AUTHORS:
Diana Khanna[1], Jyoti Khadka[1], Gang Chen[2], Christine Mpundu-Kaambwa[1], Kim Dalziel[3,4], Nancy Devlin[4], Julie Ratcliffe[1]

[1] Caring Futures Institute, Flinders University, Adelaide, South Australia, Australia
[2] Centre for Health Economics, Monash Business School, Monash University, Melbourne, Victoria, Australia
[3] Murdoch Children's Research Institute, Parkville, Melbourne, Victoria, Australia
[4] Health Economics Unit, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia

Let us talk:

A deliberative approach to improve inter-rater agreement in the assessment of Health-Related Quality of Life in children.

Abstract:

Objective: The measurement and valuation of health-related quality of life (HRQoL) in children is fundamental to quality assessment and the economic evaluation of paediatric health and social care services. This study compared the inter-rater agreement between child-self and parental-proxy HRQoL ratings (overall and domain level) and assessed the impact of a deliberation approach on improving agreement using the EQ-5D-Y-3L.

Methods: A community-based sample of child (ages 6-12 years) and parent dyads (N=85) participated in the study. For the first stage assessment, the child completed the EQ-5D-Y-3L independently of the parent who completed EQ-5D-Y-3L-Proxy 1 and Proxy 2 versions. Approximately half of the sample (N=42) dyads were invited to take part in the second stage deliberation, in which they were encouraged to discuss their responses with each other, with the option of revising their initial responses where divergences were evident. The HRQoL values were calculated using the Australian adult EQ-5D value set. The inter-rater agreement was determined using concordance correlations coefficients (CCCs) for the overall values whilst the level of agreement for the HRQoL domains were evaluated using Gwet's agreement coefficient (AC1). Child-self and proxy-rated differences in overall HRQoL across subgroups (such as child age, child gender, parent gender, child self-rated general health, presence of long-term conditions and household income) were evaluated using Wilcoxon signed-rank test.

Results: All child-parent dyads successfully completed the first stage assessment. The value-weighted EQ-5D-Y-3L profiles using Proxy versions 1 and 2 were identical, however, significantly different (diff= -2.14, p=0.02) EQ VAS scores were reported from the two proxy perspectives. No significant differences were found in the overall self and proxy values except for the subgroup with boy-children (diff= 0.02, p=0.05) and when children self-rated their general health as "very good" (diff= 0.02, p=0.02) wherein the values were underestimated by both the proxy measures. Overall, the agreement between self and proxy HRQoL ratings was poor with Proxy 1 (CCC=0.20) as well as Proxy 2 (CCC=0.17) reports. Using Proxy 1, mothers had a significant agreement (CCC=0.28, 95% CI: 0.05, 0.49) which was also higher than with fathers. When reported independently, the self-proxy concordance from the proxy-proxy perspective across domains was the lowest for "feeling worried, sad or unhappy" (AC1=0.58) followed by "having pain or discomfort" (AC1=0.68), "doing usual activities" (AC1=0.69), "looking after myself" (AC1=0.78) and highest for "walking about" (AC1=0.88). In the second stage deliberation, seventeen children and eight parents changed their responses for one or more domains. The inter-rater agreement for the deliberation sub-sample increased for the HRQoL values from fair (0.23, 95% CI: -0.07, 0.49) to significant and moderate (0.50, 95% CI: 0.24, 0.7). An improvement in agreement was also observed post deliberation across all domains.

Conclusions: This study demonstrated that the dyad deliberation approach improved agreement in child-self and proxy HRQoL reports. Further research is needed to explore child-self and proxy assessment of HRQoL and the potential for the dyad deliberation approach to improve agreement in larger and more diverse community-based samples and paediatric patient populations.

Introduction:

Cost-utility analysis (CUA) is the most prevalent form of economic evaluation, widely applied as a method for assessing the value for money of interventions and as a basis for determining resource allocation decisions across health and social care for both adults and children. In the paediatric population, child-specific health-related quality of life (HRQoL) measures are designed to directly capture quality of life across multiple dimensions (e.g., physical and psycho-social) [1]. Several child HRQoL measures have accompanied preference weights (value sets) that facilitate the calculation of QALYs. These measures have been used by regulatory bodies, for example, the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia and the National Institute for Health and Care Excellence (NICE) in England and Wales as evidence when recommending medicines eligible for government subsidies [2, 3].

Children's self-reporting capacities are supported by HRQoL measures adapted to their level of communication and understanding. Chen and Ratcliffe [3] identified nine such generic preference-based child-specific HRQoL measures for measuring and valuing HRQoL in children and adolescents, e.g., Health Utility Index Mark 2 (HUI 2), Health Utility Index Mark 3 (HUI 3), Seventeen-dimensional measure of HRQoL (17D), Child Health Utility 9 Dimensions (CHU9D), and EQ-5D Youth version (EQ-5D-Y). These child-specific HRQoL measures use a descriptive system for measuring HRQoL, in combination with a value set reflecting population-based preferences for child health states derived from the respective descriptive systems [3]. Despite the availability of such measures and accompanying value sets, in a recent review Bailey et al (2021) found that child-specific HRQoL measures were included in only four out of 62 public summary documents (PSDs) submitted to the PBAC for medicines in children. Bailey et al concluded that using child-specific HRQOL measures could have reduced decision making uncertainty in 85% of the medicines included in the review [4].

The minimum age at which children can accurately self-report their own HRQoL remains in question. Children aged 8 years and older are often considered reliable for self-reporting HRQoL measures [3, 5]. Nevertheless, studies have incorporated child-specific HRQoL measures such as the CHU9D, EQ-5D-Y and 17D in cohorts younger than 8 years old [6-9]. In a detailed investigation, Grootens-Wiegers et al. (2017) assessed the medical decision-making capacity in children based on their ability to communicate, understand, reason and appreciate. Using a neuroscientific approach, Grootens-Wiegers et al reported that children over the age of 5 years were able to communicate their choices and those over the age of 6 demonstrated the ability to remember, perform logical reasoning and appreciate using abstract thinking. However, as these traits are still maturing, a cut-off age of 12 years was recommended for overall medical decision-making competence [10]. The extent to which these more generic findings and associated recommendations in the context of medical decision making are applicable to HRQoL assessment specifically are unclear.

Although self-reported measures of HRQoL are crucial, several methodological challenges remain in assessing HRQoL in the paediatric populations. The reliability and validity of self-reports can be questioned depending on the age and cognitive capacity of the child [11]. A recent systematic review reported that children with ADHD, learning disability, speech impairments or special health care needs are associated with cognitive processing challenges, hence are more likely to have limited self-report capacity [12]. Children may also be unable to self-report if they are too young or due to episodic illness [13, 14]. In such

situations, parents, caregivers, teachers and/or health professionals may act as proxies to provide an informed estimate of the child's HRQoL on their behalf [15]. However, proxy reports, when used instead of self-reports, may be biased as adults may not have the same internal standard by which to judge HRQoL as do the children experiencing the health states [16]. Factors such as high parental stress level, the child's age and the parent's own HRQoL [13, 17-20] are known to influence parent's perception of their child's HRQoL.

There is substantial evidence to indicate that self and proxy assessed HRQoL differ, with proxy assessments typically reporting lower HRQoL than the children themselves. In addition, this discrepancy tends to be higher for the more subjective psycho-social domains (e.g., emotion or cognition) than for the domains associated with more observable attributes (e.g., mobility) of the HRQoL measures [21, 22]. While in certain situations it may be necessary for proxies, such as parents or caregivers, to report HRQoL on behalf of the child, including proxy-derived child HRQoL into CUA that estimates QALYs may be challenging in the presence of inter-rater disagreement [23]. Such discrepancies between the child's independent assessment of their HRQoL and the proxy's perception, can result in inaccurate QALYs if based only on proxy-reported outcomes [4]. Nevertheless, the potential importance of parents in their role as informants cannot be overlooked. A parent may observe a behaviour or symptom in a different context than a child, which may facilitate a better understanding of the child's HRQoL from a different perspectives [24].

A deliberative process may provide a means of strengthening the inter-rater agreement. The child and the proxy may gain a better understanding of the reasoning behind each other's HRQoL ratings after discussing their responses for each domain of the measure. Thus, deliberation entails opportunities for children and their proxies to discuss and, if possible, achieve a consented assessment of the child's HRQoL, to bridge the gap between the perceptions of each. In a study by Ungar et al, a consensus-based dyad approach was applied to children with Asthma using the HUI 2/3 measures, with promising results. When reported independently, the child-parent agreement was poor and insignificant. However, the dyad approach resulted in a moderate and significant dyad agreement [25].

To the best of our knowledge, a deliberative approach to improve child-proxy agreement has not been applied to date in a community-based sample of children from the general population using the EQ-5D-Y-3L. The aim of this study, conducted in a sample of child-parent dyads, was therefore, twofold: using the EQ-5D-Y-3L (1) To examine the level of inter-rater agreement between child self-report and parental proxy assessors in overall and domain level HRQoL, and (2) To investigate the potential impact of a deliberative method in improving concordance in child self-report and parental proxy assessed HRQoL.

Methods:
Participants and study design:
Participant recruitment was conducted through a partnership between the research team and an independent social research company (Stable Research Australia). An invitation letter outlining the details of this study was sent to an active online panel of parents to participate in research studies. Children aged 6-12 years and their parent (i.e., parent/child dyads) were eligible to participate in this cross-sectional study. A list of inclusion and exclusion criteria (Table 1) was provided to the recruitment company. To enable self-completion of the measures, only children assessed by the parents to be able to read, understand and respond to the measure were invited. Criteria for exclusion were children

diagnosed with reading disorders such as dyslexia and eye-conditions (contraindication for eye-tracking[1]) e.g., lazy eye (amblyopia), misaligned eyes (strabismus) and dancing eyes (nystagmus). The sample was selected using a stratified random sampling method to achieve broad representation of the general population in terms of socio-demographic characteristics, gender and common health conditions affecting children in general population such as asthma, anxiety disorders, conduct disorders, depressive disorders, autism spectrum disorders (ASD) and dental caries [26]. Only parents accompanied the children in the sample, and, therefore, the term "parent" in this study was limited to mothers and fathers, and did not include grandparents, extended family members or other non-parental caregivers.

Table 1: Inclusion and exclusion criteria for the sample dyads:

Inclusion criteria	Exclusion criteria
Age between 6-12 years	Children with reading disorders
Able to read and understand written language	Children with vision disorders like lazy eye (amblyopia), misaligned eyes (strabismus) and dancing eyes (nystagmus)

Written informed consent was obtained from each parent willing to participate, and a verbal assent was obtained from children. The study was conducted in South Australia and complied with the ethical guidelines of the Flinders University's Human Research Ethics Committee (Project ID 4178).

Procedure:
The child-parent dyad was invited to attend a semi-structured, face-to face interview with a trained researcher at Flinders University (Tonsley campus), South Australia. The interview consisted of two-stages. In the first stage, the child self-completed their own HRQoL assessment using the EQ-5D-Y-3L measure and a self-rated general health (SRH) item question- "In general, would you say your health is poor, fair, good, very good, or excellent?" [27], administered online via the REDCap software using a laptop computer enabled with eye-tracking technology. Following completion of the measure, the child was requested by the interviewer to retrospectively think aloud and explain their responses to the EQ-5D-Y-3L measure which was recorded.

Simultaneously, the parent was asked to complete hard copy proxy versions of EQ-5D-Y-3L (Proxy version 1 and Proxy version 2)[2] while using noise cancelling headphones such that their responses were not unduly influenced by any conversations taking place between the interviewer and the child. The parent first completed the Proxy 1 version followed by an assessment of their own HRQoL using the EQ-5D-3L. Next in sequence was the Proxy 2 measure to allow the parent time to switch between the perspectives of the two proxy versions. The parents also completed a general health SRH item, and a socio-demographic

---

[1] Eye-tracking using a screen-based eye-tracker alongside retrospective think aloud was employed for response processes validation in child-self report. This data will be discussed as the subject of a separate paper.
[2] The HRQoL questionnaires the parent completed included the proxy versions of CHU9D, PedsQL, EQ-5D-Y-3L Proxy version 1, EQ-5D-3L and EQ-5D-Y-3L Proxy version 2, in that order. The sample was divided into two groups based on the child's age: 6-7 years and 8-12 years. The order was switched between the CHU9D and the EQ-5D-Y-3L Proxy version 1 for every alternate parent participant within each group to control order effects. The data from CHU9D and PedsQL reports will be discussed in a separate paper.

questionnaire including age, gender, post code, household income, the child's long-term health condition/s if any, how well they thought they knew their child and how many waking hours they had spent with their child in the 24 hours prior to the interview.

Half[3] of the sample were assigned the EQ-5D-Y-3L measure for the deliberation task facilitated by the interviewer using only the self-completion data form the child and the Proxy version 1 data from the parent to reduce the cognitive load on the participants. The dyads were encouraged to discuss their responses, with the option of revising their initial responses where divergences were evident. This conversation was also recorded. The level of agreement was assessed for the pre- and post-deliberation responses in this sub-sample.

Measures:
The EQ-5D-Y-3L and its Proxy versions 1 and 2 were used to examine inter-rater agreement between self and proxy reporting of HRQoL. EQ-5D-Y-3L is an adapted version of the EQ-5D-3L designed for children aged 8 to 15 [28]. For proxies, in version 1, the proxy is asked to rate their child's HRQoL according to their opinion (proxy-proxy perspective), whilst in version 2, they are asked how the child would rate their own HRQoL if they were able to do so (proxy-child perspective). There are five dimensions within the EQ-5D- Y-3L-and its Proxy versions: " walking about", "looking after myself", "doing usual activities", "having pain or discomfort", and "feeling worried, sad or unhappy". For each dimension, the respondent can indicate severity on any of three levels of problems (no problems, some problems, a lot of problems). Furthermore, both the EQ-5D-Y-3L and its Proxy version also included a visual analogue scale (EQ VAS), with the respective perspectives, where the respondent can rate their child's overall health status (or that of their own when self-reporting) on a scale from 0 to 100, with 0 indicating the worst possible state and 100 the best possible state.

A value set for Australia does not currently exist for the EQ-5D-Y-3L[4]. We have therefore used the Australian adult value set for the EQ-5D-3L [29]. This value set was applied to both proxy and self HRQoL ratings. We acknowledge that value sets for adult EQ-5D-3L are known to have different properties than value sets for EQ-5D-Y-3L e.g., in terms of dimension ordering and length of value scale. However, arguably, such differences are of lesser importance here, since our purpose is to determine whether there is agreement between the dyads. The robustness of the main findings to the choice of value set was tested in sensitivity analysis using a recently published EQ-5D-Y-3L value set (for Germany) [30].

Statistical analysis:
For the first stage assessment, the differences in self and proxy reported overall and EQ VAS scores based on sociodemographic characteristics were explored. Following this, the inter-rater agreement for EQ-5D-Y-3L (Proxy 1 and Proxy 2) values and domain level HRQoL for the overall sample and as a function of subgroups were estimated. Before and after deliberation concordance between the child-parent dyads for EQ-5D-Y-3L (Proxy 1) values were estimated for the second stage deliberation sub-sample.

The self and proxy differences in HRQoL ratings were assessed for 1) age-group: children 6-7 years old, 8-10 years old and 11-age group, 2) child gender: girl-child and boy-child, 3) parent gender: mother and father, 4) child SRH: "excellent", "very good", "good" , "fair" and

---

[3] For the other half, the CHU9D measure was used, and the results will be reported in a separate paper.
[4] An EQ-5D-Y-3L value set is currently underway; results from modelling DCE data from that study are reported in a separate paper at this Plenary (Pan et al 2022).

"poor", 5) presence of long-term condition: yes and no, 4) household income: low-to-middle income (income less than $75,000) and high income (income greater than $75,000) [31] and 5) how well the parents thought they knew their child: "extremely well" and "very well or well".

Inter-rater agreement was assessed using concordance correlation coefficient (CCC) for the overall HRQoL. Since most of the participants were in relatively good health, the HRQoL scores were left skewed. CCC measures the absolute agreement between two raters and requires a minimum sample size of 10 performs. Hence, sub-groups with a sample size of less than ten were excluded from the analysis. Also, CCC does not rely on the analysis of variance (ANOVA) model assumptions [32]. Gwet's agreement coefficient (AC1) was used to analyse the domain level HRQoL. Both CCC and Gwet's AC1 take values between -1 and 1 and their magnitude was qualified using the Altman's scale for consistency of interpretation. The Altman's scale is defined as poor, fair, moderate, good and very good for values less than or equal to 0.2, 0.4, 0.6, 0.8 and 1 respectively [33]. Mann-Whitney U test was used to explore statistical differences across subgroups. A paired t-test was used to test the difference in agreement coefficients before and after deliberation [34]. In this study, the statistical significance level was set at 0.05.

Results:
Sample characteristics:

89 dyads met the inclusion criteria and were invited to participate in the study. Of these, a total of 85 dyads agreed and participated in the interview (response rate 96%). The median age of children in this sample was 9 (IQR=4), with a slight overrepresentation of girls (56%). Parents in the sample had a median age of 41 (IQR=9) and one-fifth of the dyads were father-child pairs. The majority of the children were healthy, with 31% having one of the conditions as reported by their parents: asthma (42%), autism spectrum disorder (ASD) (8%), dental caries (15%), attention deficit hyperactivity disorder (ADHD) (4%), anxiety/depression (15%), sleep problems (12%) and congenital heart disease (4%). On average, mothers reported to have spent more waking hours with their children in the 24 hours prior to the interview than fathers (diff=3.1, p-value=0.03). Approximately 23% of the respondents belonged to households with middle-to-low incomes.

First stage assessment of agreement:
Dyad EQ-5D-Y-3L value-weighted profiles and EQ VAS scores:
Table 2 describes the EQ-5D-Y-3L (Proxy 1 and Proxy 2) values and EQ VAS scores of the dyad sample based on the socio-demographic characteristics. Of the 85 dyad participants, two children did not have the EQ VAS scores and one parent did not complete the EQ-5D-Y-3L Proxy 1 measure. The mean child-self-reported utility of 0.84 (sd= 0.12) was marginally lower compared to the values derived using the proxy measures. The mean utility using Proxy 1 and Proxy 2 was 0.84 (Proxy 1: sd=0.09, Proxy 2: sd=0.1), which was almost identical. Using the Proxy 1 measure, although parents overestimated their children's HRQoL, this difference (-0.005) was not statistically significant (p-value=0.41). A similar pattern was observed in the EQ VAS scores, in which proxies consistently overestimated the health of their child from both a proxy-proxy perspective (mean= 88.39, sd= 9.82) and a proxy-child perspective (mean= 90.53, sd= 9.65) relative to child self-assessment (mean= 85.27, sd= 14.01). This difference between child-self and proxy EQ VAS scores was significant (p=0.02) when reported using the Proxy version 2.

Across the subgroups, the only statistically significant difference between child-self and proxy reported values were observed for boy-children (mean diff=0.02, p-value=0.08) and those who rated their health as "very good" on SRH item (mean diff=0.02, p-value=0.02). In both the cases the proxy reported values were lower than self-reported values. The EQ VAS score was significantly overestimated by father-proxies and when children reported "good" SRH using both Proxy 1 and Proxy 2. Parents with older children in the age-group of 11-12 years and belonging to a low-to-middle income household also reported a significantly higher EQ VAS scores using Proxy 2.

The EQ-5D-Y-3L Proxy 1 and Proxy 2 versions yielded similar value-weighted profiles. Statistical test revealed no significant differences in the overall HRQoL values (p = 0.73). However, the EQ VAS Proxy 1 scores were significantly lower than Proxy 2 scores (diff= -2.1, p-value = 0.02).

Table 2: Description of EQ-5D-Y-3L (proxy 1 and proxy 2) values based on socio-demographic characteristics

Characteristics	SELF RATED SCORES			PROXY RATED SCORES			SELF-PROXY DIFFERENCE	
	N (%)	Mean (SD)	Median (IQR)	N (%)	Mean (SD)	Median (IQR)	Mean difference (SD) Self-Proxy	p-value (MWU)
Overall								
EQ-5D-Y-3L (Proxy 1)	85	0.84 (0.12)	0.91 (0.09)	84	0.85 (0.09)	0.84 (0.09)	-0.005 (0.13)	0.40
EQ VAS (Proxy 1)	83	85.27 (14.01)	90 (23)	85	88.39 (9.82)	90 (10)	-3.12 (17.08)	0.40
EQ-5D-Y-3L (Proxy 2)				85	0.85 (0.1)	0.91 (0.09)	-0.004 (0.14)	0.64
EQ-VAS (Proxy 2)				85	90.53 (9.65)	90 (10)	-5.1 (15.50)	0.02
Age group								
6 to 7 yrs.:								
EQ-5D-Y-3L (Proxy 1)	22 (0.26)	0.83 (0.14)	0.86 (.10)	23 (0.27)	0.84 (0.11)	0.91 (0.09)	-0.02 (0.17)	0.52
EQ VAS (Proxy 1)	22 (0.26)	90.04 (14.1)	96 (13)	23 (0.27)	91.48 (7.43)	95 (10)	-1.5 (16.86)	0.52
EQ-5D-Y-3L (Proxy 2)				23 (0.27)	0.85 (0.07)	0.83 (0.09)	-0.03 (0.15)	0.64
EQ VAS (Proxy 2)				23 (0.27)	94.3 (6.05)	95 (10)	-4.23 (15.316)	0.53
8 to 10 yrs.:								
EQ-5D-Y-3L (Proxy 1)	30 (0.35)	0.83 (0.13)	0.91 (0.13)	30 (0.35)	0.84 (0.11)	0.83 (0.09)	-0.01 (0.14)	0.65
EQ VAS (Proxy 1)	29 (0.34)	84.21 (15.51)	89 (23)	30 (0.35)	87.5 (9.92)	90 (13)	-3.31 (19.78)	0.71
EQ-5D-Y-3L (Proxy 2)				30 (0.35)	0.84 (0.11)	0.91 (0.09)	-0.01 (0.14)	0.86
EQ VAS (Proxy 2)				30 (0.35)	87.33 (11.87)	90 (15)	-2.69 (17.21)	0.49
11 to 12 yrs.:								
EQ-5D-Y-3L (Proxy 1)	32 (0.38)	0.87 (0.08)	0.91 (0.08)	31 (0.36)	0.86 (0.06)	0.86 (0.09)	0.01 (0.09)	0.11
EQ VAS (Proxy 1)	32 (0.38)	82.94 (12.02)	85 (20.5)	32 (0.38)	87 (10.98)	90 (17.5)	-4.06 (14.92)	0.19
EQ-5D-Y-3L (Proxy 2)				32 (0.38)	0.85 (0.1)	0.91 (0.09)	0.02 (0.13)	0.27
EQ VAS (Proxy 2)				32 (0.38)	90.81 (8.61)	90 (7.5)	-7.87 (13.97)	0.01

Gender child

Girl-child:

EQ-5D-Y-3L (Proxy 1)	47 (0.55)	0.82 (0.14)	0.86 (0.13)	47 (0.55)	0.84 (0.11)	0.91 (0.09)	-0.02 (0.16)	0.69
EQ VAS (Proxy 1)	47 (0.55)	85.85 (14.71)	92 (25)	47 (0.55)	87.96 (9.84)	90 (13)	-2.11 (16.94)	0.98
EQ-5D-Y-3L (Proxy 2)				47 (0.55)	0.84 (0.11)	0.91 (0.09)	-0.02 (0.17)	0.71
EQ VAS (Proxy 2)				47 (0.55)	91 (9.48)	90 (15)	-5.15 (15.45)	0.14

Boy-child:

EQ-5D-Y-3L (Proxy 1)	37 (0.44)	0.87 (0.06)	0.91 (0.05)	36 (0.42)	0.85 (0.07)	0.84 (0.09)	0.02 (0.07)	0.05
EQ VAS (Proxy 1)	36 (0.42)	84.5 (13.21)	89 (17.5)	37 (0.44)	88.89 (10.05)	90 (10)	-4.44 (17.4)	0.16
EQ-5D-Y-3L (Proxy 2)				37 (0.44)	0.85 (0.08)	0.91 (0.09)	0.02 (0.08)	0.23
EQ VAS (Proxy 2)				37 (0.44)	89.81 (10.05)	90 (5)	-5.03 (15.79)	0.07

Gender parent

Mother:

EQ-5D-Y-3L (Proxy 1)	68 (0.8)	0.85 (0.10)	0.91 (0.08)	67 (0.79)	0.84 (0.1)	0.83 (0.09)	0.01 (0.12)	0.14
EQ VAS (Proxy 1)	67 (0.79)	85.78 (14.51)	90 (23)	68 (0.8)	87.5 (9.75)	90 (13.5)	-1.73 (17.25)	0.88
EQ-5D-Y-3L (Proxy 2)				68 (0.8)	0.84 (0.1)	0.83 (0.09)	0.01 (0.13)	0.22
EQ VAS (Proxy 2)				68 (0.8)	89.87 (9.65)	90 (15)	-3.94 (15.8)	0.20

Father:

EQ-5D-Y-3L (Proxy 1)	15 (0.18)	0.83 (0.16)	0.86 (0.08)	16 (0.19)	0.87 (0.07)	0.91 (0.09)	-0.05 (0.17)	0.45
EQ VAS (Proxy 1)	15 (0.18)	83.73 (12.02)	85 (19)	16 (0.19)	91.44 (9.68)	95 (6.5)	-7.8 (15.31)	0.05
EQ-5D-Y-3L (Proxy 2)				16 (0.19)	0.87 (0.07)	0.91 (0.08)	-0.056 (0.17)	0.25
EQ VAS (Proxy 2)				16 (0.19)	92.75 (9.61)	95 (9.5)	-8.87 (13.34)	0.02

Child self-reported general health
Excellent:

EQ-5D-Y-3L (Proxy 1)	20 (0.24)	0.84 (0.14)	0.91 (0.09)	19 (0.22)	0.85 (0.1)	0.91 (0.09)	-0.02 (0.17)	0.74
EQ VAS (Proxy 1)	20 (0.24)	92.95 (9.20)	96 (10)	20 (0.24)	92.95 (7.07)	95 (10)	0 (10.66)	0.70
EQ-5D-Y-3L (Proxy 2)				20 (0.24)	0.86 (0.06)	0.91 (0.09)	-0.02 (0.14)	0.95
EQ VAS (Proxy 2)				20 (0.24)	93.65 (7.51)	96.5 (10)	-0.7 (13.06)	0.75

Very good:

EQ-5D-Y-3L (Proxy 1)	44 (0.52)	0.87 (0.07)	0.91 (0.08)	44 (0.52)	0.85 (0.07)	0.83 (0.09)	0.02 (0.09)	0.02
EQ VAS (Proxy 1)	43 (0.51)	85.9 (12.04)	89 (20)	44 (0.52)	87.5 (9.77)	90 (13)	-1.60 (16.59)	0.73
EQ-5D-Y-3L (Proxy 2)				44 (0.52)	0.85 (0.1)	0.91 (0.09)	0.03 (0.12)	0.19
EQ VAS (Proxy 2)				44 (0.52)	91 (7.87)	90 (7)	-4.88 (14.75)	0.07

Good:

EQ-5D-Y-3L (Proxy 1)	16 (0.19)	0.81 (0.16)	0.86 (0.14)	16 (0.19)	0.84 (0.14)	0.91 (0.09)	-0.03 (0.18)	0.41

	n (SD)	Mean (SD)	Median (IQR)	n (SD)	Mean (SD)	Median (IQR)	Difference (SD)	p-value
EQ VAS (Proxy 1)	16 (0.19)	74.94 (17.53)	74 (23)	16 (0.19)	88.25 (10.4)	90 (10)	-13.31 (21.39)	0.02
EQ-5D-Y-3L (Proxy 2)				16 (0.19)	0.83 (0.13)	0.87 (0.09)	-0.03 (0.18)	0.61
EQ VAS (Proxy 2)				16 (0.19)	89.25 (10.79)	90 (8.5)	-14.31 (15.96)	0.02
Fair:								
EQ-5D-Y-3L (Proxy 1)	3 (0.04)	0.68 (0.05)	0.68 (0.10)	3 (0.04)	0.75 (0.07)	0.72 (0.13)	-0.06 (0.09)	0.27
EQ VAS (Proxy 1)	3 (0.04)	78.33 (18.93)	70 (35)	3 (0.04)	70.67 (4.04)	70 (8)	7.67 (22.50)	1
EQ-5D-Y-3L (Proxy 2)				3 (0.04)	0.77 (0.05)	0.74 (0.08)	-0.08 (0.02)	0.10
EQ VAS (Proxy 2)				3 (0.04)	68.33 (17.56)	70 (35)	10 (22.91)	0.66
Long-term condition								
No:								
EQ-5D-Y-3L (Proxy 1)	58 (0.68)	0.85 (0.11)	0.91 (0.08)	58 (0.68)	0.86 (0.06)	0.91 (0.09)	-0.01 (0.11)	0.73
EQ VAS (Proxy 1)	58 (0.68)	85.76 (13.54)	89.5 (23)	59 (0.69)	88.69 (9.65)	90 (10)	-2.91 (17.18)	0.59
EQ-5D-Y-3L (Proxy 2)				59 (0.69)	0.86 (0.07)	0.91 (0.09)	-0.02 (0.12)	0.83
EQ VAS (Proxy 2)				59 (0.69)	91.27 (7.72)	90 (12)	-5.45 (15.86)	0.06
Yes:								
EQ-5D-Y-3L (Proxy 1)	26 (0.31)	0.83 (0.13)	0.91 (0.13)	26 (0.31)	0.81 (0.13)	0.82 (0.17)	0.02 (0.17)	0.32
EQ VAS (Proxy 1)	25 (0.29)	84.12 (15.27)	90 (20)	26 (0.31)	87.69 (10.38)	90 (14)	-3.6 (17.17)	0.51
EQ-5D-Y-3L (Proxy 2)				26 (0.31)	0.8 (0.14)	0.82 (0.17)	0.02 (0.18)	0.30
EQ VAS (Proxy 2)				26 (0.31)	88.85 (13.06)	90 (10)	-4.28 (14.91)	0.21
Household income								
Less than $75,000:								
EQ-5D-Y-3L (Proxy 1)	18 (0.21)	0.83 (0.09)	0.83 (0.15)	18 (0.21)	0.81 (0.13)	0.82 (0.1)	0.02 (0.14)	0.47
EQ VAS (Proxy 1)	18 (0.21)	79.89 (16.57)	83.5 (20)	18 (0.21)	87.89 (10.63)	87.5 (15)	-8 (19.14)	0.12
EQ-5D-Y-3L (Proxy 2)				18 (0.21)	0.8 (0.16)	0.83 (0.17)	0.03 (0.18)	0.88
EQ VAS (Proxy 2)				18 (0.21)	91.22 (8.89)	90 (15)	-11.33 (15.11)	0.02
Greater than $75,000:								
EQ-5D-Y-3L (Proxy 1)	60 (0.71)	0.85 (0.12)	0.91 (0.08)	60 (0.71)	0.86 (0.08)	0.82 (0.14)	-0.005 (0.13)	0.76
EQ VAS (Proxy 1)	59 (0.69)	86.78 (13.04)	91 (22)	60 (0.71)	88.07 (10.05)	92.5 (14)	-1.30 (16.65)	0.93
EQ-5D-Y-3L (Proxy 2)				60 (0.71)	0.86 (0.06)	0.87 (0.14)	-0.01 (0.12)	0.45
EQ VAS (Proxy 2)				60 (0.71)	90.05 (10.36)	92.5 (12.5)	-3.10 (15.26)	0.24
Know								
Very well or well								
EQ-5D-Y-3L (Proxy 1)	39 (0.46)	0.83 (0.13)	0.86 (0.1)	39 (0.46)	0.82 (0.1)	0.82 (0.09)	0.003 (0.15)	0.24
EQ VAS (Proxy 1)	38 (0.45)	81.24 (15.46)	78.5 (25)	39 (0.46)	86.79 (10.53)	90 (15)	-5.47 (19.02)	0.20
EQ-5D-Y-3L (Proxy 2)				39 (0.46)	0.82 (0.12)	0.82 (0.09)	0.01 (0.17)	0.40

EQ VAS (Proxy 2)				39 (0.46)	88.41 (11.19)	90 (10)	-7 (17.29)	0.08
Extremely well								
EQ-5D-Y-3L (Proxy 1)	45 (0.53)	0.86 (0.11)	0.91 (0.08)	44 (0.52)	0.86 (0.08)	0.91 (0.09)	-0.01 (0.12)	0.89
EQ VAS (Proxy 1)	44 (0.52)	89 (11.7)	91 (15)	45 (0.53)	89.51 (9.05)	90 (10)	-0.55 (14.86)	0.77
EQ-5D-Y-3L (Proxy 2)				45 (0.53)	0.87 (0.06)	0.91 (0.09)	-0.01 (0.11)	0.98
EQ VAS (Proxy 2)				45 (0.53)	92.16 (7.81)	90 (10)	-2.98 (13.53)	0.23

Dyad agreement for EQ-5D-Y-3L (Proxy 1 and Proxy 2) values:

The agreement using concordance correlation coefficient between self and proxy values using EQ-5D-Y-3L was slightly higher for Proxy 1 (0.20) as compared to Proxy 2 (0.17) (Table 3). A fair and significant agreement was observed between mothers and children (0.28) using Proxy 1. In contrast, the agreement between father-child dyads was negligible using both Proxy 1 (0.06) and Proxy 2 (0.09). Within each of the following subgroups, a higher dyad agreement was consistently observed using both the proxy versions: children aged 8-10 years, boys, and children with "good" SRH. Also, children with "very good" SRH consistently reported lowest agreement with both Proxy 1 and Proxy 2. However, no such pattern was observed for children with a long-term condition, low-to-middle income household and parents who reported they knew their children well/very well.

Table 3: Overall and subgroup agreement between self and proxy values using EQ-5D-Y-3L (Proxy 1 and Proxy 2)

		EQ-5D-Y-3L (PROXY 1)		EQ-5D-Y-3L (PROXY 2)	
	No. of targets	CCC	95% CI	CCC	95% CI
Overall	84	0.20	0, 0.39	0.17	-0.04, 0.37
Subgroups					
Age group					
6-7 yrs	23	0.11	-0.29, 0.48	0.14	-0.2, 0.44
8-10 yrs	30	0.27	-0.09, 0.56	0.29	-0.07, 0.58
11-12 yrs	31	0.16	-0.18, 0.47	-0.01	-0.34, 0.32
Child gender					
Girl-child	47	0.17	-0.1, 0.43	0.14	-0.14, 0.4
Boy-child	36	0.29	-0.02, 0.54	0.23	-0.08, 0.49
Parent gender					
Mother	67	0.28	0.05, 0.49	0.23	-0.01, 0.44
Father	16	0.06	-0.29, 0.4	0.09	-0.26, 0.42
Child-Self rated general health					
Excellent	19	0.11	-0.32, 0.51	0.16	-0.15, 0.44
Very good	44	0.05	-0.23, 0.33	0.01	-0.25, 0.28
Good	16	0.26	-0.24, 0.65	0.24	-0.26, 0.63
Fair	3	0.02	-0.79, 0.8	0.30	-0.36, 0.76
Long term condition					
No	58	0.16	-0.05, 0.36	0.17	-0.06, 0.37
Yes	26	0.21	-0.18, 0.54	0.14	-0.25, 0.49
Household income					
Less than $75,000:	18	0.22	-0.23, 0.59	0.05	-0.35, 0.43
Greater than $75,000:	60	0.19	-0.04, 0.41	0.23	0.03, 0.41

Know					
Very well or well	39	0.20	-0.11, 0.47	0.13	-0.19, 0.42
Extremely well	45	0.18	-0.11, 0.44	0.20	-0.06, 0.43

Dyad agreement for EQ-5D-Y-3L (Proxy 1 and Proxy 2) domains:

Table 4 presents the dyad agreement across domains for both proxy versions of EQ-5D-Y-3L along with the 95% confidence intervals. The agreement ranged between good and very good for all domains except the "feeling worried, sad or unhappy" wherein a moderate agreement was seen using Proxy 1. The highest level of agreement was reported in the "walking about" domain (Proxy 1= 0.88, Proxy 2= 0.86). The agreement was similar for all domains using the two proxy versions. However, for the "feeling worried, sad or unhappy" domain Proxy 2 provided a higher agreement estimate than Proxy 1.

Table 4: Overall domain-level agreement for EQ-5D-Y-3L (Proxy 1 and Proxy 2)

	Domains									
	Walking about		Looking after myself		Doing usual activities		Having pain/discomfort		Feeling worried, sad or unhappy	
	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI
Overall (N=85) EQ-5D-Y-3L (Proxy 1)	0.88	0.8, 0.95	0.78	0.68, 0.88	0.69	0.57, 0.81	0.68	0.55, 0.8	0.58	0.43, 0.72
EQ-5D-Y-3L (Proxy 2)	0.86	0.78, 0.94	0.81	0.71, 0.91	0.74	0.62, 0.85	0.64	0.51, 0.77	0.70	0.58, 0.82

Altman's scale interpretation: Less than or equal to 0.2=Poor, between 0.21 & 0.4=Fair, between 0.41 & 0.6=Moderate, between 0.61 & 0.8=Good, between 0.81 & 1=Very Good.

The level of agreement observed within the domains was almost consistent for both Proxy 1 and 2 versions and is presented in Appendix 1. The "walking about" domain showed very good agreement across all subgroups, with boy-children reporting the highest agreement level and the youngest age group reporting the lowest. In the domains "looking after myself" and "doing usual activities", overall, the lowest agreement was seen among children aged 6-7 years and their parents. 11–12-year-olds reported the highest dyad agreement using Proxy 1 for both "looking after myself" and "doing usual activities" and only for "looking after myself" when Proxy 2 was used. Girl-child and parent dyads reported the highest agreement level for the "doing usual activities" domain using Proxy 2. Children who reported an "excellent" SRH had the best overall agreement within the "having pain/discomfort" domain while it was lowest but moderate for children from a low-to-middle income household. The overall agreement in the "feeling worried, sad or unhappy" domain was the highest for parents reporting they knew their children extremely well and the lowest for "very well/well".

Second stage assessment-deliberation:

Response change post deliberation:

42 dyads completed the second stage deliberation using the EQ-5D-Y-3L (Proxy 1) measure. A response change[5] was observed in 32 (76%) of the dyad assessments of the child's HRQoL in one or more domains. As shown in Table 5, during deliberation, almost twice as many

---

[5] Apart from the response shifts mentioned in the results, three children changed their responses in the absence of disagreement. This has not been included in the calculation of response shifts as 1) it does not reflect the impact of deliberation, and 2) to avoid confusion.

children (47%) as parents (25%) agreed to change their responses. However, seventeen children (53%) did not change their responses despite the discrepancy. On the other hand, three-quarters of parent-proxy respondents (75%) chose to remain with their initial responses. When divergences were present, the largest movements in child-self responses (12.5%) were observed in the domains "doing usual activities" and "having pain/discomfort", followed by "walking about" and "looking after myself" (9.4%). Only one of the 15 children who disagreed with their parent in the "feeling worried, sad or unhappy" domain changed their response. In comparison to other domains, a higher proportion of response shift (9.4%) among parents was observed within the "having pain/discomfort" and "feeling worried, sad or unhappy" domains. Most of the child-self response shift was observed as a one-level change in magnitude. A two-level response shift (two level decrease) was seen only in the "having pain/discomfort" domain. In the proxy report, whilst a one-level change (increase) was observed in only one domain, a two-level magnitude shift (decrease) was observed within every other domain where a response shift was reported. A detailed profile-level child and proxy responses is provided in Appendix 2.

Table 5: Percentage of post-deliberation response shift with respect to direction and level in the presence of child-proxy disagreement within each domain

Domains	SELF						PROXY					
	N	-2	-1	0	1	2	-2	-1	0	1	2	N
Walking about	7			57	43				86		14	7
Looking after myself	10			70	30				100			10
Doing usual activities	16		19	75	6				94		6	16
Having pain/discomfort	12		17	67	8	8		25	75			12
Feeling worried, sad or unhappy	15			93	7				80		20	15
Total	32			53					25			

The numbers in the second row from -2 to 2 represent all the possible direction and level of shift. The sign determines the direction of change, negative for level increase and positive for level decrease (for example, changing response from level 2 to level 1 would have a positive sign for level decrease and vice versa), whereas the number represents the magnitude of level change. 0 represents no change in response.

Dyad agreement for EQ-5D-Y-3L (Proxy 1) values:

Table 6 describes the overall and subgroup agreement for the EQ-5D-Y-3L (Proxy 1) scores in the deliberation sub-sample. A statistically significant moderate agreement was observed post deliberation (0.50, 95% CI: 0.24 to 0.7) between child-parent dyad after each of them deliberated and, if willing, changed, their responses. This was a substantial improvement as compared to the fair, non-significant agreement (0.23, 95% CI: -0.07, 0.49) before deliberation. Likewise, across all subgroups, the level of agreement increased after deliberation relative to before deliberation. Furthermore, none of the subgroups in the deliberation sub-sample showed significant agreement before deliberation. However, post deliberation, the agreement was statistically significant for the dyad pairs with the following characteristics: 6–7-year-olds, 11–12-year-olds, girl-children, mothers, children reporting "good" SRH, children with no long-term condition, high income households, and parents who reported that they knew their children very well/well. Overall, the best agreement post deliberation was found between parents and children who reported a "good" SRH. On the other hand, the worst agreement with parents, both before and after deliberation, was noted in children who reported a "very good" SRH.

Table 6: Overall and subgroup agreement between self and proxy values using EQ-5D-Y-3L (Proxy 1) before and after deliberation

		Before deliberation		After deliberation	
	No. of targets	CCC	95% CI	CCC	95% CI
Overall	42	0.23	-0.07, 0.49	0.5	0.24, 0.7
Subgroups					
Age group					
6-7 yrs	12	0.09	-0.48, 0.61	0.56	0.02, 0.85
8-10 yrs	15	0.2	-0.31, 0.62	0.46	-0.04, 0.78
11-12 yrs	15	0.26	-0.24, 0.65	0.39	0.01, 0.67
Child gender					
Girl-child	23	0.23	-0.18, 0.57	0.55	0.19, 0.78
Boy-child	18	0.24	-0.21, 0.6	0.32	-0.15, 0.66
Parent gender					
Mother	34	0.3	-0.03, 0.57	0.52	0.23, 0.73
Father	7	-0.12	-0.75, 0.62	0.51	-0.28, 0.89
Child-Self rated general health					
Excellent	7	-0.05	-0.73, 0.68	0.25	-0.55, 0.81
Very good	17	-0.13	-0.44, 0.21	0.05	-0.39, 0.47
Good	13	0.23	-0.32, 0.66	0.63	0.15, 0.87
Fair	3	0.02	-0.79, 0.8	0.32	-0.91, 0.98
Long term condition					
No	30	0.19	-0.17, 0.5	0.48	0.15, 0.71
Yes	12	0.16	-0.42, 0.64	0.5	-0.06, 0.82
Household income					
Less than $75,000:	10	0.17	-0.42, 0.66	0.31	-0.35, 0.76
Greater than $75,000:	29	0.2	-0.12, 0.48	0.48	0.15, 0.71
Know					
Very well or well	23	0.2	-0.22, 0.55	0.56	0.21, 0.78
Extremely well	18	0.12	-0.35, 0.54	0.35	-0.11, 0.68

Dyad agreement for EQ-5D-Y-3L (Proxy 1) domains:

An improvement in the child-parent dyad agreement was observed post deliberation across all domains (Table 7). The highest level of agreement was noted within the domain "walking about" followed by "having pain/discomfort", "looking after myself", "doing usual activities" and finally, "feeling worried, sad or unhappy", with the lowest level of agreement. The only domain which showed a significant improvement in agreement was the "having pain/discomfort" which improved from good to very good. The agreement in "feeling worried, sad or unhappy" and "doing usual activities" domains also improved from moderate to good albeit not significantly.

Table 7: Agreement between self and proxy domain-level HRQoL using EQ-5D-Y-3L (Proxy 1) before and after deliberation

	Before deliberation		After deliberation		Difference in agreement
	AC1	95% CI	AC1	95% CI	P-value
Walking about	0.82	0.69, 0.96	0.93	0.84, 1	0.05
Looking after myself	0.74	0.58, 0.9	0.8	0.65, 0.94	0.33

Doing usual activities	0.56	0.35, 0.76	0.66	0.48, 0.84	0.14
Having pain/discomfort	0.66	0.48, 0.85	0.86	0.73, 0.98	0.01
Feeling worried, sad or unhappy	0.55	0.35, 0.75	0.62	0.43, 0.81	0.37

Sensitivity analysis:

The findings using the German EQ-5D-Y value-set indicate similar inter-rater agreements in terms of overall HRQoL. The agreement was 0.29 (0.08, 0.47) for the overall sample. In the deliberation sample, the post-deliberation agreement was 0.56 (0.31, 0.73), an improvement from a fair agreement of 0.29 (-0.01, 0.54). A comparison of the distribution of the child-self and proxy reported HRQoL values using the EQ-5D-Y-3L (Proxy 1) measure derived from the EQ-5D (Australian adult) and EQ-5D-Y (German) value sets is provided in Appendix 3.

Discussion:

Previous studies have used a deliberation approach in the valuation of HRQoL with adult populations [35, 36]. However, employing a dyad consensus approach in the measurement of child HRQoL is very rare. This is the first study to our knowledge to evaluate child-parent agreement using the two proxy perspectives for the EQ-5D-Y-3L, Proxy 1 and Proxy 2, in the general population via a deliberation approach. Consistent with two previous studies examining inter-rater agreement between parents and children [21, 22], this study found poor inter-rater agreement among child-parent dyads using both proxy versions. It was expected a priori that the overall and domain level HRQoL reports would converge by applying a deliberative approach and allowing the informants to consider each other's frames of reference for completing the measure. The findings of this study suggest that a better inter-rater agreement can be achieved by using a deliberative technique with children aged 6-12 years and their parents using the EQ-5D-Y-3L measure.

In the overall sample, we found that proxies reported slightly higher HRQoL regardless of the proxy version used, as measured by preference weighted profiles, than the children themselves. However, proxy measures underestimated child HRQoL with a statistically significant difference for boy-children and for children reporting a "very good" general health on the SRH item. The EQ VAS, which generates HRQoL scores using an overall assessment of HRQoL on a 0-100 scale by the child and proxy, consistently yielded higher proxy-reported scores regardless of the perspective adopted. In particular, proxies overestimated their child's HRQoL on the EQ VAS when asked to rate it from Proxy 1 perspective (diff=-3.1, p-value=0.39) and, significantly, from Proxy 2 (proxy-child) perspective (diff=-5.1, p-value=0.02). A statistically significant difference in the self and proxy VAS scores was also reported in a study by Jelsma and Ramma involving school children using the EQ-5D-Y-3L Proxy 2 measure [37]. These findings underscore the fact that children apply a different set of internal standards to evaluate their health than parents [16]. According to one study, children under the age of 7 may lack the conceptual ability to use a VAS [38]. However, as compared to the EQ-5D-Y values, the EQ VAS scores were more consistent with the SRH categories especially in the younger age-group.

The concordance correlation coefficients for the overall HRQoL values between child-parent dyads were 0.20 for Proxy 1 and 0.174 for Proxy 2. A significant agreement was observed

only between child-mother dyads (0.28, 95%CI: 0.05 to 0.69), whilst the agreement was among the lowest for father-child pairs. A study by Perez Sousa et al in children with cerebral palsy also reported a higher agreement with mothers than fathers [39]. The authors hypothesized that this could be because they spend less time with their children than mothers and mothers take bigger role as the primary carer of their children than fathers. In this study, we found that mothers had spent significantly more waking time with their children in the previous 24 hours prior to the interview.

Shiroiwa et al. assessed the domain level agreement in Japanese children/adolescents aged 8-15 using Cohen's Kappa [40]. Compared with their findings, the agreement as measured by Gwet's AC1 in this study for the domain HRQoL was better (as interpreted on Altman's scale) for all domains except "looking after myself" and "doing usual activities". Cohen's Kappa is more widely used than Gwet's AC as an indicator of inter-rater agreement [21, 22]. However, the appropriateness of the Kappa has been questioned [33]. An important consideration is whether the estimate of the inter-rater agreement significantly differs within the same population measured using the two above mentioned statistics.

Children in this study were sampled from the general community and thus tended not to present with any obvious physical impairment. Therefore, very good child-proxy agreement using both the proxy versions was noted for the domain "walking about". Further, the agreement was good in all other domains except for "feeling worried, sad or unhappy" (AC1=0.58) when Proxy 1 was used which showed moderate agreement. Despite the significant inter-rater agreement for all subgroups, the youngest group of children aged 6-7 years had the lowest inter-rater agreement of all subgroups within each domain with observable attributes, namely, "walking about", "looking after myself" and "doing usual activities". Larger discrepancies in child-self and proxy reports have been commonly observed within this age-group [41]. This has been attributed to either the inability of young children to accurately self-report or a reporting different interpretation of the same construct [42]. Agreement was observed to be low within the "having pain or discomfort" domain if the dyads were from a low-to-middle income household. On the question of how well they knew their children, parents who responded "extremely well" showed a higher inter-rater agreement (better with Proxy 2 than Proxy 1) in the "feeling worried, sad or unhappy" domain than those who reported "very well/well".

In the second stage, when divergence existed, both the child and the parent were able to express their viewpoint and explain their responses. Children re-evaluated their initial answers more frequently than parents. As previously reported by Ungar et al. [25], we observed that parents helped children to recall events relevant to their responses to the questions. In one instance, the child reported no pain or discomfort. After being reminded by the parent, however, the child agreed to feeling some pain and changed their response. Additionally, we found that children did not align their responses with parental proxy reports if they believed that their initial response reflected the 'true' response. The child, in the example above, also reported that they were not feeling worried, sad or unhappy, whereas the parent responded that the child was a bit worried, sad or unhappy. Whilst the proxy report diverged, the child regarded their response as credible and did not alter it. These observations warrant a need for further research to investigate the reasons for response shifts in more detail. Further research is underway to fully analyse the qualitative data to explore the reasons for the shift in child and proxy responses.

Post deliberation, we found a statistically significant inter-rater agreement for the overall HRQoL and across all domains between the child-parent dyads. Children and parents with "good" SRH demonstrated a significant and the highest level of agreement across subgroups. Relative to other categories of SRH, this group was characterised by a higher proportion of girls, older children (age group 11-12 years), children without current illnesses, and parents who reported they knew their children "very well/well". Independently, each of these subgroups had also shown a significant dyad agreement after deliberation.

Consistent with earlier studies evaluating domain level HRQoL agreement using other measures such as HUI 2/3 [43-46], the agreement was the low for the less observable "feeling worried, sad or unhappy" domain. However, in the deliberation sub-sample a low agreement was also noted for the "doing usual activities" domain, which has been documented to achieve a better agreement [40, 47, 48]. The deliberation approach resulted in a statistically significant improvement in child-parent agreement in the "having pain or discomfort" domain. Within this domain, five out of seven responses post deliberation increased by one level, suggesting that both the child's and the parent's perspectives were valid when the level of pain or discomfort experienced by the child was underestimated by either of the informants.

It is important to note that prior studies have utilised adult weights to compute child values due to the absence of country-specific EQ-5D-Y-3L valuation sets [9, 49]. Given that the EQ-5D-Y-3L valuation set for Australia is not yet available and our aim is not to assess the HRQoL of children in this sample, we used Australian EQ-5D adult weights to calculate self and proxy child values [29].The same value set was applied to both child and proxy reports such that the values are comparable. In addition, a German value-set was used to check the robustness of the analysis. Although the deliberative approach resulted in a higher level of agreement between child and parent using the EQ-5D-Y-3L Proxy 1 measure, some important factors need to be explored. Further investigation would be undertaken to determine the validity of the self-report in this sample. Additionally, the deliberation process would be analysed to understand the factors which influence the child and proxy perceptions when responses are altered. Since this study sample comprised an overall healthy population, we were unable to examine the impact of deliberation with more severe levels of HRQoL domains. Therefore, a study using a different population or other child-specific utility instruments would provide additional insight into the use of the deliberation method to improve inter-rater agreement.

Conclusion:
This study has demonstrated that the dyad deliberation approach may improve agreement in child-self and proxy HRQoL reports by reconciling disparate accounts of HRQoL by children and their proxies. However, there is a need to explore ways of further refining the deliberative approach to make it more readily implementable (e.g., online mode of administration) since this approach may be resource intensive. Future research should investigate child-self and proxy assessment of HRQoL and the potential for the dyad deliberation approach to improve agreement in larger and more diverse community-based samples and paediatric patient populations.

References:

1. Kwon J, Freijser L, Huynh E, Howell M, Chen G, Khan K, et al. Systematic Review of Conceptual, Age, Measurement and Valuation Considerations for Generic Multidimensional Childhood Patient-Reported Outcome Measures. PharmacoEconomics, 2022;40(4):379-431. 10.1007/s40273-021-01128-0.

2. Brazier J, Ratcliffe J, Saloman J,Tsuchiya A. Measuring and Valuing Health Benefits for Economic Evaluation: Oxford University Press; 2016.

3. Chen G,Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. Pharmacoeconomics, 2015;33(10):1013-28. https://doi.org/10.1007/s40273-015-0286-7.

4. Bailey C, Dalziel K, Cronin P, Devlin N,Viney R. How are Child-Specific Utility Instruments Used in Decision Making in Australia? A Review of Pharmaceutical Benefits Advisory Committee Public Summary Documents. Pharmacoeconomics, 2022;40(2):157-182. 10.1007/s40273-021-01107-5.

5. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res, 2010;19(6):875-86. 10.1007/s11136-010-9648-y.

6. Niemensivu R, Roine RP, Sintonen H,Kentala E. Health-related quality of life in hearing-impaired adolescents and children. Acta Otolaryngol, 2018;138(7):652-658. 10.1080/00016489.2018.1438661.

7. Gusi N, Perez-Sousa MA, Gozalo-Delgado M,Olivares PR. [Validity and reliability of the spanish EQ-5D-Y proxy version]. An Pediatr (Barc), 2014;81(4):212-9. https://doi.org/10.1016/j.anpedi.2013.11.028.

8. Bharij A, Neighbors K, Alonso EM,Mohammad S. Health utility and quality of life in pediatric liver transplant recipients. Pediatric Transplantation, 2020;24(4):e13720. https://doi.org/10.1111/petr.13720.

9. Bray N, Noyes J, Harris N,Edwards RT. Measuring the health-related quality of life of children with impaired mobility: examining correlation and agreement between children and parent proxies. BMC Res Notes, 2017;10(1):377. https://doi.org/10.1186/s13104-017-2683-9.

10. Grootens-Wiegers P, Hein IM, van den Broek JM,de Vries MC. Medical decision-making in children and adolescents: developmental and neuroscientific aspects. BMC Pediatr, 2017;17(1):120. 10.1186/s12887-017-0869-x.

11. Matza LS, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. Value in Health, 2013;16(4):461-479. https://doi.org/10.1016/j.jval.2013.04.004.

12. Bevans KB, Ahuvia IL, Hallock TM, Mendonca R, Roth S, Forrest CB, et al. Investigating child self-report capacity: a systematic review and utility analysis. Qual Life Res, 2020;29(5):1147-1158. 10.1007/s11136-019-02387-3.

13. White-Koning M, Arnaud C, Dickinson HO, Thyen U, Beckung E, Fauconnier J, et al. Determinants of child-parent agreement in quality-of-life reports: a European study of children with cerebral palsy. Pediatrics, 2007;120(4):e804-14. 10.1542/peds.2006-3272.

14. Varni JW, Limbers CA,Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age

subgroups using the PedsQL 4.0 Generic Core Scales. Health Qual Life Outcomes, 2007;5:2. 10.1186/1477-7525-5-2.

15. Mpundu-Kaambwa C, Chen G, Dalziel K, Devlin N,Ratcliffe J. International guidelines for self-report and proxy completion of paediatric health-related quality of life measures: a protocol for a systematic review. BMJ Open, 2021;11(12):e052049. 10.1136/bmjopen-2021-052049.

16. Theunissen NC, Vogels TG, Koopman HM, Verrips GH, Zwinderman KA, Verloove-Vanhorick SP, et al. The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res, 1998;7(5):387-97. 10.1023/a:1008801802877.

17. Ooi YP, Weng S-J, Magiati I, Ang RP, Goh TJ, Fung DS, et al. Factors Influencing Agreement between Parent and Child Reports of Anxiety Symptoms among Children with High-Functioning Autism Spectrum Disorders. Journal of Developmental and Physical Disabilities, 2016;28(3):407-424. 10.1007/s10882-016-9481-5.

18. Cremeens J, Eiser C,Blades M. Factors influencing agreement between child self-report and parent proxy-reports on the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health Qual Life Outcomes, 2006;4:58. 10.1186/1477-7525-4-58.

19. Rensen N, Steur LMH, Schepers SA, Merks JHM, Moll AC, Kaspers GJL, et al. Determinants of health-related quality of life proxy rating disagreement between caregivers of children with cancer. Quality of Life Research, 2020;29(4):901-912. 10.1007/s11136-019-02365-9.

20. Roddenberry A,Renk K. Quality of life in pediatric cancer patients: The relationships among parents' characteristics, children's characteristics, and informant concordance. Journal of Child and Family Studies, 2008;17(3):402-426. https://doi.org/10.1007/s10826-007-9155-0.

21. Jiang M, Ma Y, Li M, Meng R, Ma A,Chen P. A comparison of self-reported and proxy-reported health utilities in children: a systematic review and meta-analysis. Health and Quality of Life Outcomes, 2021;19(1):45. https://doi.org/10.1186/s12955-021-01677-0.

22. Khadka J, Kwon J, Petrou S, Lancsar E,Ratcliffe J. Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic evaluation: A systematic review. Soc Sci Med, 2019;240:112543. https://doi.org/10.1016/j.socscimed.2019.112543.

23. Petrou S. Methodological challenges surrounding QALY estimation for paediatric economic evaluation. Cost Effectiveness and Resource Allocation, 2022;20(1):10. 10.1186/s12962-022-00345-4.

24. Dirks MA, De Los Reyes A, Briggs-Gowan M, Cella D,Wakschlag LS. Annual research review: embracing not erasing contextual variability in children's behavior--theory and utility in the selection and use of methods and informants in developmental psychopathology. J Child Psychol Psychiatry, 2012;53(5):558-74. 10.1111/j.1469-7610.2012.02537.x.

25. Ungar WJ, Boydell K, Dell S, Feldman BM, Marshall D, Willan A, et al. A parent-child dyad approach to the assessment of health status and health-related quality of life in children with asthma. PharmacoEconomics, 2012;30(8):697-712. https://doi.org/10.2165/11597890-000000000-00000.

26. Welfare AIoHa. Health of children. Canberra: AIHW; 2020.

27. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav, 1997;38(1):21-37.

28. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Quality of Life Research, 2010;19(6):887-897. 10.1007/s11136-010-9649-x.

29. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, et al. Time trade-off derived EQ-5D weights for Australia. Value Health, 2011;14(6):928-36. 10.1016/j.jval.2011.04.009.

30. Kreimeier S, Mott D, Ludwig K, Greiner W, Prevolnik Rupel V, Ramos-Goñi JM, et al. EQ-5D-Y Value Set for Germany. PharmacoEconomics, 2022. 10.1007/s40273-022-01143-9.

31. STATISTICS ABO. Socio-Economic Indexes for Areas (SEIFA) 2033.0.55.001 CANBERRA2018.

32. Chen C-C, Barnhart HX. Comparison of ICC and CCC for assessing agreement for data without and with replications. Computational Statistics & Data Analysis, 2008;53(2):554-564. https://doi.org/10.1016/j.csda.2008.09.026.

33. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol, 2013;13:61. 10.1186/1471-2288-13-61.

34. Klein D. Implementing a general framework for assessing interrater agreement in Stata. Stata Journal, 2018;18(4):871-901.

35. Stein K, Ratcliffe J, Round A, Milne R, Brazier JE. Impact of discussion on preferences elicited in a group setting. Health and Quality of Life Outcomes, 2006;4(1):22. 10.1186/1477-7525-4-22.

36. Gansen F, Klinger J, Rogowski W. MCDA-based deliberation to value health states: lessons learned from a pilot study. Health and Quality of Life Outcomes, 2019;17(1):112. 10.1186/s12955-019-1189-7.

37. Jelsma J, Ramma L. How do children at special schools and their parents perceive their HRQoL compared to children at open schools? Health and quality of life outcomes, 2010;8:72-72. https://doi.org/10.1186/1477-7525-8-72.

38. Shields BJ, Palermo TM, Powers JD, Grewe SD, Smith GA. Predictors of a child's ability to use a visual analogue scale. Child Care Health Dev, 2003;29(4):281-90. 10.1046/j.1365-2214.2003.00343.x.

39. Perez Sousa M, Olivares Sánchez-Toledo PR, Gusi Fuerte N. Parent-child discrepancy in the assessment of health- related quality of life using the EQ-5D-Y questionnaire. Arch Argent Pediatr, 2017;115(6):541-546. https://doi.org/10.5546/aap.2017.eng.541.

40. Shiroiwa T, Fukuda T, Shimozuma K. Psychometric properties of the Japanese version of the EQ-5D-Y by self-report and proxy-report: reliability and construct validity. Qual Life Res, 2019;28(11):3093-3105. https://doi.org/10.1007/s11136-019-02238-1.

41. Conijn JM, Smits N, Hartman EE. Determining at What Age Children Provide Sound Self-Reports: An Illustration of the Validity-Index Approach. Assessment, 2019;27(7):1604-1618. 10.1177/1073191119832655.

42. Woolley ME, Bowen GL, Bowen NK. Cognitive Pretesting and the Developmental Validity of Child Self-Report Instruments: Theory and Applications. Research on Social Work Practice, 2004;14(3):191-200. 10.1177/1049731503257882.

43. Barr RD, Simpson T, Whitton A, Rush B, Furlong W,Feeny DH. Health-related quality of life in survivors of tumours of the central nervous system in childhood--a preference-based approach to measurement in a cross-sectional study. Eur J Cancer, 1999;35(2):248-55. https://doi.org/10.1016/s0959-8049(98)00366-9.

44. Glaser A, Kennedy C, Punt J,Walker D. Standardized quantitative assessment of brain tumor survivors treated within clinical trials in childhood. Int J Cancer Suppl, 1999;12:77-82. https://doi.org/10.1002/(sici)1097-0215(1999)83:12+<77::aid-ijc14>3.0.co;2-x.

45. Morrow AM, Hayen A, Quine S, Scheinberg A,Craig JC. A comparison of doctors', parents' and children's reports of health states and health-related quality of life in children with chronic conditions. Child Care Health Dev, 2012;38(2):186-95. https://doi.org/10.1111/j.1365-2214.2011.01240.x.

46. Verrips GH, Stuifbergen MC, den Ouden AL, Bonsel GJ, Gemke RJ, Paneth N, et al. Measuring health status using the Health Utilities Index: agreement between raters and between modalities of administration. J Clin Epidemiol, 2001;54(5):475-81. https://doi.org/10.1016/s0895-4356(00)00317-6.

47. Lin J, Wong CKH, Cheung PWH, Luo N,Cheung JPY. Feasibility of Proxy-Reported EQ-5D-3L-Y and Its Agreement in Self-reported EQ-5D-3L-Y for Patients With Adolescent Idiopathic Scoliosis. Spine (Phila Pa 1976), 2020;45(13):E799-e807. https://doi.org/10.1097/brs.0000000000003431.

48. Zhou W, Shen A, Yang Z, Wang P, Wu B, Herdman M, et al. Patient-caregiver agreement and test-retest reliability of the EQ-5D-Y-3L and EQ-5D-Y-5L in paediatric patients with haematological malignancies. Eur J Health Econ, 2021;22(7):1103-1113. https://doi.org/10.1007/s10198-021-01309-w.

49. Kirkham FJ, Vigevano F, Raspall-Chaure M, Wilken B, Lee D, Le Reun C, et al. Health-related quality of life and the burden of prolonged seizures in noninstitutionalized children with epilepsy. Epilepsy Behav, 2020;102:106340. 10.1016/j.yebeh.2019.04.058.

Appendices:

Appendix 1.1: Domain level agreement for EQ-5D-Y-3L (Proxy 1) by subgroups.

		Domains								
		Walking about		Looking after myself		Doing usual activities		Having pain/discomfort		Feeling worried, sad or unhappy	
Subgroups	No. of subjects	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI
Age group											
6-7 yrs	23	0.81	0.61, 1	0.54	0.24, 0.83	0.42	0.1, 0.73	0.75	0.52, 0.98	0.59	0.3, 0.87
8-10 yrs	30	0.90	0.77, 1	0.82	0.65, 0.98	0.69	0.48, 0.91	0.67	0.44, 0.89	0.62	0.39, 0.85
11-12 yrs	32	0.90	0.79, 1	0.90	0.79, 1	0.87	0.73, 1	0.64	0.42, 0.86	0.52	0.28, 0.77
Child gender											
Girl-child	47	0.82	0.68, 0.95	0.74	0.58, 0.89	0.71	0.55, 0.87	0.66	0.49, 0.84	0.63	0.45, 0.81
Boy-child	37	0.97	0.92, 1	0.86	0.72, 0.99	0.69	0.51, 0.88	0.72	0.54, 0.9	0.53	0.3, 0.76
Parent gender											
Mother	68	0.89	0.81, 0.97	0.79	0.68, 0.9	0.68	0.54, 0.82	0.71	0.57, 0.84	0.57	0.41, 0.73
Father	16	0.87	0.66, 1	0.72	0.42, 1	0.72	0.42, 1	0.62	0.29, 0.96	0.56	0.2, 0.92

	No. of subjects	Walking about AC1	95% CI	Looking after myself AC1	95% CI	Doing usual activities AC1	95% CI	Having pain/discomfort AC1	95% CI	Feeling worried, sad or unhappy AC1	95% CI
Child-Self rated general health											
Excellent	20	0.84	0.64, 1	0.78	0.55, 1	0.66	0.37, 0.94	0.89	0.73, 1	0.63	0.32, 0.94
Very good	44	0.91	0.81, 1	0.85	0.73, 0.97	0.85	0.73, 0.97	0.63	0.44, 0.81	0.50	0.3, 0.71
Good	16	0.87	0.66, 1	0.80	0.54, 1	0.37	-0.04, 0.78	0.70	0.39, 1	0.68	0.37, 1
Fair	3			-0.33	-0.33, -0.33	0.14	-1, 1	0.61	-1, 1	1.00	1, 1
Long term condition											
No	59	0.87	0.78, 0.97	0.78	0.65, 0.9	0.76	0.62, 0.89	0.70	0.56, 0.85	0.61	0.45, 0.78
Yes	26	0.88	0.73, 1	0.79	0.6, 0.98	0.54	0.26, 0.82	0.62	0.37, 0.88	0.49	0.22, 0.77
Household income											
Less than $75,000:	18	0.82	0.6, 1	0.61	0.3, 0.93	0.68	0.37, 0.98	0.44	0.09, 0.8	0.49	0.16, 0.83
Greater than $75,000: **Know**	60	0.91	0.84, 0.99	0.86	0.76, 0.96	0.72	0.58, 0.86	0.79	0.66, 0.91	0.64	0.49, 0.8
Very well or well	39	0.89	0.78, 1	0.71	0.53, 0.89	0.61	0.4, 0.82	0.59	0.38, 0.8	0.43	0.21, 0.65
Extremely well	45	0.88	0.78, 0.99	0.83	0.71, 0.96	0.75	0.6, 0.91	0.77	0.62, 0.92	0.69	0.51, 0.87

Appendix 1.2: Domain level agreement for EQ-5D-Y-3L (Proxy 2) by subgroups

		Domains									
		Walking about		Looking after myself		Doing usual activities		Having pain/discomfort		Feeling worried, sad or unhappy	
Subgroups	No. of subjects	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI	AC1	95% CI
Age group											
6-7 yrs	23	0.81	0.61, 1	0.54	0.24, 0.83	0.48	0.17, 0.78	0.64	0.37, 0.9	0.64	0.37, 0.9
8-10 yrs	30	0.86	0.71, 1	0.90	0.77, 1	0.74	0.54, 0.94	0.70	0.49, 0.92	0.71	0.5, 0.92
11-12 yrs	32	0.90	0.79, 1	0.90	0.79, 1	0.90	0.79, 1	0.59	0.36, 0.82	0.74	0.54, 0.93
Child gender											
Girl-child	47	0.82	0.68, 0.95	0.94	0.86, 1	0.76	0.62, 0.91	0.64	0.45, 0.82	0.68	0.51, 0.85
Boy-child	37	0.94	0.86, 1	0.89	0.77, 1	0.73	0.55, 0.91	0.67	0.48, 0.87	0.75	0.58, 0.92
Parent gender											
Mother	68	0.88	0.79, 0.96	0.83	0.72, 0.93	0.74	0.61, 0.86	0.66	0.52, 0.8	0.71	0.58, 0.84
Father	16	0.87	0.66, 1	0.72	0.42, 1	0.72	0.42, 1	0.62	0.29, 0.96	0.64	0.31, 0.97
Child-Self rated general health											
Excellent	20	0.84	0.64, 1	0.78	0.55, 1	0.72	0.46, 0.98	0.77	0.53, 1	0.76	0.52, 1
Very good	44	0.88	0.77, 0.99	0.85	0.73, 0.97	0.88	0.77, 0.99	0.61	0.42, 0.8	0.74	0.58, 0.9
Good	16	0.87	0.66, 1	0.87	0.66, 1	0.38	-0.01, 0.78	0.70	0.39, 1	0.60	0.26, 0.94
Fair	3			0.14	-1, 1	0.61	-1, 1	0.61	-1, 1	0.61	-1, 1
Long term condition											
No	59	0.86	0.75, 0.96	0.82	0.7, 0.93	0.76	0.62, 0.89	0.68	0.53, 0.83	0.70	0.56, 0.85
Yes	26	0.88	0.73, 1	0.79	0.6, 0.98	0.70	0.47, 0.93	0.56	0.29, 0.83	0.70	0.47, 0.93
Household income											
Less than $75,000:	18	0.82	0.6, 1	0.68	0.39, 0.97	0.75	0.49, 1	0.42	0.07, 0.78	0.79	0.53, 1
Greater than $75,000:	60	0.91	0.84, 0.99	0.88	0.78, 0.97	0.76	0.63, 0.89	0.76	0.63, 0.9	0.70	0.56, 0.85

Know													
Very well or well	39	0.89	0.78, 1	0.74	0.57, 0.91	0.68	0.49, 0.87	0.55	0.34, 0.77	0.57	0.36, 0.78
Extremely well	45	0.86	0.74, 0.97	0.86	0.74, 0.97	0.78	0.63, 0.93	0.74	0.58, 0.9	0.8	0.66, 0.94

Appendix 2: A detailed description of the profiles -level differences in the child and proxy ratings of HRQoL using the EQ-5D-Y-3L (Proxy version 1) before and after deliberation in the presence of disagreement.

No.	Child-self ratings		Proxy ratings	
	Before deliberation	After deliberation	Before deliberation	After deliberation
1	21111	11111	11112	11111
2	11211	12211	11112	11111
3	12221	11221	11121	11121
4	11111	11111	11112	11112
5	11122	11121	12112	12112
6	11111	11121	11121	11121
7	11111	11111	11112	11111
8	11111	11111	21221	11221
9	12222	12222	11122	11122
10	11112	11111	11111	11111
11	11212	11212	12322	12222
12	12111	11111	11112	11112
13	22221	22221	11112	11122
14	11111	11211	11211	11211
15	11211	11111	11112	11112
16	11111	11121	12122	12122
17	11111	11111	11112	11112
18	11121	11121	11111	11121
19	11211	11211	11111	11111
20	21212	21212	12111	12111
21	12122	11222	11222	11222
22	21121	11111	11111	11111
23	11232	11212	11112	11112
24	21121	11221	11111	11111
25	11111	11111	11112	11112
26	11121	11221	12211	12211
27	11211	11211	11112	11112
28	21211	21211	11111	11111
29	11111	11111	11112	11112
30	11111	11111	11112	11112
31	11121	11121	11111	11111
32	11121	11121	11111	11121

The profiles are in order of the EQ-5D-Y domains: "walking about", "looking after myself", "doing usual activities", "having pain or discomfort", and "feeling worried, sad or unhappy". 1= No problems, 2= Some problems, 3= A lot of problems.

Appendix 3: A comparison of the distribution of the child-self and proxy reported HRQoL values using the EQ-5D-Y-3L (Proxy 1) measure derived from the EQ-5D (Australian adult) and EQ-5D-Y (German) value sets.

Distribution of child-self and proxy reported HRQoL values derived using EQ-5D (Australian) and EQ-5D-Y (German) value sets