

Assessing quality of life in children and adolescents: development and validation of the Italian version of the EQ-5D-Y

Luciana Scalone^{1,2}, Carlo Tomasetto³, Maria C. Matteucci³, Patrizia Selleri³, Serena Broccoli⁴, Barbara Pacelli⁵, Giulia Cavrini⁴

¹Research Centre on Public Health, Department of Clinical Medicine and Prevention, University of Milano – Bicocca, Italy; ²CHARTA Foundation, Milan, Italy; ³Department of Education, University of Bologna, Italy; ⁴Department of Statistics, University of Bologna, Italy; ⁵Epidemiology Unit, Local Health Authority, Bologna, Italy

Correspondence to: Luciana Scalone, Research Centre on Public Health, University of Milano - Bicocca, Department of Clinical Medicine and Prevention, Villa Serena, Via Pergolesi 33, 20052 Monza, Italy. E-mail: luciana.scalone@unimib.it

Abstract

Background: Although assessment of Health Related Quality of Life (HRQoL) in paediatric populations is gaining interest, it is not sufficiently assessed. While a number of specific paediatric instruments have been developed, some users appear to prefer generic tools such as the EQ-5D, which is a widely used and recommended tool to describe and value health across many different adult populations. We adapted the EQ-5D generic instrument into the EQ-5D-Y(youth) for the assessment of HRQoL in children and adolescents, and investigated the feasibility, acceptability, validity and reliability of this new version of the tool.

Methods: The Italian version of the EQ-5D-Y was administered to 415 children and adolescents from a general population aged between 8 and 15, and to 25 paediatric patients diagnosed with Acute Lymphoblastic Leukaemia (ALL).

Results: The Italian version of the EQ-5D-Y was found to be feasible and acceptable for self-completion in the target age-group, with less than 1% refusing to complete it and no invalid answers given. Convergent and divergent validity tested with a child specific standard instrument was satisfactory overall. The test-retest reliability was moderate to good in all the domains of the descriptive system, and the Visual Analogue Scale (VAS) showed optimal levels of reliability (Intraclass Correlation Coefficient = 0.82). As regards known-group validity, compared with the youths from general population, the ALL patients reported more difficulties in four of the five domains of the descriptive system and, on average, had a lower VAS score.

Conclusions: The Italian version of the EQ-5D-Y shows to be a promising tool for assessing HRQoL in children and adolescents from 8 to 15 years of age. Future studies should further investigate and optimize its applicability to clinical research and carry out economic evaluations within the health system.

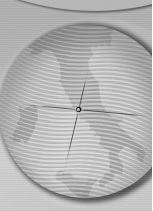
Key words: Health-Related Quality-of-Life (HRQoL), EQ-5D-Y, health, children, adolescents

Introduction

Health-Related Quality-of-Life (HRQoL) in paediatric populations is gaining particular interest in the contexts of research, health care and policy making, and a number of instruments for measuring HRQoL in youths have been developed in the last decade [1-3]. Most of the available instruments for measuring HRQoL in paediatric populations have been specifically developed for children or adolescents. However, previous research suggests that it is also possible to measure HRQoL in youths by modifying existing HRQoL instruments used for investigating HRQoL in adults, as long as the wording of instructions, items and response options are adapted to the needs of these younger respondents [2]. This latter approach has some potential advantages

for clinical research and health economic evaluations, such as facilitating the comparison of HRQoL between different age groups (e.g., adults and youths), or enabling the tracking of individual changes in HRQoL over time. It should be noted that this would be particularly beneficial in longitudinal studies investigating HRQoL in chronic or progressive conditions, or in long-term treatments that continue from childhood through to adulthood.

The EQ-5D is a brief, generic measure of HRQoL designed for use in adult populations [4]. It consists of two parts: a descriptive system including five domains – mobility, self-care, usual activities, pain or discomfort, anxiety or depression – and a Visual Analogue Scale (VAS), which quantifies general HRQoL with a score



ranging from 0 (worst imaginable health) to 100 (best imaginable health). The EQ-5D has been widely used in clinical trials, population surveys and health care policy evaluations, and it is recommended for use in cost-effectiveness analyses [5]. The EQ-5D is commonly used as a complement to other disease-specific outcome measures, its shortness of length being a major factor in its wide use in studies concerning a variety of health conditions. Interest in the assessment of HRQoL in young people has, in the last few years, led some users to also adopt the adult version of the EQ-5D to younger respondents [6,7], thus suggesting that the contents of the EQ-5D tap into aspects which are also appropriate for the assessment of children's and adolescents' health [8-11]. Accordingly, in 2006, an international taskforce was established within the EuroQoL group with the aim of developing and validating a child-friendly version of the EQ-5D to be used with children and adolescents aged 8 to 15: the EQ-5D-Y [12]. With respect to the standard adult version, the wording of the instructions, items and response options of the EQ-5D-Y were rephrased to improve comprehensibility in the age-specific target population. As a result of this multicentre international study, the EQ-5D-Y has been found to be feasible, reliable, and valid [13], and is now available in different languages from the EuroQoL Group website (<http://www.euroqol.org>). The purpose of the present paper is to contribute, within the framework of the multinational study aimed at developing the EQ-5D-Y, to the cross-cultural adaptation of the instrument. To do so, we translated, pre-tested, adjusted and finally investigated the psychometric properties of the EQ-5D-Y in Italian children and adolescents aged between 8 and 15 years.

Methods

Participants

To test the psychometric properties of the tool, we involved youths from the general population and from a population of severely ill patients.

Participants from the general population were recruited from a total of twenty-seven classes of primary and secondary schools randomly selected using a stratified sampling procedure from all eligible public schools in Northern Italy. After the schools had given their consent, the parents or guardians of all children and adolescents attending these classes were sent a letter in which their children/adolescents were invited to participate in the study. The goal was to reach a minimum target number of 200 youths enrolled, according to power calculations performed

within the International main project [12].

Regarding the ill patients, who had been diagnosed with Acute Lymphoblastic Leukaemia (ALL) and in chemotherapy maintenance treatment, they were recruited through public paediatric hospitals in two cities in the north of Italy. For ethical reasons, the patients not in maintenance therapy were considered too sick to participate in the project. With regards to sample size, because of the rareness of the disease and because only youths in maintenance chemotherapy could participate, we enrolled all the available patients that were treated during the study period in the specialist hospitals that participated in the project.

Participants of the validation study had to meet the following criteria: (a) age between 8 and 15 years, (b) absence of cognitive impairment, (c) ability to understand the Italian language.

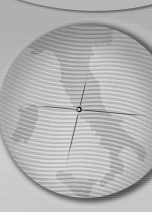
Procedure

This study was conducted in agreement with the National Regulatory Requirements, the International Conference on Harmonization Guidelines for Good Clinical Practice and the Declaration of Helsinki [14]. The protocol was approved by the boards of the participating schools and by the Ethical Committees of two hospitals. The participant and one of their parents or guardians both signed an informed written consent.

Overall, the study was conducted in 3 phases: 1) translation, 2) pre-test, and 3) validation study.

As regards phase 1), an Italian version of the EQ-5D-Y was obtained using a forward-backward-forward translation procedure according to the available guidelines [15,16]. Firstly, two native Italian speakers with expertise in quality of life, and one Italian professional translator, independently translated the questionnaire into Italian. The three versions were compared and reconciled. This preliminary Italian version was translated back into English by an independent native English speaker who was not familiar with the original version of the questionnaire. Some minor changes were made to the Italian version to solve slight discrepancies between the back-translated English version and the original.

Phase 2) followed: the resulting Italian version was pre-tested on a sample of thirty-seven children, twenty of whom were clinically healthy and seventeen who were chronically ill. The chronically ill participants were selected from patients of two types of chronic conditions: haemophilia and skin diseases (e.g., atopic dermatitis). The adapted EQ-5D-Y was satisfactorily understood by the young respondents and only minor adjustments



were made: the word "problems" was replaced by "difficulties", which was suggested by the respondents as a clearer term; with respect to the standard EQ-5D version for adults, the wording of the VAS instructions was changed as follows: "Please mark on the numbered line with ONE X the point that shows how good or bad your health is TODAY" [17].

Finally, phase 3) of the study was performed: at school the questionnaires were collectively administered to children and adolescents whose parents had provided them with informed consent. Although an assistant was available for any clarification, the participants were asked to complete the questionnaire autonomously. Ten-day delay test-retest data were collected from 129 (31%) of the participants from the general population. ALL patients were identified through the database of the medical centres involved in the study. Children and adolescents, whose parents or guardians provided them with consent, were asked to participate during inpatient hospitalization or an outpatient appointment. A member of the medical staff asked patients to autonomously complete the questionnaire.

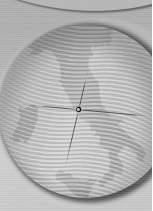
Measures

- *Socio-demographic measures.* Information about age, gender, and school grade was collected at the time of data collection. From the ALL patients, further information on co-existing health conditions, co-morbidity, and past and current treatments was assessed.
- *EQ-5D-Y.* The EQ-5D-Y consists of a descriptive system that is comprised of five items referring to mobility ('walking about'), self-care ('looking after myself'), usual activities ('doing everyday activities'), pain or discomfort ('having pain or discomfort') and anxiety or depression ('feeling worried, sad or unhappy'). Each item has three levels of problems reported (no difficulty, some difficulty, a lot of difficulty). The EQ-5D-Y also includes a vertical, graduated Visual Analogue Scale (VAS) where the respondent rates his or her overall health status on a scale from 0 to 100, with 0 representing the worst and 100 the best health state he or she can imagine. All items refer to the health state 'today'.
- *PedsQLTM.* The Italian version of the Pediatric Quality of Life Inventory Generic Core Scales (PedsQL [18]) was also used to assess children's quality of life. It consists of twenty-three items that can be grouped into four multidimensional scales: physical functioning, emotional functioning, social functioning, and school functioning. These four scales can be grouped into one summary score. Scores range from 0 to 100, with 0 meaning the worst and 100 meaning the best levels of HrQoL.
- *General Health Item.* The General Health Item [19] consists of one question regarding the individual's own general health and which has five possible response levels: excellent (coded as 5), very good (4), good (3), fair (2), poor (1).
- *Ten-points Health Score.* Finally, as schoolchildren in Italy are familiar with a school evaluation scale ranging from 0 to 10, where zero corresponds to the worst mark and 10 to the best mark, we asked participants to "Give a mark from 0 to 10 to your health TODAY".

Data analyses

Descriptive statistics of the items and average score of the VAS were calculated. The acceptability and feasibility of the EQ-5D-Y were assessed by calculating the percentage of missing values on the descriptive system and the VAS.

We tested construct validity by testing convergent and divergent validity between the items of the descriptive system of the EQ-5D-Y and the items of the PedsQL [20], using Spearman-rank's correlation coefficients. To be considered relevant for convergence, the correlation coefficients between equivalent items were required to be higher than those with items considered to be different, and in any case higher than 0.20 and statistically significant. To be considered relevant for divergence, the correlation coefficients between domains identified as being conceptually different had to be lower than those with the items considered as equivalent, and in any case lower than 0.20 [21]. The VAS was assessed for convergence by calculating Pearson's product-moment correlation coefficients with the total PedsQL score, the General Health Item and the Ten-points Health Score. Known-group validity was tested by matching one ALL patient with three participants from the general population (controls) of same age and sex, to adjust for these potential HRQoL determinants. The percentage of reported difficulties with each item of the descriptive system and the related Odds Ratios (ORs) were calculated. ORs higher than 1 correspond to a higher risk of reporting difficulties among ALL patients; ORs lower than 1 corresponds to a lower risk of reporting difficulties among ALL patients. Regarding the VAS, mean, median, minimum and maximum, floor (percentage of respondents marking 0) and ceiling (percentage of respondents marking 100) effects were compared in the 2 subgroups. Differences between the mean scores of the ALL patients and the controls were compared



using a paired *t*-test. For test-retest reliability, the percentage of agreement (i.e., the percentage of respondents giving the same answers) and Cohen's weighted Kappa were calculated for each item of the descriptive system [22,23]. To identify the strengths of agreement, we adopted the Landis and Koch standards [24]: a *K* of 0 indicates absence of agreement, from 0.01 to 0.20 slight agreement, from 0.21 to 0.40 fair agreement, from 0.41 to 0.60 moderate agreement, from 0.61 to 0.80 substantial agreement and from 0.81 to 1 almost perfect agreement. To assess the test-retest reliability of the VAS, the Intraclass Correlation Coefficient (ICC [25]) was calculated and ICC scores lower than 0.20 indicated poor agreement, those from 0.20 to 0.40 fair agreement, from 0.41 to 0.60 moderate agreement, from 0.61 to 0.80 good agreement and higher than 0.81 excellent agreement [26]. The Bland-Altman approach [27] was also adopted to identify possible bias. The method proposed by Bland & Altman consists of plotting the difference against the mean of scores for each patient: if no systematic difference exists between the first and the second test, 95% of all the differences are expected to be within 2 SD of the mean, according to the standard threshold for repeatability established by the British Standards Institute [26].

SPSS 15.0 for Windows was used for data analyses and $p < 0.05$ was considered to be the threshold value for statistical significance.

Results

Sample description

We recruited 457 children and adolescents from the general population and 25 patients with ALL in maintenance chemotherapy. Forty-two participants by the general population did not meet inclusion/exclusion criteria for the study, and hence analyses were performed on the data provided from 415 valid participants from the general population. All the 25 recruited leukaemia patients met the required criteria and were considered for the analyses.

Participants from the general population were comprised of 199 males (47.9%) and 216 females (52.1%) aged between 8 and 15 years (Median_{age} = 11.8 years). Participants from the ALL sample were 18 males (72%) and 7 females (28%) aged between 8 and 13 (Median_{age} = 9.4 years).

Descriptive statistics

The frequency of participants from the general population reporting "no", "some/a bit", or "a lot/very" difficulties with each of the five items of

the descriptive system of the EQ-5D-Y are shown in Table 1. Less than 10% of participants from the general population reported difficulties with the "walking about" and "looking after myself" items, while 39% reported "some" or "a lot of" difficulties with "having pain or discomfort" and "feeling worried, sad or unhappy", suggesting that a ceiling effect may be present. In contrast, a higher percentage of difficulties was reported by the ALL sample. Descriptive statistics of the PedsQL sub-scales and total scores of the General Health Item and of the Ten-points Health Score are also presented in Table 1.

Acceptability and feasibility

The EQ-5D-Y was found to be acceptable by the large majority of the participants. Only one participant out of 415 from the general population (0.24%) refused to self-complete the EQ-5D-Y and the set of related instruments. Some participants in hospital refused to self-complete the instrument and asked to answer the questions when read aloud by an assistant. Regarding feasibility, no missing or incomplete answers were given either to the items of the EQ-5D-Y profile or to the VAS.

Convergent and divergent validity

The correlation coefficients between the items of the EQ-5D-Y profile and the twenty-three items of the PedsQL are reported in Table 2. The items "mobility", "self-care" and "pain or discomfort" of the EQ-5D-Y showed good levels of convergent and divergent validity as higher correlation coefficients with equivalent items and lower coefficients with non equivalent items of the PedsQL were found. For "usual activities" and "anxiety and depression", most of the correlation coefficients with the PedsQL items, deemed to be equivalent, were higher than 0.20, and were also higher than those with most of the items deemed to be different.

With respect to the VAS, convergent validity was found to be moderate with the PedsQL total score ($r = 0.37$, $p < 0.001$), good with the General Health item ($r = -0.52$, $p < 0.001$), and optimal with the Ten-points Health Score ($r = 0.77$, $p < 0.001$).

Known-group validity

As shown in Table 3, compared to children from the general population (controls) the ALL patients more frequently reported difficulties with "looking after myself", "doing everyday activities", "having pain or discomfort" and "feeling worried, sad or unhappy". Except for "walking about", all the related ORs were higher

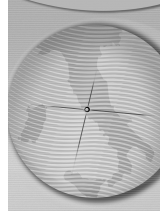


Table 1. Descriptive statistics of the HRQoL measures.

Measure	General population (N=415)	ALL patients (N=25)
EQ-5D-Y profile domains: N(%)		
1.Mobility		
No difficulty	388 (93.5)	23 (92.0)
Some difficulty	26 (6.3)	2 (8.0)
A lot of difficulty	1 (0.2)	0
2.Self-care		
No difficulty	397 (95.7)	20 (80.0)
Some difficulty	17 (4.1)	5 (20.0)
A lot of difficulty	1 (0.2)	0
3.Usual activities		
No difficulty	350 (84.3)	18 (72.0)
Some difficulty	62 (14.9)	7 (28.0)
A lot of difficulty	3 (0.7)	0
4.Pain and discomfort		
No	253 (61.0)	11 (44.0)
Some	159 (38.3)	12 (48.0)
A lot of	3 (0.7)	2 (8.0)
5.Anxiety and depression		
No	253 (61.0)	13 (52.0)
A bit	144 (34.7)	10 (40.0)
Very	18 (4.3)	2 (8.0)
EQ-5D-Y VAS: Median (min-max)	90 (30-100)	85 (0-100)
PedsQL: Median (min-max)		
Physical functioning	81.2 (37.5-100)	56.2 (10.0-71.9)
Emotional functioning	70.0 (15.0-100)	56.2 (10.0-75.0)
Social functioning	90.0 (25.0-100)	60.0 (0.0-75.0)
School functioning Summary scale	75.0 (10.0-100)	50.0 (25.0-71.7)
	78.3 (37.0-100)	54.3 (14.3-71.7)
General Health Item: N(%)		
Excellent	118 (28.5)	3 (12.0)
Very good	184 (44.3)	12 (48.0)
Good	98 (23.6)	8 (32.0)
Fair	14 (3.4)	2 (8.0)
Poor	0	0
Ten-points Health Score: Median (min-max)	9 (2-10)	9 (6-10)

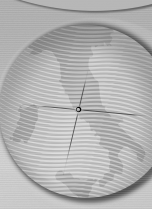
than 2, thus indicating that the probability of reporting difficulties was from twice to three times higher for the ALL patients compared to the controls. However, perhaps because of the limited size of the sample, the differences fell short of significance for “looking after myself” ($p > 0.20$), “doing everyday activities” ($p > 0.13$), and “having pain or discomfort” ($p > 0.12$). In contrast, the difference was significant for the “feeling worried, sad or unhappy”.

For the VAS, the minimum score in the ALL sample was 0, whereas the minimum score in

the controls was 50. The highest score (100 in both groups) was reported by five patients from the ALL sample (20%) and by twenty-five controls (33.3%). The mean VAS score was lower for the ALL patients respect to the controls, and the difference was statistically significant, $t(25) = 13.7$, $p < 0.001$.

Reliability

Test-retest agreement ranged from 69.8% for the item “pain or discomfort” to 95.4% for “self-care”. However, the K values appear to be



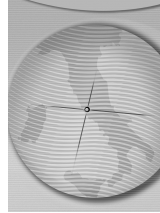
only fair to moderate (see Table 4). It should be noted, for instance, that the *K* values may be biased because of high ceiling effects and should be interpreted with caution. Regarding the VAS, the ICC value was 0.82, *p* < 0.001, thus

suggesting a good level of test-retest reliability. According to the Bland-Altman approach, the mean of the differences between pairs is close to zero (-3.25). The standard deviation of the differences (10.8) is lower than the standard

Table 2. Convergent and divergent validity of EQ-5D-Y profile compared with the PedsQL.

PedsQL items	EQ-5D-Y profile items				
	1	2	3	4	5
1.1 To walk more than one block	0.23***	0.13**	0.16**	0.09	0.14**
1.2 To run	0.22***	0.03	0.15**	0.18***	0.19***
1.3 Sports activity or exercise	0.09*	-0.06	0.15**	0.11*	0.30***
1.4 To lift something heavy	-0.06	-0.03	0.16***	0.07	0.10*
1.5 To take a bath or shower by myself	0.09	0.37***	0.21***	0.04	0.04
1.6 To do chores around the house	0.09*	0.05	0.22***	0.03	0.13**
1.7 Hurt or ache	0.09*	-0.01	0.17***	0.40***	0.27***
1.8 Low energy	0.01	-0.06	0.21***	0.19***	0.36***
2.1 Afraid or scared	0.06	0.04	0.20***	0.15**	0.30***
2.2 Sad or blue	0.08	0.00	0.17***	0.14**	0.52***
2.3 Angry	0.00	-0.03	0.10*	0.08	0.21***
2.4 Trouble sleeping	0.09*	0.08	0.15**	0.06	0.26***
2.5 Worry about what will happen to me	0.06	0.04	0.16***	0.16***	0.34***
3.1 Trouble getting along with other children	0.06	0.06	0.17***	0.04	0.14*
3.2 Other children do not want to be my friend	0.07	0.08	0.04	0.03	0.13**
3.3 Other children tease me	0.10*	0.14**	0.09	0.04	0.09
3.4 Cannot do things that other children can do	0.01	0.05	0.21***	-0.02	0.11*
3.5 Keep up when I play with other children	0.08	0.08	0.26***	0.02	0.12*
4.1 Pay attention in class	0.04	0.05	0.17***	0.16***	0.30***
4.2 Forget things	0.07	0.14**	0.19***	0.12*	0.18***
4.3 Trouble keeping up with my schoolwork	0.11*	0.07	0.22***	0.08	0.30***
4.4 Miss school because of not feeling well	0.14**	0.10*	0.10*	0.20***	0.19***
4.5 Miss school to go to doctor or hospital	0.08	-0.02	0.10*	0.12**	0.04

Note. Spearman Rank-Correlations coefficients representing the intercorrelations between the items of the EQ-5D-Y profile and the twenty-three items of the PedsQL coefficients expected to be relevant (*r*_S > 0.20 and *p* < 0.05) are reported in grey cells. Coefficients expected to be non relevant (*r*_S < 0.20) are reported in white cells. EQ-5D-Y items: 1 = walking about; 2 = looking after myself; 3 = doing everyday activities; 4 = having pain or discomfort; 5 = feeling worried, sad or unhappy.
p* < 0.05 *p* < 0.01 ****p* < 0.001.



deviation of the measurements themselves (14.8 and 13.9 in the first and second assessment, respectively), as expected. In addition, 94% of the differences between pairs are lower than 2 *SD* above or below the mean (Figure 1). Although this value is slightly lower than the standard threshold for repeatability (95%) established by the BSI (1979), these findings suggest that there was no systematic bias between the two waves of assessment.

Discussion

An Italian version of the EQ-5D-Y was developed so as to be suitable for Italian children and adolescents from 8 to 15 years old. After a standardized translation procedure [28] and an accurate pre-test, the EQ-5D-Y was tested on a sample of children and adolescents from the general population and on a clinical sample of paediatric patients. Paediatric patients were children and adolescents with Acute Lymphoblastic Leukaemia, a rare but severe and disabling condition.

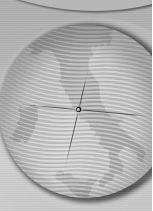
The Italian version of the EQ-5D-Y appeared to be suitable for self-completion in the target age-group, with extremely few refusals to complete the instrument, and no missing or incorrect answers. An overall small proportion of missing or inappropriate responses have been found in the validation studies of the EQ-5D-Y in other languages [2], thus supporting acceptability and feasibility for the EQ-5D-Y.

Regarding convergent and divergent validity, the Italian version of the EQ-5D-Y displayed distinct patterns of association with other comparable child-specific measures of HRQoL. For each domain of the descriptive system of the EQ-5D-Y, we identified one or more conceptually equivalent PedsQL items, and most of the PedsQL items (17 out of 23) were found to be equivalent with the EQ-5D-Y domains. These findings suggest that the EQ-5D-Y includes components of quality of life which are relevant for children/adolescents and appropriate for assessing HRQoL in paediatric populations.

Table 3. EQ-5D-Y results of ALL patients compared with youths from general population.

Measure	ALL Patients (n = 25)		Controls (n = 75)		OR (Lower/Upper 95% CI)	p-value
	n	%	n	%		
Walking about						
No difficulty	23	92.0	68	90.7	0.85 (0.16/4.36)	0.8401
Some difficulty	2	8.0	6	8.3		
A lot of difficulty	0	0	1	1.3		
Looking after myself						
No difficulty	20	8.0	67	89.3	2.09 (0.61/7.12)	0.2295
Some difficulty	5	2.0	8	1.7		
A lot of difficulty	0	0	0	0		
Doing everyday activities						
No difficulty	18	72.0	64	85.3	2.26 (0.77/6.68)	0.1329
Some difficulty	7	28.0	10	13.3		
A lot of difficulty	0	0	1	1.3		
Having pain or discomfort						
No	11	44.0	46	61.3	2.02 (0.81/5.05)	0.1295
Some	12	48.0	26	34.7		
A lot of	2	8.0	3	4.0		
Feeling worried, sad or unhappy						
Not	13	52.0	58	77.3	3.15 (1.21/8.17)*	0.0156
A bit	10	4.0	16	21.3		
Very	2	8.0	1	1.3		
VAS**						
Mean	76.9		9.6		-	-
Median	85.0		95.0			
Min-max	0-100		50-100			

Note. Frequency of reported difficulties with the EQ-5D-Y profile items by children diagnosed with Acute Lymphoblastic Leukaemia (ALL) and by children from the general population of same age and sex (controls). Each ALL patient was matched with three participants from the general population of same age and sex (controls). Odds Ratios (ORs) represent the relative risk of reporting difficulties by the ALL patients relative to the controls. Confidence Intervals (CIs) of OR values are reported in parentheses. To calculate the ORs the response levels "some/a bit" and "a lot of/very" were collapsed into a unique level (i.e., "reporting difficulties"). Paired t-test was performed to compare the mean VAS scores of participants from the two samples. The VAS difference was significant, with $p < 0.05$



As regards to the VAS, a moderate to good convergence was found between the PedsQL summary score and the General Health Item, respectively. Furthermore, an optimal convergence was found between the VAS and the Ten-points Health Score. It should be noted that this latter measure was modelled upon the school evaluation scale which is currently adopted in the Italian school system in order to make it as comprehensible as possible to our participants. Consequently, the optimal convergence found between the Ten-points Health Score and the VAS suggests that the VAS is an easily understood measure of the overall quality of life perception in school-age children and adolescents. The EQ-5D-Y also showed promising discriminatory properties between severely ill paediatric patients and youths from the general population. Compared to children and adolescents from the general population, the paediatric ALL patients reported more difficulties in four of the five items of the EQ-5D-Y descriptive system (i.e., "looking after myself", "doing everyday activities", "having pain or discomfort" and "feeling worried, sad

or unhappy"), and reported lower scores on the VAS. Overall these findings suggest that the EQ-5D-Y could be capable of discriminating between severely ill patients (such as ALL children and adolescents) and youths from the general population. However, the difference between the scores of the two groups attained the conventional level of statistical significance only for one of the items of the descriptive system (i.e., feeling worried, sad or unhappy) and for the VAS score, but not for the other four items of the descriptive system. One possible reason for this partial limitation could be that we investigated known-group validity with respect to only one clinical condition and the rareness of the disease together with a limited sample size may have contributed to this finding. Hopefully, further studies should extend these results by investigating the discriminatory properties of the EQ-5D-Y with different samples of other types of acutely or chronically ill paediatric patients.

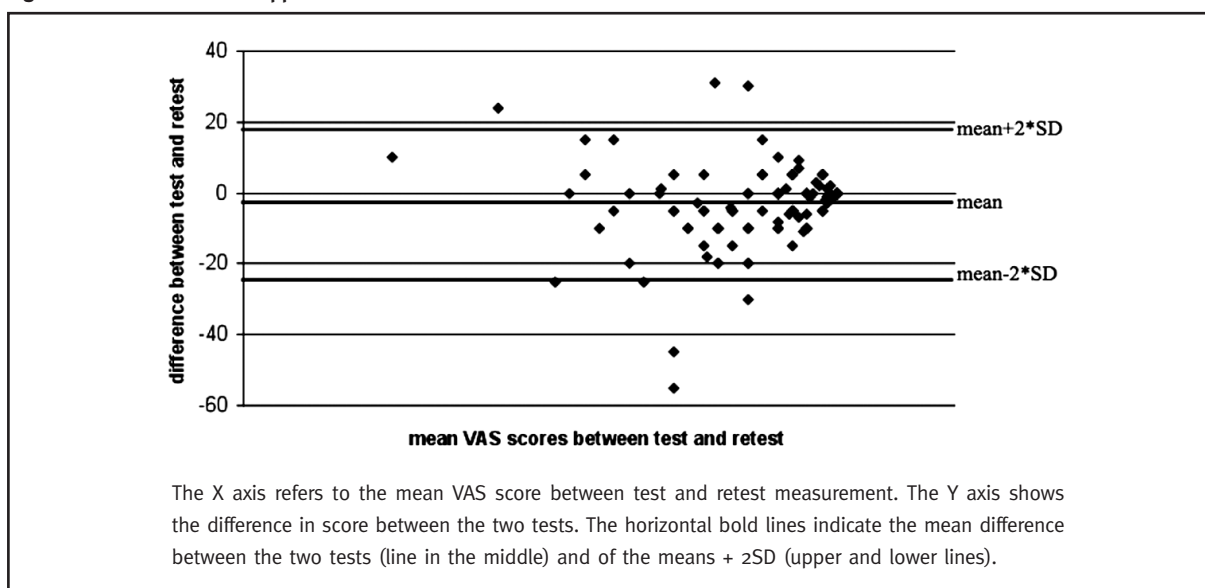
The Italian EQ-5D-Y showed acceptable test-retest reliability. All the items of the descriptive system had moderate to good levels of test-

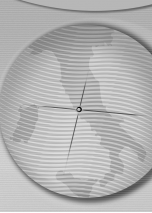
Table 4. Test-retest reliability for the EQ-5D-Y profile.

<i>EQ-5D-Y profile items</i>	<i>Agreement (%)</i>	<i>Kappa Coefficient</i>
Walking about	91.5	.22*
Looking after myself	95.4	.48**
Doing everyday activities	82.9	.35**
Having pain or discomfort	69.8	.36**
Feeling worried, sad or unhappy	76.7	.52**

* $p < .01$ ** $p < .001$

Figure 1. Bland and Altman approach results.





retest agreement, and percentages of agreement were lower than 80% with respect to only two items of the descriptive system, i.e., "having pain or discomfort" and "feeling worried, sad or unhappy". In contrast, an optimal level of stability (ICC = 0.82) was found with respect to the VAS. It should not be overlooked that the EQ-5D-Y explicitly instructs respondents to make reference to their health status at the moment of measurement (i.e., "mark one box that best describes your health *today*"), and this specification may have limited, in part, the agreement between the two measurement times.

In summary, combined with the results of other studies conducted in other countries (e.g., Jelsma, 2010 [10] and Burström et al., 2010 [29]) and clinical populations (e.g., Eidt-och et al., 2009 [30]), the present findings confirm the cross-cultural validity of the EQ-5D-Y. However, a potential limitation should also be discussed, with respect to the large ceiling effects produced in the five domains of the descriptive system. The ceiling effect was especially large among children and adolescents from the general population (up to 95.7% in the "looking after myself" item). Comparable findings were also observed in validation studies of other HRQoL instruments for children and adolescents [31,32], as well as in a validation study of the EQ-5D-Y in other languages (e.g. Burström et al., 2010 [29]). However, supporting its ability to detect health impairments, the percentage of children reporting "no difficulties" were lower in the ALL sample, suggesting that the instrument could be sensitive in detecting problems where actually present. To reduce ceiling effects, it might be useful to provide respondents with more differentiated response options. In this respect, a five level response choice of the adult version of the EQ-5D is currently in development within the EuroQol group [33]. Nonetheless, the ceiling effects imply that we should be cautious about applying this tool to the paediatric populations with mild illness severity, and further research is necessary to reduce the skewed distribution of the responses seen in this investigation.

Another goal that deserves future attention is to investigate the longitudinal stability of the instrument, as well as its responsiveness

to changes in the individuals' health status. A particularly interesting potential added value of EQ-5D-Y is related to the possibility of measuring HRQoL in longitudinal, cross-sectional and cross-cultural studies. Currently, few instruments allow these comparisons, whereas the EQ-5D-Y, and its corresponding adult version, could be an essential tool for providing a comparison of HRQoL in different groups (e.g., adults and youths), and for tracking individual changes in HRQoL over time, especially for lifelong diseases and treatments.

HRQoL has been widely assessed in several adult sub-populations in Italy (e.g., Deli et al., 2009 [34], Scalone et al., 2008 [35], Olivieri et al., 2008 [36], Scalone et al., 2006 [37]) with the EQ-5D being one of the most successfully used generic instruments in clinical and economic studies. However, HRQoL assessment is still much less frequently conducted in paediatric populations, resulting in a dramatic shortage of data available to guide health care providers and other professionals involved in policy decision-making. Furthermore, available instruments specific for age are not suitable for the computation of health utility indexes. In this regard, the EQ-5D-Y may be of particular interest for health-economic evaluations also.

This relates to the growing importance of the economic evaluation of health utilities in younger populations (e.g., Griebisch et al., 2005 [38], Ungar, 2009 [39]), and future studies should investigate the appropriateness of deriving age-specific values sets from the EQ-5D-Y.

In conclusion, despite some possible limitations, our findings support the feasibility, validity and reliability of the Italian version of EQ-5D-Y and it presents itself as a promising instrument for the assessment of HRQoL in children and adolescents from the general population or paediatric patients.

N.B. The EQ-5D-Y is a copyrighted instrument. All requests for EQ-5D-Y translations should be sent to the EuroQol Executive Office in Rotterdam, the Netherlands (userinformationservice@euroqol.org).

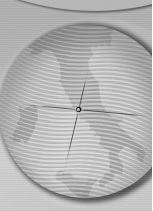
Funding

This project received funds from the EuroQol Group Foundation.

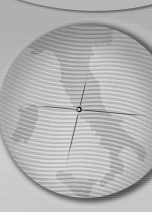
References

1) Clarke S, Eiser C. The measurement of health-related quality of life (QoL) in paediatric clinical trials: a systematic review. *Health & Health Qual Life Outcomes* 2004;2:66.

2) Ravens-Sieberer U, Erhart M, Wille N, et al. Generic Health-Related Quality of life assessment in children and adolescents: methodological considerations. *Pharmacoeconomics* 2006;24:1199-20.



- 3) Varni JW, Burwinkle TM, Lane MM. Health-related quality of life measurement in pediatric clinical practice: An appraisal and precept for future research and application. *Health Qual Life Outcomes* 2005;3:34.
- 4) Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33(5):337-43.
- 5) National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisals (Publication reference N1618). London: NICE, 2008.
- 6) Polinder S, Meerding WJ, Toet H, Mulder S, Essink-Bot ML, van Beeck EF. Prevalence and prognostic factors of disability after childhood injury. *Pediatrics* 2005;116(6):e810-7.
- 7) Sach TH, Barton GR. Interpreting parental proxy reports of (health-related) quality of life for children with unilateral cochlear implants. *Int J Pediatr Otorhinolaryngol*, 2007;71(3):435-45
- 8) Burström K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. *Qual Life Res* 2001;10:621-35.
- 9) Hennessy S, Kind P. Measuring health status in children: developing and testing a child-friendly version of EQ-5D. Proceedings of the 19th Plenary Meeting of the EuroQol Group. York, UK, 2002.
- 10) Jelsma J. A comparison of the performance of the EQ-5D and the EQ-5D-Y Health-Related Quality of Life instruments in South African children. *Int J Rehabil Res* 2010;33(2):172-7.
- 11) Stolk EA, Busschbach JJ, Vogels T. Performance of the EuroQol in children with imperforate anus. *Qual Life Res* 2000;9:29-38.
- 12) Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res* 2010;19(6):875-86.
- 13) Ravens-Sieberer U, Wille N, Badia X, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res* 2010;19(6):887-97.
- 14) World Medical Association. Declaration of Helsinki. As amended by the 59th WMA General Assembly, Seoul, Korea, 2008. Available from: <http://www.wma.net/en/30publications/10policies/b3/17c.pdf>. [Accessed on February 28, 2010].
- 15) Herdman M, Fox-Rushby J, Rabin R, Badia X, Selai C. Producing other language versions of the EQ-5D. In: Brooks RRR, de Charro F, editor. The measurement and valuation of health status using EQ-5D: A European perspective. chapter 11. Dordrecht: Kluwer Academic Publishers, 2003:183-9.
- 16) Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patients-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8(2):94-104.
- 17) Pacelli B, Cavrini G, Scalone L, et al. A comparison of different Italian CF-EQ VAS version: marking, scoring or drawing a line? Poster presented at the 23rd Plenary Meeting of the EuroQoL Group. Barcelona, Spain, September 2006.
- 18) Varni J, Seid M, Kurtin P. PedsQL 4.0: Reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001;39:800-12.
- 19) McDowell I, Newell C. Measuring Health. A Guide to Rating Scales and Questionnaires, Second Edition. Oxford University Press, 1996.
- 20) Campbell DT, Fiske DW. Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychol Bull* 1959;56:81-105.
- 21) Cohen J. Statistical power analysis for the behavioural sciences (2nd ed.). Hillsdale, NJ: Erlbaum, 1988.
- 22) Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213-20.
- 23) Krippendorff K. Reliability in content analysis: some common misconceptions and recommendations. *Human Comm Res* 2004;30:411-33.
- 24) Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- 25) Bartko JJ. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep* 1966;19:3-11.
- 26) British Standard Institution. Precision of test methods I: Guide for the determination and reproducibility for a standard test method (BS 5497, part 1). London: BSI, 1979.
- 27) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307-10.
- 28) Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Soc Sci Med* 2003;57(7): 1289-306.
- 29) Burström K, Svartengren M, Egmar AC. Testing a Swedish child-friendly pilot version of the EQ-5D instrument—initial results. *Eur J Public Health* 2011;21(2):178-83.
- 30) Eidt-Koch D, Mittendorf T, Greiner W. Cross-sectional validity of the EQ-5D-Y as a generic health outcome instrument in children and adolescents with cystic fibrosis in Germany. *BMC Pediatrics* 2009;28:9:55.
- 31) Felder-Puig R, Frey E, Proksch K, Varni JW, Gadner H, Topf R. Validation of the German version of the Pediatric Quality of Life Inventory™ (PedsQL™) in childhood cancer patients off treatment and children with epilepsy. *Qual Life Res* 2004;13:223-34.
- 32) Huguet A, Miro J. Development and Psychometric Evaluation of a Catalan Self- and Interviewer-Administered Version of the Pediatric Quality of Life Inventory™ Version 4.0. *J Pediatr Psychol* 2008;33(1):63-79.
- 33) Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the Standard EQ-5D Three-Level System with a Five-Level Version. *Value Health* 2008;11:275-84.
- 34) Deli R, Macri LA, Mannocci A, La Torre G. Measuring Quality of Life in TMD: use of SF-36. *Ital J Public Health* 2009;6(2):158-65.
- 35) Scalone L, Mantovani LG, Krol M, et al. Costs, quality of life, treatment satisfaction and compliance in patients with β -thalassaemia major undergoing iron chelation therapy: the ITHACA study. *Curr Med Res Opin* 2008;24(7):1905-17.
- 36) Olivieri I, de Portu S, Salvarani C, et al. Costs and outcomes in Psoriatic Arthritis Patients With Inadequate Response to Conventional Therapy treated with Tumor Necrosis Factor Inhibitors: the Psoriatic Arthritis Cost Evaluation (PACE) Study. *Reumatology* 2008;47(11):1664-70.
- 37) Scalone L, Mantovani LG, Mannucci PM, Gringeri A. Quality of Life is Associated to the Orthopaedic Status



in Haemophilic Patients with Inhibitors. *Haemophilia* 2006;12:154-62.

38) Griebisch I, Coast J, Brown J. Quality Adjusted Life Years lack quality in pediatric care: a critical review of

published cost-utility studies in child health. *Pediatrics* 2005;115(5): e600-14.

39) Ungar W. *Economic Evaluation in Child Health*. Oxford Scholarship Online Monographs, 2009.