11
Comparing the Standard EQ-5D 3L versus 5L Version for the Assessment of Health of Patients with Liver Diseases

Scalone L[1,2], Ciampichini R[2], Fagiuoli S[3], Gardini I[4], Gaeta L[3], Del Prete A[3], Magini G[3], Baldan A[3], Mantovani L[2,5]

ABSTRACT

Rationale

Preliminary investigations suggest that the 5L new version is promising as a more powerful version of the EQ-5D. However, more extensive work is encouraged to assess its properties in different sub-populations and contexts.

Objective

To assess the performance of the 5L version in a naturalistic context targeted to patients with chronic hepatic diseases.

Methods

We introduced the 5L descriptive system into a Cost-of-Illness study involving patients with different severe chronic hepatic diseases. The participants were asked to self-report in a questionnaire their QoL, socio-demographic information and productivity loss. We included in the questionnaire the Italian 5L descriptive system (previously finalized within the EQ group) followed by the socio-demographic questions to minimize the memory effect, then the 3L standard descriptive system, finally the VAS.

We test the following properties: *feasibility*, tested by computing the percentage of missing/not unique answers; *inconsistency*, estimated by computing the percentage of responses given in each 5L domain that are not expected according to responses given

1. CESP – Research Centre on Public Health, University of Milano Bicocca, Milano, Italy
2. CHARTA Foundation, Milano, Italy
3. Ospedali Riuniti, Bergamo, Italy
4. Patients' Organization - EpaC ONLUS, Italy
5. CIRFF – Centre of Pharmacoeconomics, University of Naples Federico II, Napels, Italy

Yfantopoulos J, editor.
27th Scientific Plenary Meeting of the EuroQol Group - Proceedings:213-227 © 2011 EuroQol Group

with the 3L version; *redistribution properties among consistent responses*, computed by testing the association of the mean and median VAS scores with the responses given in each profile domain; *ordinality*, tested by assessing transitivity between 5L scores, using the VAS score as reference; *convergent validity*, tested by computing the Spearman rank correlation coefficient between the scores obtained with the 3L, 5L and VAS; *discriminatory power*, quantified in terms of ceiling effect and of absolute (H') and relative informativity (J'), using the Shannon index.

Results

Data from 426 patients are analysed: 69% male, 19-84 (median=57) years old. They are affected with one or more between: chronic hepatitis C virus, chronic hepatitis B virus, cirrhosis, liver transplantation, hepatic carcinoma, NASH, and other hepatic diseases. One-hundred fifty-three patients (33.1%) described their health as 11111 with the 3L version, those reporting 11111 with the 5L version were 143 (33.6%). Median VAS was 70 (15-100). Seventeen patients (4.0%) did not complete the 5L profile. Missing answers were more frequent with the 3L version. The average proportion of inconsistency is 2.5%, ranging from 1.7% (Usual-Activities and Pain/Discomfort) to 3.2% (Anxiety/Depression). Regarding consistency, for each profile domain, the mean and median VAS scores are higher among subgroups of patients reporting better health with both the 3L and 5L version. *Ordinality* test overall shows fair results. Strong correlation coefficients are found between 3L and 5L domains, ranging from 0.83 (Self-Care) to 0.91 (Usual-Activities). The coefficient between 5L domains and VAS are moderate to strong, ranging from -0.35 (Self-Care) to -0.57 (Usual-Activities), higher than the coefficients between the 3L domains and VAS, except for Self-Care (-0.41). The ceiling effect remains similar in all but one domain (mobility). Absolute informativity is higher for the 5L version in every domain, however, relative informativity is very similar between the two versions.

Discussion

Our results show that the EQ-5D-5L is feasible and suggest that it can be a more powerful version to assess QoL in patients with chronic hepatic diseases.

11.1 INTRODUCTION

In more than 20 years of its use, the EQ-5D has shown to be valid and reliable generic instrument to assess QoL in the many different populations. However, some drawbacks that have been ascribed to the tool, such as a lacking descriptive richness, a restricted ability to discriminate small but clinical relevant different health status between and within individuals, the presence of ceiling effect, have encouraged the development and investigation of the properties of a 5L version of the EQ-5D [Pickard et al, 2007a and 2007b Janssen et al, 2008a and 2008b among the most recent publications].

The results obtained so far, conducted on a preliminary version of the 5L system, are encouraging and suggest that the 5L system is a promising improved version of the EQ-5D [Pickard et al, 2007a and 2007b Janssen et al, 2008a and 2008b among the most recent publications].

However, the research conducted so far tested the properties of two preliminary language versions (US English and Dutch) of the EQ-5D-5L and few data were available of the real individual's health status, reported by patients affected with different types of cancer [Pickard et al, 2007a and 2007b Janssen et al, 2008a and 2008b].

Accordingly, more extensive research is encouraged to be performed for a more comprehensive assessment of the 5L psychometric properties, in comparison with the standard 3L system, in the different sub-populations and contexts.

11.2 OBJECTIVE

We aimed to assess the performance of an updated 5L version, applied in a naturalistic context targeted to patients with different severe hepatic diseases.

11.3 METHODS

Study sample and setting

A sample of patients with different severe hepatic diseases were invited to participate in a Cost of Illness study carried out at the Gastroenterology Unit of the Ospedali Riuniti di Bergamo (northern part of Italy).

To participate in the study, any patient sequentially accessing at the hospital, with a diagnosis of hepatic disease was considered valid. These patients were enrolled in the study if accepted to participate after receiving information on the scope of the study, the type of information and the mode of data collection.

Instruments and data

Two data collection instruments were used in the COI study: an electronic questionnaire to be completed by the physicians, containing questions on the clinical status and on the health care resources consumption (drugs, hospitalizations, medical visits, examinations) through the six months before the enrollment, and a paper questionnaire to be completed in autonomy by each patient.

The latter questionnaire contained the following parts in this order: the Italian version of the 5L descriptive system, previously finalized within the EQ group, then some socio-demographic questions, then the Italian 3L standard descriptive system, the VAS, finally questions on productivity loss and on the use of alternative treatments for the hepatic disease, during the previous six months.

Into the questionnaire, the 5L was introduced before the 3L according to previous research findings, which revealed that when respondents scored 3L first, there can be a tendency to not use the 'in-between' levels 2 and 4 of 5L [Janssen et al, 2008a]. Also, before receiving the 3L, to minimize memory effects between the 5L and 3L, the participants were asked socio-demographic questions and then were specified that the 3L and 5L are independent ways of assessing their health state.

The participants completed by themselves this questionnaire while waiting for the visit; however a physician was available to help to comprehend the questionnaire, if requested.

Ethical issues

This study was conducted in agreement with the National Regulatory Requirements, International Conference on Harmonization Guidelines for Good Clinical Practice and the 18th World Medical Assembly [WMA, 1964] and all subsequent amendments. The study protocol was notified at the Ethical Committee of the Ospedali Riuniti di Bergamo. Each patient had to sign an informed written consent to participate.

Data analyses

The study sample is described according to demographic (age and gender), clinical characteristics (type of disease) and QoL, with both the 5L and the 3L system.

The following psychometric properties of the 5L descriptive system are tested by following, whenever possible according to our data, the same approach used by Janssen et al, 2008a in their study:

Feasibility.

This property is tested by computing the percentage of missing or not clear answers obtained with the 5L, in comparison with the standard 3L version. The percentage of blank 5L profiles, compared with the 3L profiles and the VAS is also calculated.

Inconsistency, consistency and redistribution properties

An inconsistent response can be defined as a 3L response that is at least 2 levels away from the 5L response given by the same individual. We expect that responses are consistent but also that response given with the 3L can redistribute between adjacent 5L levels that are logically consistent with the 3L ones. Validation of the 5L extension of EQ-5D is supported when the redistribution results show that respondents make sufficient use of the extra levels in 5L.

Inconsistency is estimated by computing the percentage of responses given in each 5L domain that are considered inconsistent with the responses given with the 3L version,

and by computing the size of inconsistencies, according to the weight defined for each level of inconsistency. Figure 11.1.a shows a redistribution diagram proposed for the 5L-3L comparison by Janssen et al [2008a]. The diagram shows all the potential 3L-5L response pairs, using arrows. Each arrow (gx.y) represents a projection of a 3L response (x=1, 2, 3) on a 5L response (y=1, 2, 3, 4, 5): long arrows (g1.1, g1.2, g2.2, g2.3, g2.4, g3.4, g3.5) are defined for consistent levels, as they connect 5L levels adjacent with 3L levels. The short arrows are defined for inconsistent levels, as the adjacent levels of the 5L system are skipped from the 3L one.

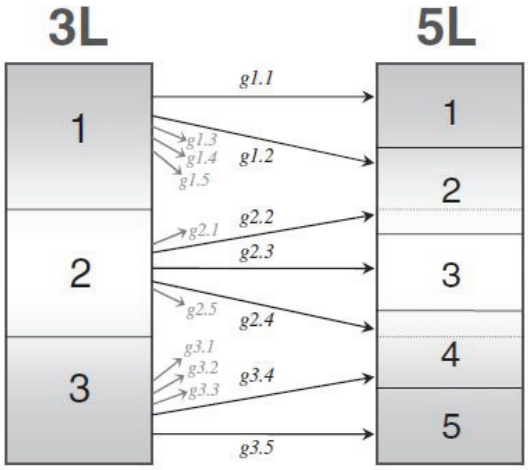

Figure 11.1a. Redistribution diagram:
expected relationship between 3L and 5L for one dimension.
Figure drawn from Pickard et al, 2008a

The weights of inconsistencies are calculated according to the approach shown in the inconsistency box (Figure 11.1.b) proposed by Janssen et al [2008a]: the 3L response scale is projected on the 5L one: hence 3L-1=5L-1, 3L-2=5L-3, 3L-3=5L-5 and weights between the two systems levels are defined as 0 if consistent (e.g., if g1.1 and g1.2 and g2.2, g2.3 and g2.4 are consistent), or 1 to 3 if inconsistent, with higher weights as long as 5L levels are farer from the expected 3L levels: e.g., g1.5 is more inconsistent than g1.3 (weight = 3 versus 1).

3L 5L	Level 1	Level 2	Level 3	Level 4	Level 5
Level 1	0	0	1	2	3
Level 2	1	0	0	0	1
Level 3	3	2	1	0	0

Figure 11.1b. **Consistency box: degree of inconsistency for each pair of responses.** Figure drawn from Pickard et al, 2008a (The dark cells represent inconsistent responses).

Consistency corresponds to the complementary proportions of inconsistency responses, i.e. the proportion of the 3L-5L response pairs within each 3L response level.

Within the consistent responses, we analyse the *redistribution properties* of the 5L compared with the 3L system, by calculating the proportions of 3L responses grouped within the corresponding 5L levels (i.e., 5L levels 1 and 2 for 3L level 1, 5L levels 2, 3 and 4 for 3L level 2, and 5L levels 4 and 5 for 3L level 3), and calculating the corresponding mean and median VAS values, for each dimension separately. We expect a decreasing mean and median VAS for each domain group.

Ordinality is tested by assessing transitivity between 5L scores of VAS dependent subgroups.

In particular, within each dimension the respondents are assigned to one out of seven 5L subgroups according to the 3L level and the corresponding VAS scores given by the patients reporting that level. In particular, the patients choosing 3L-1 are assigned to an upper or a lower subgroup, depending on whether their VAS is upper or lower to the median obtained for that subgroup. The patients choosing 3L-2 are assigned to an upper, middle or lower subgroup, depending on which tertile subgroup the individual VAS belongs. The 3L-3 is managed as the 3L-1 one. For each of the 7 subgroups created, the corresponding median and mean 5L responses are calculated: ordinality of 5L is confirmed if a weak and a strong transitivity holds for the median and mean 5L scores, respectively. Specifically, weak transitivity holds when the median 5L among lower half $\leq$ the median 5L among upper half subgroups (for 3L-1 and 3L-3 scores) or the median 5L among lower tertile $\leq$ median 5L among middle tertile $\leq$ median 5L among upper tertile subgroups (for 3L-2 scores). Strong transitivity holds when the 5L mean scores among lower half < upper half (for 3L-1 and 3L-3 scores) or the mean 5L among lower tertile < middle tertile < upper tertile (for 3L-2 scores).

Convergent validity

This property is tested to determine the extent to which the 5L descriptive system can measure the same concept than the 3L descriptive system and the VAS. The Spearman's rank correlation coefficient is applied between 5L and 3L responses, 5L and VAS and between 3L and VAS. Strength of correlation is interpreted as follows: absent ($r < 0.20$), weak ($0.20 \leq r < 0.35$), moderate ($0.35 \leq r < 0.50$), and strong ($r \geq 0.50$) [Juniper et al, 1996].

Discriminatory power (or informativity) is assessed in two ways:

(i) we compare the ceiling effect of the 5L with the 3L systems, i.e. the percentage of patients reporting at each domain "no problems". Also the frequencies of patients reporting 11111 with the two systems are compared.
(ii) we calculate the Shannon index (H') and the Shannon Evenness index (J') of informativity.

Informativity is assessed by dimension using the Shannon index formula [Janssen et al, 2008a] :

$$H' = -\sum_{i=1}^{L} p_i \log_2 p_i$$

where H' represents the absolute amount of informativity captured, L is the number of possible levels, and p_i is the proportion of the i^{th} level in the sample.
The higher the index, the more information is captured by the instrument. The upper limit of the Shannon index is reached when all levels of the instrument are evenly filled. The following formula can be applied:

$$H' \max = \log_2 L$$

which amounts to 1.58 in 3L and to 2.32 in 5L. If the number of levels increases, H' max increases accordingly but the related informativity only increases if the newly added levels are actually used.

J' allows to adjust H' by H' max and reflects the rectangularity of a distribution, i.e., the maximum level, 1, is reached when all levels are filled with the same frequency. In other words, J' indicates the informativity of the instrument relative to the number of possible levels and it is measured as follows:

$$J' = H' / H' \max$$

The percentage change in informativity between the two versions is also calculated as follows:

$$\frac{J(5L) - J(3L)}{J(5L)} * 100$$

Negative values show a loss of informativity while positive ones show gain of informativity of 5L compared to 3L.

11.4 RESULTS

Sample description

From January to May 2010, 426 patients (69% male) aged 19-84 (median= 57) years, affected with different chronic hepatic conditions were enrolled. The diagnosis reported is the one more recently developed or the last event (transplantation) occurred.

Table 11.1. Sample description according to clinical condition.

Disease Condition	Patients N (%)	Male (%)	Age Median (min-max)	Duration from diagnosis Median (min – max)
Chronic Hepatitis C	108 (25.4)	69	59 (19-80)	10.5 (0-37)
Chronic Hepatitis B	96 (22.5)	70	56 (28-76)	11.5 (1-47)
Cirrhosis	89 (20.9)	64	58 (36-84)	5.0 (1-30)
Liver transplanation	81 (19.0)	75	57 (21-71)	4.0 (0-23)
Other Hepatic Diseases*	20 (4.7)	50	49 (22-77)	3.0 (0-15)
Hepatic Carcinoma	19 (4.5)	79	64 (46-80)	3.0 (1-10)
Nash	8 (1.9)	75	54 (35-58)	1.5 (0-6)
Chronic Hepatitis C and B	5 (1.2)	60	65 (37-80)	20.0 (2-30)
TOT	426	69	57 (19-84)	6.0 (0-47)

* Sclerosis cholangitis, liver and kidney cysts, ductal hyperplasia, hypertransaminase disease, autoimmune hepatitis

As shown in Figure 11.2 the patients' reported relatively good levels of perceived health. The VAS was on average 70.4 (SD=20.3), with a median of 70. It ranges from a minimum of 15 (for 0.5% of patients) and a maximum of 100 (for 10.0% of patients).

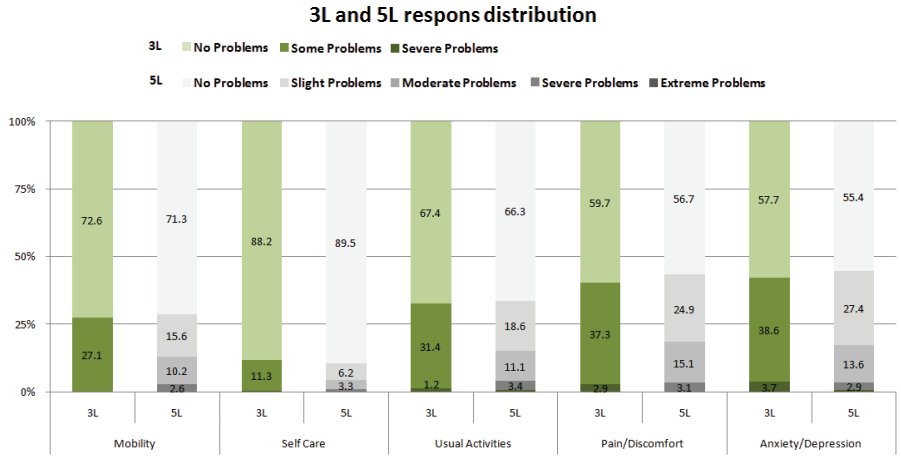

Figure 11.2. Health description with the 3L (green bars) and the 5L (grey bars) systems

Psychometric properties of the 5L compared to the 3L descriptive system

Feasibility

Four percent of the questionnaires received back, had missing data on the 5L system – in particular, 0.7% 5L descriptive systems were returned completely blank. As regards the 3L system, 6.6% of the profiles were partially incomplete (3.5%) or returned blank (3.1%). However, 11 out of 13 patients who returned the 3L system profile blank, returned the 5L system profile complete – only two persons did leave blank both the 5L and the 3L descriptive system. The VAS data was missing for 3.3% of patients but 10 out of the 14 patients not completing the VAS gave back both the 5L and the 3L profile completed.

Inconsistency

Table 11.2 shows the proportion of inconsistent responses and the average size of inconsistency for each dimension. The highest frequency of inconsistent answers were found on the Anxiety/Depression domain, followed by sel-care and then by mobility. Anxiety/depression and self-care, followed by mobility, show also the highest average sizes of inconsistencies. The overall proportion of inconsistent responses was 2.5% with a weighted average size of inconsistency equal to 1.48%.

Table 11.2. Proportion of inconsistent response pairs between 3L and 5L and the weighted average size of inconsistency for each dimension.

Dimension	N of total observations	N (%) of inconsistent observations		Average size of inconsistency weights
Mobility	408	11	(2.7)	1.63
Self Care	405	12	(3.0)	1.75
Usual Activities	402	7	(1.7)	1.13
Pain/Discomfort	404	7	(1.7)	1.13
Anxiety/Depression	407	13	(3.2)	1.75
TOT	2.026	50	(2.5)	1.48

Consistency and redistribution properties

Table 11.3 shows the mean and median VAS values for each of the 3L-5L consistent response pairs. Due to the presence of only one or none data at g3.4 or g3.5 level, in 4 domains, it was not possible to calculate the VAS mean and median in these cases. However, the available results show a decreasing value when moving from the highest (g1.1) to the lowest (g3.5) 3L-5L level, indicating that consistency is present between the 3L and 5L descriptive systems in every domain.

However, because the patients reported relatively good levels of their own QoL, a skewed distribution of frequency responses is found at the g1.1 and g1.2 levels, with the g1.1 proportion being from 92.7% (anxiety/depression) to 99.2% (self-care). Furthermore, at the 3L-2 subgroups the proportions are higher as long as the 5L levels are high, showing some unbalanced redistribution of the 5L levels through the 3L levels. Finally, no or few observations are found at the 3L-3 levels.

Ordinality

The results for the ordinality test are shown in Table 11.4. All the 5L median subgroup values show decreasing or equal values on every dimension, indicating a weak transitivity on the 3L within-group level. Strong transitivity holds for 18 of the 20 comparisons on the mean 5L scores: 18 means within the 3L groups are in descending order in every dimension when going from lower half 3L-1 to upper half 3L-3. The strong transitivity is not reached only in two cases: in the Usual-Activities domain the lower tertile of the 3L-2 subgroup is equal to the middle one (2.70), and in Pain/Discomfort, the lower half of the 3L-3 subgroup is lower (3.57) than the upper one (3.67).

Table 11.3. Redistribution properties of consistent responses, from the 3L to 5L system.

Dimension	3L	5L	Subgroup	Observations N (%)		VAS Mean*	Std.Dev.	VAS Median*
Mobility	1	1	g1.1	283	(96.3)	76.1	17.9	80
		2	g1.2	11	(3.7)	64.5	11.3	70
	2	2	g2.2	53	(51.5)	62.2	18.9	62
		3	g2.3	39	(37.9)	52.2	13.8	50
		4	g2.4	11	(10.7)	37.0	19.3	30
	3	4	g3.4	0	(0.0)	--	--	--
		5	g3.5	0	(0.0)	--	--	--
Self-Care	1	1	g1.1	355	(99.2)	73.7	18.6	77
		2	g1.2	3	(0.8)	66.7	15.3	70
	2	2	g2.2	21	(60.0)	48.0	11.9	50
		3	g2.3	11	(31.4)	47.7	15.6	49
		4	g2.4	3	(8.6)	21.7	7.6	20
	3	4	g3.4	0	(0.0)	--	--	--
		5	g3.5	0	(0.0)	--	--	--
Usual/Activities	1	1	g1.1	263	(96.3)	78.8	16	80
		2	g1.2	10	(3.7)	64.5	15.4	70
	2	2	g2.2	66	(55.0)	58.2	16.1	60
		3	g2.3	41	(34.2)	51.8	15.1	50
		4	g2.4	13	(10.8)	41.0	21.3	40
	3	4	g3.4	1	(50.0)	--	--	--
		5	g3.5	1	(50.0)	--	--	--
Pain/Discomfort	1	1	g1.1	223	(92.5)	79.0	16.4	80
		2	g1.2	18	(7.5)	74.9	18.7	70
	2	2	g2.2	84	(57.1)	66.8	16.4	70
		3	g2.3	58	(39.5)	54.7	17.3	50
		4	g2.4	5	(3.4)	43.7	7.5	40
	3	4	g3.4	8	(88.9)	28.7	8.7	30
		5	g3.5	1	(11.1)	--	--	--
Anxiety/Depression	1	1	g1.1	216	(92.7)	77.9	17.8	80
		2	g1.2	17	(7.3)	69.1	18.4	70
	2	2	g2.2	95	(62.9)	67.0	16.8	70
		3	g2.3	51	(33.8)	56.4	19.1	60
		4	g2.4	5	(3.3)	36.0	11.4	40
	3	4	g3.4	7	(70.0)	51.3	15.7	49
		5	g3.5	3	(30.0)	26.7	5.8	30

*Mean and median VAS cannot be estimated in case of only 1 or 0 observation

Table 11.4. Ordinality of the 5L descriptive system. Median and mean 5L level scores of VAS dependent subgroups.

Dimension	3L	VAS subgroups	VAS percentile	Obs N	5L median	5L mean
Mobility	1	Lower half	(0-80)	188	1	1.10
		Upper half	(81-100)	102	1	1.00
	2	Lower tertile	(0-50)	49	3	2.78
		Middle tertile	(51-65)	21	2	2.38
		Upper tertile	(66-100)	34	2	2.15
	3	Lower half	(0-50)	1	3	3.00
		Upper half	(51-100)	0	--	--
Self-Care	1	Lower half	(0-75)	175	1	1.03
		Upper half	(75-100)	174	1	1.00
	2	Lower tertile	(0-40)	18	2	2.39
		Middle tertile	(40-50)	10	2	2.30
		Upper tertile	(50-100)	13	2	2.08
	3	Lower half	(0-50)	1	3	3.00
		Upper half	(50-100)	1	1	1.00
Usual/Activities	1	Lower half	(0-80)	164	1	1.09
		Upper half	(80-100)	105	1	1.00
	2	Lower tertile	(0-50)	60	3	2.70
		Middle tertile	(51-60)	20	2	2.70
		Upper tertile	(61-100)	38	2	2.39
	3	Lower half	(0-38)	2	4	4.50
		Upper half	(39-100)	2	1	2.00
Pain/Discomfort	1	Lower half	(0-80)	146	1	1.12
		Upper half	(81-100)	91	1	1.02
	2	Lower tertile	(0-50)	56	2	2.59
		Middle tertile	(51-70)	52	2	2.35
		Upper tertile	(71-100)	36	2	2.25
	3	Lower half	(0-30)	7	4	3.57
		Upper half	(31-100)	3	4	3.67
Anxiety/Depression	1	Lower half	(0-80)	135	1	1.14
		Upper half	(81-100)	93	1	1.04
	2	Lower tertile	(0-55)	51	3	2.65
		Middle tertile	(56-70)	53	2	2.32
		Upper tertile	(71-100)	48	2	2.08
	3	Lower half	(0-40)	8	4	4.12
		Upper half	(41-100)	6	4	3.50

Convergent Validity

Table 11.5 shows a strong Spearman rank correlation coefficients between 3L and 5L version: it ranges from 0.83 (Self-Care) to 0.91 (Usual-Activities). A moderate to strong correlation coefficient is found between 5L and VAS, with a minimum of -0.35 for self-care, and a maximum of -0.57 for usual activities, in all but one (self-care) domains slightly higher than the coeeficients between the VAS and the 3L system.

Table 11.5. Spearman rank correlation coefficients for 3L, 5L and VAS by dimension.

	5L and 3L		VAS and 5L		VAS and 3L	
Dimension	Correlation coefficient	N. Patients	Correlation coefficient	N. Patients	Correlation coefficient	N. Patients
Mobility	0.87	408	-0.46	409	-0.41	396
Self-Care	0.83	405	-0.35	406	-0.41	395
Usual Activities	0.91	402	-0.57	404	-0.56	397
Pain/Discomfort	0.88	404	-0.50	404	-0.47	394
Anxiety/Depression	0.86	407	-0.44	406	-0.41	396

Discriminatory Power

With the 3L system, between 57.7% and 96.1% of patients reported 1 in every domain. This percentage ranges from 55.4 and 89.5% of patients with the 5L system. With the 3L system, 33.1% of the patients showed a 11111 profile, and with the 5L version, the percentage of patients with 11111 profile was 33.6%. Thirty-one percent of patients reported a 11111 profile with both 3L and 5L systems.

However, mobility shows that 25% less patients report no problems with the 5L than with the 3L system, and these distribute their response level between 5L-2 and 5L-3. Among the other domains, the frequency of patients reporting no problems remains similar with the two systems, although the frequency of those reporting some problems with the 3L system seems redistributed through the levels 2 and 3 with the 5L system.

Table 11.6 shows the informativity results of 3L and 5L. In terms of absolute informativity (H'), the 5L system shows higher results with an average difference from 3L system of 0.11. Relative informativity (J') indicates that both systems reach less than 20-22% of the potential informativity of the system itself. Relative informativity (J') shows also very similar results between 3L and 5L systems, with a poor improvement in discriminatory power using the 5L system.

Percentage gain in informativity ranges from -25% (Self-Care) to 0 (Mobility and Usual Activities), showing that 5L actually looses in relative informativity, compared with the 3L version.

Table 11.6. Shannon's index (H') and Shannon's Evenness index (J') values for 3L and 5L systems.

	3L			5L		
Dimension	N. Patients*	H'	J'	N. Patients*	H'	J'
Mobility	408	0.26	0.16	408	0.38	0.16
Self-Care	405	0.16	0.10	405	0.19	0.08
Usual Activities	402	0.29	0.18	402	0.43	0.18
Pain/Discomfort	404	0.34	0.21	404	0.46	0.20
Anxiety/Depression	407	0.35	0.22	407	0.48	0.21

*Number of patients responding to both 3L and 5L system

11.5 DISCUSSION

The aim of this study was to compare the performance of the standard 3L EQ-5D with the newly developed 5L version for the assessment of QoL in patients with severe chronic hepatic diseases. To perform this study we introduced the 5L new system into a paper questionnaire to be self-completed, which was given to more than 400 patients participating into a COI study. Accordingly, to the best of our knowledge, this is the first work showing the performance of the 5L system introduced into a naturalistic research context.

Overall, the patients reported, with both the 5L and the 3L systems and with the VAS, relatively good levels of QoL.

As regards its feasibility, this is very similar or slightly higher than of the 3L system, although it is important to remember that the 5L system was presented before the 3L one, which may have affected this slight difference. However, by giving the 5L system before the 3L, we have an unbiased estimated of the not low level of feasibility of this system.

The total proportion of inconsistent responses can be considered low (2.5%), indicating that the respondents were able to make a coherent use of the two systems. As a result of a relatively good level of QoL of the participants, among the consistent responses, in every profile domain the largest proportion of patients are found on the g1.1 levels, showing high levels of ceiling effects, while few or no patient reported g3.5 levels in every domain. However, despite the skewed distribution of responses, it was possible to find good level of redistribution properties, with higher mean and median VAS from the subgroup of patients reporting better QoL, compared to the one reporting worse QoL level.

Ordinality of 5L is achieved in almost all cases, although the low proportions of the most disable levels partially compromised a comprehensive testing of this property.

Convergent validity shows a strong correlation between 3L and 5L version and moderate to strong correlation of the 5L with the VAS. These results are slightly better than those obtained from the 3L domains with the VAS.

The Shannon's indexes show a modest or low gain in discrimination and informational richness associated with the 5L system, but also this property could be negatively affected by the relative good health condition reported by the sample.

The main limitations of this study are attributable to the skewed distribution of QoL levels among the participants, which causes some difficulties in testing some psychometric properties of the tool. However, the naturalistic peculiarity of the COI study conducted at the hospital level has allowed to collect the data from patients spontaneously accessing at the participating hospital.

To try to overcome this possible drawback, as the data collection within the main COI study is progressing, we plan to perform more extensive analyses on a bigger sample.

Despite these limitations, the 5L system overall shows an improved performance than the standard 3L system, for the assessment of QoL among chronically ill patients with hepatic conditions. This study results, together with the previous research ones, are encouraging and suggest to use the 5L system as a more powerful of the tool.

REFERENCES

Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the Standard EQ-5D Three-Level System with a Five-Level Version. Value in Health, 2008. 11 (2): 275-284. (a)

Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and five-level version according to two methods. Qual Life Res, 2008; 15: 463-473 (b)

Juniper EF, Guyatt GH, Jaeshke R. How to develop and validate a new health-related quality of life instrument. In: Spilker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, PA: Lippincott- Raven Publishers; 1996

Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Medical Care, 2007. 45 (3): 259-263. (a)

Pickard AS, Kohlmann T, Janssen MF, Bonsel G, Rosenbloom S, Cella David. Evaluating equivalency between response systems. Application of the rasch model to a 3-level and 5-level EQ-5D. Medical Care,2007; 45(9): 812-819 (b)

World Medical Association (WMA) Declaration of Helsinki. Adopted by the 18th WMA General Assembly, held in Helsinki in June 1964