asked to place a mark on the line to indicate their status. The minimum and maximum scores for each question were 0 and 10 respectively.

The global HRQOL scale was calculated by summing the 20 VAS scores for the physical/ functional dimension such as locomotion, sexual activity, appetite and sleeping (6 questions), mental health status such as behavior, cognitive and emotional aspects (6 questions), social health status such as activity, social participation and personal relationship (6 questions) and global (overall) quality of life (2 questions). The sum of the scores for each group of questions gave three corresponding subscores (minimum 0, maximum 60) and an overall score (minimum 0, maximum 200) was calculated by combining the three subscores and those of the two general questions. Three questions explored the patients' perceived sexual status: two were from the physical/ functional subscore (sexual desire, quality of erection) and one from the social subscore (satisfaction with sexual life). They were analyzed separately by constructing a sexual score (minimum 0, maximum 30) from the responses to these three questions.

Procedure

Patients were required to complete the HRQOL-20 after obtaining their informed consent. All questionnaires were self-administered although guidance was available from one of the authors (K.F.Q) of this study. All patients included in the validity study were retested at twelve weeks after the first administration of the HRQOL-20. In assessing

the responsiveness, patients completed the questionnaires a week prior to surgery and were retested twelve weeks after TURP.

Data Analysis

Cronbach's alpha coefficient was used to assess the internal consistency of the HRQOL-20[6]. The internal consistency shows the resulting values of Cronbach's alpha for the scale when individual item are excluded from the analysis. Test-retest reliability was assessed using the intraclass correlation coefficient (ICC), which is derived from analysis of variance (ANOVA) repeated measures model. The values of ICC varies from 1 (perfectly reliable) to 0 (totally unreliable)[7]. The responsiveness was assessed by calculating the difference between HRQOL-20 before and after TURP and dividing it by the mean standard deviation of the scores before TURP (effect size)[8].

Guyatt statistic was used to employ the mean differences in HRQOL-20 scores before and after TURP by calculating each individual item by means of a paired t test or by dividing it by the mean standard deviation of stable patients with LUTS.

Comparison between the mean of pre-treatment and post-treatment items scores of patients who undergo TURP were done to assess the sensitivity of the HRQOL-20 whereas comparison between the mean pre-treatment and post-treatment items scores in control group was to assess the specificity.

Results

A total of 237 respondents participated in this study. The mean age of the patients in the medical group was 63.67 (SD=8.57) years, surgical group 70.01 (SD=8.17) and control group 50.04 (SD=12.29). In terms of ethnicity, the Chinese formed the largest ethnic groups in all three groups; medical (66.70%), surgical (75.90%) and control (58.00%) followed by Indians (28.70%, 22.80% and 42.00%) respectively and Others (4.60%, 1.30%, 0%) respectively. Patients with LUTS and higher HRQOL-20 scores indicating deterioration of quality of life ($p<0.001$). Of the 237,108 patients with LUTS have the following mean scores: physical score of 22.31 (SD=8.65), mental score of 18.99 (SD=9.21), social score of 20.13 (10.21), global score of 6.26 (SD=3.71) and HRQOL-20 of 68.19 (SD=26.52) while 79 patients undergoing TURP have mean physical score of 31.16 (SD=9.04), mental score of 25.67 (SD=8.91), social score of 34.72 (11.12), global score of 9.27 (SD=3.21), and HRQOL-20 score of 110.00 (SD=27.17).

A high internal consistency for the HRQOL-20 was observed for all domains of HRQOL-20 indicating a high level of homogeneity among items in the scale. The domains of physical mental, social, global and HRQOL-20 had an ICC of 0.87, 0.90, 0.89, 0.74 and 0.91 ($p<0.001$) respectively (Table I).

Table II showed the pre and post scores, mean difference, effect size, and the Guyatt statistic for individual items and for the total scores. The mean pre intervention score for the domains on the physical, mental, social, global and HRQOL-20 score were 31.16 (SD=9.04), 25.67 (SD=8.91), 34.72 (SD=11.12), 9.27 (SD=3.21) and 110.00 (SD=27.17) respectively while the mean post-intervention score were 26.49 (SD=6.37), 20.55 (SD=6.42), 26.94 (SD=8.19), 7.94 (SD=2.63) and 81.94 (SD=18.27, $p<0.0001$) respectively. This suggests an average improvement after TURP on the physical domain with a mean score of 4.67 (SD=17.32), mental 5.12 (SD=6.01), social 7.77 (SD=7.66), global 1.33 (SD=2.79) and HRQOL-20 19.05 (SD=19.58). Overall the high effect size and sensitivity to change (responsiveness), suggest an improvement in the TURP-induced urinary symptoms in these patients.

In treatment responsiveness, all items of the HRQOL-20 demonstrated a high degree of sensitivity and specificity to the effects of treatment (Table III). All items and domains in the

Table I
Validity and Reliability:
Mean Test-Retest Score, Intraclass Correlation Coefficient and Internal Consistency for Individual HRQOL-20 Items

HRQOL -20 items	Intraclass Correlation Coefficient* (ICC)	Internal Consistency**	Mean Test Score	SD	Mean Retest Score	SD	Mean Difference	SD	95% Confidence Interval Lower	Higher
2	0.86	0.86	1.92	1.98	1.86	1.73	0.066	1.32	-0.18	0.32
3	0.82	0.82	2.98	2.32	3.08	2.31	0.11	0.82	-0.45	0.24
4	0.77	0.77	2.03	1.83	1.77	1.61	0.26	1.49	-0.02	0.54
5	0.85	0.86	5.63	2.76	5.29	2.70	0.35	1.93	-0.02	0.71
9	0.78	0.79	4.32	2.65	3.92	2.65	0.40	2.21	-0.02	0.83
11	0.91	0.90	5.27	2.76	5.18	2.79	0.089	1.64	-0.22	0.40
6	0.82	0.81	3.32	2.24	3.25	2.12	0.075	1.73	-0.25	0.40
7	0.80	0.80	3.02	2.11	2.83	1.99	0.18	1.66	-0.13	0.50
12	0.81	0.81	2.79	2.02	2.69	1.82	0.10	1.53	-0.19	0.39
13	0.82	0.83	3.96	2.14	3.61	2.12	0.35***	1.64	-0.03	0.66
15	0.73	0.73	2.86	2.04	2.69	2.02	0.17	1.87	-0.19	0.52
18	0.83	0.84	3.02	2.32	2.70	2.27	0.31	1.72	-0.018	0.64
1	0.83	0.83	2.81	2.29	2.62	2.23	0.19	1.71	-0.14	0.51
8	0.75	0.76	4.25	2.72	3.72	2.56	0.53***	2.33	0.09	0.98
10	0.83	0.83	2.78	2.26	2.86	2.38	0.079	1.78	-0.42	0.26
14	0.83	0.83	2.93	2.08	2.77	1.98	0.16	1.55	-0.14	0.45
17	0.83	0.87	4.70	2.82	4.64	3.00	0.064	1.96	-0.31	0.44
20	0.79	0.79	2.55	1.86	2.45	1.75	0.099	1.51	-0.19	0.39
16	0.68	0.68	3.67	2.27	3.91	2.41	0.24	2.29	-0.68	0.19
19	0.58	0.58	2.69	2.07	2.58	1.93	0.12	2.18	-0.30	0.54
Overall physical	0.87	0.88	22.31	8.65	21.02	9.19	1.30	5.93	0.16	2.42
Overall Mental	0.90	0.90	18.99	9.21	17.77	9.62	1.21	5.64	0.14	2.29
Overall social	0.89	0.90	20.13	10.21	19.02	10.73	1.11	6.43	-0.12	2.34
Overall global	0.74	0.74	6.26	3.71	6.41	3.64	0.15	3.33	-0.79	0.48
Overall HRQOL-20	0.91	0.91	68.19	26.52	64.14	28.38	4.05	15.53	1.09	7.01

p<0.001 for all ICCs(*)
Cronbach's alpha: note that Cronbach's alpha value given for each item represent the effect of removing that item from the calculation of the alpha value (eg if item 1 is omitted, the resulting value for the scale is 0.86, if item 2 is omitted, it is 0.82, and so forth)(**)
t test for paired comparisons significant (***)

Table II
Responsiveness: Mean Scores Before and After TURP, Effect Size and Guyatt Statistic

HRQOL-20 Items	PreTURP Mean	SD	PostTURP Mean	SD	Mean Differences*	SD	Effect Size	Guyatt statistic
2	2.69	2.51	1.81	1.86	0.87	2.07	0.35	0.44
3	4.96	2.09	3.61	1.74	1.35	1.85	0.65	0.58
4	3.46	1.99	2.53	1.37	0.92	1.36	0.46	0.50
5	7.66	2.41	7.28	2.08	0.38**	2.02	0.16	0.14
9	4.83	2.23	3.92	2.01	0.90	1.94	0.40	0.34
11	7.16	2.50	7.26	2.36	0.098	1.39	0.04	0.036
6	3.70	1.72	3.57	1.64	0.13**	1.07	0.08	0.058
7	4.16	1.98	3.24	1.53	0.92	1.22	0.47	0.44
12	3.84	1.84	2.97	1.44	0.86	1.51	0.47	0.43
13	5.21	1.47	4.29	1.36	0.92	1.18	0.63	0.43
15	4.17	2.04	3.22	1.64	0.95	1.61	0.47	0.47
18	4.41	2.32	3.20	1.88	1.21	1.89	0.52	0.52
1	6.36	2.89	4.31	2.33	2.05	2.06	0.71	0.90
8	7.34	2.41	5.45	2.12	1.89	2.22	0.78	0.70
10	6.20	2.78	4.15	2.16	2.05	2.04	0.74	0.91
14	4.25	2.17	3.35	1.48	0.90	1.56	0.42	0.43
17	6.90	2.16	6.94	2.07	0.044**	1.76	0.02	0.016
20	3.71	1.87	2.77	1.10	0.94	1.48	0.50	0.51
16	5.71	2.12	5.10	1.95	0.61	1.67	0.29	0.27
19	3.58	1.80	2.83	1.36	0.75	1.33	0.42	0.36
Overall physical	31.16	9.04	26.49	6.37	4.67	7.32	0.52	0.54
Overall mental	25.67	8.91	20.55	6.42	5.12	6.01	0.58	0.56
Overall social	34.72	11.12	26.94	8.19	7.77	7.66	0.70	0.76
Overall global	9.27	3.21	7.94	2.63	1.33	2.79	0.41	0.36
Overall HRQOL-20	110.00	27.17	81.94	18.27	19.05	19.58	0.70	0.72

Effect size=Mean difference/SD PreTURP
Guyatt statistics=Mean difference/SD of stable LUTS patients (medication group)
**p<0.0001*
***Not significant*

Table III
HRQOL-20 Items Characteristics of Patients Undergoing TURP and the Control Group: Sensitivity and Specificity

Sensitivity Items	N	Mean Changes	SEM	t statistics	p value
2	79	0.87	0.23	3.76	0.0001
3	79	1.35	0.21	6.49	0.0001
4	79	0.92	0.15	6.04	0.0001
5	79	0.38	0.23	1.69	0.094
9	79	0.90	0.22	4.15	0.0001
11	79	0.098	0.16	0.63	0.532
6	79	0.13	0.12	1.07	0.288
7	79	0.92	0.14	6.69	0.0001
12	79	0.86	0.17	5.07	0.0001
13	79	0.92	0.13	6.90	0.0001
15	79	0.95	0.18	5.23	0.0001
18	79	1.21	0.21	5.69	0.0001
1	79	2.05	0.23	8.85	0.0001
8	79	1.89	0.25	7.56	0.0001
10	79	2.05	0.23	8.91	0.0001
14	79	0.89	0.18	5.10	0.0001
17	79	0.044	0.02	0.22	0.826
20	79	0.94	0.17	5.69	0.0001
16	79	0.61	0.19	3.22	0.002
19	79	0.75	0.15	5.02	0.0001
Overall physical	79	4.67	0.82	5.68	0.0001
Overall mental	79	5.12	0.68	7.58	0.0001
Overall social	79	7.77	0.86	9.02	0.0001
Overall global	79	1.33	0.31	4.22	0.0001
Overall HRQOL-20	79	19.05	2.20	8.65	0.0001
2	50	0.45	0.19	2.34	0.024
3	50	0.33	0.24	1.34	0.185
4	50	0.31	0.26	1.15	0.254
5	50	0.34	0.29	1.17	0.247
9	50	0.22	0.35	0.64	0.528
11	50	0.19	0.27	0.71	0.483
6	50	0.05	0.18	0.26	0.794
7	50	0.40	0.27	1.52	0.135
12	50	0.16	0.28	0.59	0.559
13	50	0.32	0.24	1.32	0.192
15	50	0.29	0.28	1.04	0.303
18	50	0.098	0.30	0.33	0.744
1	50	0.057	0.30	0.19	0.849
8	50	0.43	0.29	1.46	0.149
10	50	0.35	0.22	1.58	0.119
14	50	0.17	0.24	0.74	0.464
17	50	0.037	0.30	0.12	0.901
20	50	0.24	0.23	1.04	0.305
16	50	0.10	0.32	0.31	0.758
19	50	0.29	0.18	1.65	0.106
overall physical	50	1.64	0.96	1.69	0.096
overall mental	50	1.55	1.18	1.32	0.194
overall social	50	0.85	1.18	0.72	0.475
overall global	50	0.18	0.42	0.44	0.661
overall HRQOL-20	50	4.55	3.14	1.45	0.154

Table IV
HRQOL Items Characteristics: Discriminant Validity

Items	Pre TURP patients		Control						
	Mean	SEM	Mean	SEM	Mean Difference	SEM	95% Confidence Interval Lower	Higher	p value
2	2.69	0.28	1.73	0.28	0.96	0.40	0.17	1.75	0.05
3	4.96	0.23	2.58	0.29	2.38	0.37	1.64	3.12	0.0001
4	3.46	0.22	2.18	0.29	1.28	0.36	0.56	2.00	0.001
5	7.66	0.27	4.81	0.39	2.85	0.46	1.94	3.75	0.0001
9	4.83	0.25	3.40	0.33	1.43	0.41	0.61	2.24	0.001
11	7.16	0.28	3.37	0.34	3.79	0.45	2.90	4.67	0.0001
6	3.70	0.19	2.49	0.28	1.21	0.33	0.55	1.86	0.0001
7	4.16	0.22	3.50	0.37	0.66	0.43	0.20	1.51	0.13
12	3.84	0.21	2.88	0.30	0.95	0.35	0.26	1.65	0.01
13	5.21	0.17	3.75	0.26	1.46	0.30	0.86	2.07	0.0001
15	4.17	0.23	3.16	0.32	1.01	0.38	0.25	1.76	0.01
18	4.41	0.26	3.54	0.36	0.87	0.44	0.0038	1.73	0.05
1	6.36	0.32	3.01	0.34	3.35	0.49	2.37	4.32	0.0001
8	7.34	0.27	4.24	0.42	3.09	0.50	2.10	4.09	0.0001
10	6.20	0.31	2.90	0.32	3.30	0.45	2.41	4.19	0.0001
14	4.25	0.24	2.46	0.26	1.78	0.37	1.05	2.52	0.0001
17	6.90	0.24	3.75	0.35	3.15	0.41	2.33	3.97	0.0001
20	3.71	0.21	2.34	0.25	1.37	0.33	0.72	2.03	0.0001
16	5.71	0.24	3.25	0.30	2.46	0.38	1.70	3.21	0.0001
19	3.58	0.20	2.36	0.25	1.23	0.32	0.59	1.87	0.0001
Overall physical	31.16	1.02	18.09	1.11	13.07	1.55	9.99	16.15	0.0001
Overall mental	25.67	1.00	19.55	1.35	6.12	1.65	2.85	9.39	0.0001
Overall social	34.72	1.25	18.70	1.47	16.01	1.96	12.14	19.89	0.0001
Overall global	9.27	0.36	5.61	0.49	3.66	0.60	2.48	4.84	0.0001
Overall HRQOL-20	101.00	3.06	62.01	3.50	38.98	4.75	29.59	48.38	0.0001

LUTS group exhibited significant changes. The lowest magnitude of change was observed in item 17. In contrast, none of the comparison in the treatment of the control group except for item 2 ($p<0.05$) approached significant level.

Table IV showed the discriminant validity between the surgically treated and the control group. Significant differences were observed between the two groups for all the items and domains.

Table V
Comparison of the HRQOL-20 Validated in Malaysia Versus HRQOLL-20 Validated in France

	HRQOL-20 validated in Malaysia		HRQOL-20 validated in France	
	Test-Retest Reliability (ICC)	Internal Consistency (Cronbach's alpha)	Test-Retest Reliability (ICC)	Internal Consistency (Cronbach's alpha)
Physical score	0.87	0.88	0.89	0.88
Mental score	0.90	0.90	0.88	0.96
Social score	0.89	0.90	0.83	0.78
Global score	0.74	0.74	0.87	0.81
HRQOL-20	0.91	0.91	0.78	0.95

The comparison of the reliability and validity of the Health-Related Quality Of Life (HRQOL-20) validated in the Malaysian population and Health-Related Quality Of Life (HRQOL-20) validated in France was shown in Table V.

Discussion

The HRQOL-20 was able to discriminate between patients with LUTS and those without. It is a valid and reliable tool in assessing patients with LUTS in a Malaysian population. All the items showed high levels of sensitivity and specificity. The large effect size obtained when the questionnaire was administered before and after an intervention of known efficacy (TURP) indicated a high degree of responsiveness. Similar findings were also noted in other studies[8,9].

In the present study, the correlation between the changes in the HRQOL-20 scores and global ratings of patient improved following TURP. There was a mean reduction of 19 points. These results suggest that changes in HRQOL-20 scores might reflect meaningful clinical changes and provide substantial assurance that scores obtained using the HRQOL-20 are reliable.

The reliability and validity of the Health-Related Quality Of Life (HRQOL-20) validated in the Malaysian population was consistent with scores of Health-Related Quality Of Life (HRQOL-20) validated in France[3]. The performance properties of the HRQOL-20 validated in the Malaysian population and the HRQOL-20 validated in France[3] showed that they are virtually identical with respect to their measurement properties. Although test-retest exhibited statistical significance in some items of HRQOL-20, this could be due to the fact that retesting was done at 12 weeks interval thus allowing the symptoms to be improved or worsen. In contrast, the minimal changes of symptoms would occur if test-retest were done at one week, two week or the latest one-month after the initial assessment. The twelve weeks interval was chosen for retesting because most of the patients in the TURP group would then have achieved the maximum symptoms improvement and this facilitate comparisons between the LUTS patients and those undergoing TURP at this interval.

Conclusion

The reliability and validity of the HRQOL-20 for 237 patients were tested in a sample of urological patients. The test retest exhibited good reliability. The intraclass correlation coefficient for total scores of the domains of the HRQOL-20 indicated an excellent intraclass reliability. The high Cronbach's alpha coefficient indicated that HRQOL-20 showed a good internal consistency and high discriminant validity. These findings indicated that the HRQOL-20 is a useful and accurate tool for assessing patients with LUTS in the Malaysian populations.

Acknowledgement

The authors would like to express their gratitude to the University Malaya's Research and Development Management Unit for providing the research grant for this study.

References

1. Ware WE. Conceptualizing disease impact and treatment outcomes. Cancer 1984; 53: 2316-323.
2. Barry MJ, Mulley AG, Fowler FJ, Wennberg JW. Watchful waiting vs immediate transurethral resection for symptomatic prostatism. The importance of patients' preferences. JAMA 1988; 259: 3010.
3. Lukacs B, Leplege A, Mc Carthy C and Comet, D. Construction and validation of a BPH specific health related quality of life scale, for medical outcome research studies. In Cockett ATK, Khoury S, Aso Y et al., eds, Proceedings of the Second International Consultation on Benign Prostatic Hyperplasia (BPH). Paris. SCI, 1994; 139-43.
4. Naughton MJ, Wyman JF. Quality of life in geriatric patients with lower urinary tract dysfunction. Am J Med Sci 1997; 314; 219-27.
5. Ware JE, Sherbourne, C. The MOS 36-item Short Form Health Survey (SF-36): I.Conceptual framework and item selection. Med Care 1992; 30: 473-83.
6. Cronbach LJ. Coefficient alpha and the internal structure of test. Psychometrika 1951; 16: 297.
7. Deyo RA, Dichr P, Patrick DL. Reproducibility and responsiveness of health status measures. Control Clin Trials 1991; 12 (supp 1): 142S-158S.
8. Cohen J. Statistical Power analysis for the Behavioral Sciences. New York, Academic Press, 1977; 75-105.
9. Guyatt G, Walter S, Norman G. Measuring changes overtime: assessing the usefulness of evaluative instruments. J Chronic Dis 1987; 40: 171-78.