Original article

Appropriateness of self- and caregiver reports in measuring health utility in Thai pediatric cancer patients

Ramida Yoykaew[1], Phantipa Sakthong[1], Piti Techavichit[2] and Trai Tharnpanich[1]

[1]Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Chulalongkorn University; [2]Division of Pediatric Hematology and Oncology, Faculty of Medicine, Chulalongkorn University

Abstract:

Background: *The EQ-5D-Y-3L is a standardized instrument used to assess health utility in pediatric populations for health economic evaluations. However, no studies have examined the agreement between, or the appropriateness of choosing, self- and proxy-report versions in Thai pediatric cancer patients.* ***Objective:*** *This study aimed to examine the agreement between self- and proxy-report versions of the EQ-5D-Y-3L, evaluate their appropriateness across respondent types and assess known-groups validity based on ECOG performance status.* ***Materials and methods:*** *A cross-sectional analytic study was conducted among 109 Thai pediatric cancer patients aged ≥ 8 years and their caregivers at King Chulalongkorn Memorial Hospital (March to April 2025). Agreement of health utility scores from the Thai EQ-5D-Y-3L (self- and proxy-report version 1) was assessed, along with known-groups validity based on ECOG status and ceiling/floor effects.* ***Results:*** *Overall, EQ-5D-Y-3L scores showed moderate agreement (ICC = 0.52). Agreement was higher among children aged 8-11 years (ICC = 0.68) compared to those aged 12-17 years (ICC = 0.35). The dimensions of mobility and looking after myself showed moderate agreement, with weighted kappa values of 0.52 and 0.58, respectively, while other dimensions and EQ-VAS scores showed poor agreement. A large effect size was observed for known-groups validity based on ECOG performance status, particularly in proxy reports, which yielded values comparable to or greater than self-reports, along with lower ceiling effects across all age groups.* ***Conclusion:*** *Proxy-report versions of the EQ-5D-Y-3L are appropriate for assessing health utility in Thai pediatric cancer patients, demonstrating larger effect sizes in known-groups validity based on ECOG status and lower ceiling effects in most subgroups compared to self-reports.*

Keywords : ● *EQ-5D-Y-3L questionnaire* ● *Health utility scores* ● *Agreement* ● *Pediatric cancer patients*

J Hematol Transfus Med. 2026;36:39-51.

Received 10 May 2025 Corrected 13 May 2025 Accepted 11 June 2025

Correspondence should be addressed to Trai Thampanich, Department of Pharmacy Practice, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Rama 4 Rd., Pathumwan, Bangkok 10330

นิพนธ์ต้นฉบับ

ความเหมาะสมของการรายงานโดยผู้ป่วยและผู้ดูแลในการวัดค่าอรรถประโยชน์ด้านสุขภาพในผู้ป่วยเด็กโรคมะเร็งชาวไทย

รมิดา ย้อยแก้ว[1] พรรณทิพา ศักดิ์ทอง[1] ปิติ เตชะวิจิตร์[2] และ ตรัย ธารพานิช[1]

[1]ภาควิชาเภสัชกรรมปฏิบัติ คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย [2]แผนกโลหิตวิทยาและมะเร็งเด็ก คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อ

ที่มาและความสำคัญ *แบบสอบถาม EQ-5D-Y-3L เป็นเครื่องมือสำหรับประเมินค่าอรรถประโยชน์ด้านสุขภาพในประชากรเด็ก เพื่อการวิเคราะห์ทางเศรษฐศาสตร์สาธารณสุข อย่างไรก็ตาม ยังไม่มีข้อมูลการศึกษาที่ประเมินความสอดคล้องและความเหมาะสมในการเลือกใช้ฉบับที่ผู้ป่วยและผู้ดูแลตอบในผู้ป่วยเด็กโรคมะเร็งในประเทศไทย* ***วัตถุประสงค์*** *เพื่อศึกษาความสอดคล้องและความเหมาะสมในการเลือกใช้แบบสอบถาม EQ-5D-Y-3L ฉบับที่ผู้ป่วยตอบหรือผู้ดูแลตอบ และเพื่อประเมินความตรงจากการเทียบกับกลุ่มที่รู้ตามระดับสมรรถภาพการทำกิจวัตรประจำวัน โดยพิจารณาร่วมกับค่า ceiling และ floor effect* ***แนวทางและวิธีการ*** *งานวิจัยนี้เป็นการศึกษาเชิงวิเคราะห์ที่จุดเวลาใดเวลาหนึ่งในผู้ป่วยเด็กโรคมะเร็งชาวไทยอายุ 8 ปีขึ้นไป และผู้ดูแล จำนวน 109 คู่ ณ โรงพยาบาลจุฬาลงกรณ์ สภากาชาดไทย ช่วงเดือนมีนาคมถึงเมษายน พ.ศ. 2568 โดยประเมินความสอดคล้องของค่าอรรถประโยชน์ด้านสุขภาพที่ได้จากแบบสอบถาม EQ-5D-Y-3L ทั้งฉบับที่ผู้ป่วยตอบและฉบับตัวแทนที่ผู้ดูแลตอบ (proxy version 1) พร้อมวิเคราะห์ความตรงเชิงกลุ่มที่รู้ โดยพิจารณาร่วมกับค่า ceiling และ floor effect* ***ผลลัพธ์*** *ค่า EQ-5D-Y-3L โดยรวมมีความสอดคล้องในระดับปานกลาง (ICC = 0.52) โดยกลุ่มเด็กอายุ 8-11 ปีมีความสอดคล้องสูงกว่า (ICC = 0.68) เมื่อเทียบกับกลุ่มอายุ 12-17 ปี (ICC = 0.35) มิติการเคลื่อนไหวและการดูแลตนเองมีความสอดคล้องในระดับปานกลาง โดยมีค่า weighted kappa เท่ากับ 0.52 และ 0.58 ตามลำดับ ขณะที่มิติอื่น ๆ และค่า EQ-VAS มีความสอดคล้องในระดับต่ำ นอกจากนี้ พบว่า effect size มีขนาดใหญ่ ในการวิเคราะห์ความตรงเชิงกลุ่มรู้ตามระดับคะแนน ECOG โดยเฉพาะในฉบับที่ผู้ดูแลตอบ ซึ่งให้ค่าที่ใกล้เคียงหรือสูงกว่าฉบับผู้ป่วยตอบ และมีค่า ceiling effect ต่ำกว่าในทุกช่วงอายุ* ***สรุป*** *แบบสอบถาม EQ-5D-Y-3L ฉบับที่ผู้ดูแลตอบมีความเหมาะสมในการประเมินค่าอรรถประโยชน์ด้านสุขภาพในผู้ป่วยเด็กโรคมะเร็งชาวไทย โดยแสดงค่า effect size ที่สูงกว่าในการวิเคราะห์ความตรงเชิงกลุ่มรู้ตามระดับคะแนน ECOG และมีค่า ceiling effect ต่ำกว่าเมื่อเปรียบเทียบกับฉบับที่ผู้ป่วยตอบด้วยตนเอง*

คำสำคัญ : ● *แบบสอบถาม EQ-5D-Y-3L* ● *ค่าอรรถประโยชน์ด้านสุขภาพ* ● *ความสอดคล้อง* ● *ผู้ป่วยเด็กโรคมะเร็ง*

วารสารโลหิตวิทยาและเวชศาสตร์บริการโลหิต. 2569;36:39-51.

Introduction

Childhood cancer constitutes a significant public health burden.[1,2] In 2022, Thailand reported the highest age-standardized incidence rate in Southeast Asia, at 13.6 per 100,000 population, along with a mortality rate of 5.0 per 100,000.[3] Given this burden, healthcare professionals must consider not only clinical findings but also the impact of illness and treatment on physical, psychological, social, and other aspects of patients' lives, which together constitute health-related quality of life (HRQoL) from the patient's perspective.[4]

In health economic evaluations, HRQoL is commonly quantified using health utility scores, which indicate the health status of individuals and populations.[4] The EQ-5D-Y-3L is a generic instrument designed to measure health utility scores among both children and adolescents[5, 6] and demonstrates good psychometric properties.[7-9] For children aged 8 years and older, self-reporting of health status using the EQ-5D-Y-3L is generally recommended; however, proxy reporting may be used when the child is unable to respond. Moreover, the use of the EQ-5D-Y-3L (proxy version 1) may be important among older children with cognitive impairments, behavioral problems or neurological conditions.[5,10] In proxy version 1, the proxy (parent, other caregiver or informant) is asked to respond to the questionnaire by providing their own impression of the child or adolescent's health status on the day of administration.[5]

Only one study was identified that assessed health utility scores in pediatric cancer patients using the EQ-5D-Y-3L. Conducted in China among children with hematologic malignancies, it reported a mean health utility score of 0.88±0.10 and an EQ-VAS score of 85.8±15.1. This study may underestimate health utility scores, as it included only inpatients. Health status was reported solely by patients, and agreement with proxy-reported outcomes was not evaluated.[11]

A literature review indicated that studies using the EQ-5D-Y-3L have demonstrated a wide spectrum of agreement, ranging from low to high, between self-reports by pediatric patients and proxy assessments, both across the five dimensions and the EQ-VAS score.[12-14]

Limited generalizability has been observed in studies using small, condition-specific samples, such as children with cerebral palsy,[12] and among those with non-representative demographics of children with mobility impairments.[13] The inclusion of predominantly healthy children from the general population may have influenced responses and reduced the applicability of the findings to clinical settings.[14]

A review of existing literature indicates considerable variation in agreement between self- and proxy-reported EQ-5D-Y-3L outcomes. Proxy reporting is often used when children are unable to respond, and proxy version 1 may be useful in older children with cognitive or behavioral issues. However, it remains unclear which version is more appropriate for clinical use. Therefore, this study aimed to evaluate the agreement between self- and proxy-reported health utility scores, determine the appropriateness of each version across respondent types, and assess known-groups validity based on ECOG status, including ceiling and floor effects.

The findings will provide evidence on the agreement between self-reported and proxy-reported health utility scores and the appropriateness of using the EQ-5D-Y-3L across respondent types. Additionally, the study will inform the validity of the instrument in distinguishing clinical status, thereby guiding future assessments of health utility in pediatric oncology populations.

Materials and methods

Study protocol

This cross-sectional analytic study was conducted among pediatric cancer patients at King Chulalongkorn Memorial Hospital, between March and April 2025. The study protocol was reviewed and approved by the Institutional Review Board of King Chulalongkorn Memorial Hospital (IRB No. 0959/67).

Patients and procedures

Participants were eligible for inclusion if they met the following criteria: 1) aged between 8 and 17 years; 2) diagnosed with cancer and 3) both the pediatric patients and their caregivers were Thai nationals capable of reading, understanding, and communicating in Thai.

Exclusion criteria were: 1) central nervous system metastasis; 2) diagnosis of psychiatric disorders and/or use of psychotropic medications; 3) absence of a caregiver accompanying the child on the day of data collection and 4) refusal to participate or failure to provide written informed consent from both patient and caregiver.

Sample size calculation

The required sample size was calculated based on an expected correlation coefficient greater than 0.4, as suggested by Rosner's criteria, to assess agreement between self-reported and proxy-reported outcomes.[15] With a two-sided α of 0.05 and a two-sided β of 0.05, indicating the use of a two-sided test for determining statistical significance, the final sample size was determined to be 109 pediatric cancer patients aged 8 years and older and 109 corresponding proxies.

Data collection

Caregivers and pediatric patients were invited to participate in the study while attending scheduled medical appointments at the pediatric hematology and oncology outpatient clinic or during inpatient chemotherapy treatment. Pediatric patients completed the EQ-5D-Y-3L, while their caregivers completed the EQ-5D-Y-3L (proxy version 1).

Measures

Measures included a general information form for participants and a clinical information form for pediatric patients. Physical performance status was evaluated using the Eastern Cooperative Oncology Group (ECOG) scale.

The EQ-5D-Y-3L

In Thailand, the Thai Health Technology Assessment guideline recommends the EQ-5D as the standard instrument for health utility scores in the Thai population.[16,17] The EQ-5D-Y-3L comprises five health dimensions: mobility, self-care, usual activities, pain/discomfort and feeling worried, sad or unhappy. Each dimension consists of a single item with three response levels: no problems, some problems and a lot of problems. In addition, the questionnaire includes the EuroQol Visual Analogue Scale (EQ-VAS), allowing respondents to rate their overall health on a vertical scale from 0 (the worst imaginable health state) to 100 (the best imaginable health state).[5] Currently, the EQ-5D-Y-3L has been widely used in clinical settings across multiple countries.[11,18-26]

The EQ-5D-Y-3L questionnaire (both self-reported and proxy version 1) used in this study was the officially translated Thai version, developed and certified by the EuroQol Research Foundation. The Thai version was produced following the standardized EuroQol translation methodology, including forward and backward translation, cognitive debriefing, and linguistic validation, in accordance with ISPOR guidelines. The translation was carried out by qualified bilingual professionals in Thailand and tested on Thai-speaking respondents. A translation certificate was issued by the EuroQol group to confirm the semantic and cultural equivalence of the Thai version (Version 2.0).[5] All respondents in this study completed the officially translated Thai version of the EQ-5D-Y-3L questionnaire, which is provided in full in the appendix.[5]

The EQ-5D-Y-3L is an instrument used to calculate health utility scores by applying a predefined scoring algorithm that assigns weights to each dimension.[5] Health utility scores reflect the health status of individuals and populations and are reported as index values ranging from 0 (death) to 1 (full health).[4] As of 2024, the EQ-5D-Y-3L has not yet been used to assess health utility scores among pediatric cancer patients in Thailand. In the absence of a Thai-specific value set, health utility scores in this study were calculated using the Indonesian value set, ranging from 1.000 for the best health state (coded as 11111) to -0.086 for the worst possible state (coded as 33333).[22] Health utility

scores were assessed using the EQ-5D-Y-3L, which includes both the descriptive system (reported as EQ-5D-Y-3L scores) and the visual analogue scale (reported as EQ-VAS scores).

Statistical analysis

Qualitative data were presented as frequencies, proportions and percentages. Quantitative data were expressed as mean and standard deviation (SD) or median and interquartile range (IQR), depending on the data distribution. The Shapiro-Wilk test was used to assess normality.

The Paired t-test or Wilcoxon signed-rank test was used, depending on data distribution, to examine differences in mean scores between patient self-reports and proxy-reports, and for subgroup comparisons within on-treatment and off-treatment groups. EQ-VAS agreement was assessed using intraclass correlation coefficients (ICCs), based on a two-way mixed-effects model with absolute agreement and single measurement, and weighted kappa values were used for the five dimensions. Validity was assessed through known-groups validity, using ECOG performance status as the grouping variable. For this analysis, ECOG performance status scores were categorized into two groups: 0–1 (indicating good functional status) and ≥ 2 (indicating poor functional status).

Known-groups validity was analyzed using data from all patients aged 8-17 years, for each respondent type (patient- and proxy-reported data) and stratified by age groups (8-11 and 12-17 years). Effect sizes were calculated using Cohen's d, defined as the difference in mean EQ-5D-Y-3L scores between groups divided by the pooled standard deviation. Statistical comparisons of EQ-5D-Y-3L scores between ECOG groups were conducted using the Mann-Whitney U test due to non-normal distribution. If the data had been normally distributed, an independent t-test would have been considered appropriate. Effect sizes were calculated using Cohen's d, with pooled standard deviations from group-specific SDs. These analyses aimed to assess the ability of the EQ-5D-Y-3L to discriminate between groups defined by ECOG performance status.

A floor effect was defined as the percentage of participants reporting the lowest possible score and a ceiling effect as the percentage reporting the highest possible score. Both were calculated as proportions of the total sample for health utility scores and EQ-VAS. Ceiling and floor effects were evaluated using the maximum and minimum achievable scores of each measure as reference points. For the EQ-5D-Y-3L score, a value of 1.00 was defined as the ceiling and 0.00 as the floor. For the EQ-VAS, scores of 100 and 0 were used as the ceiling and floor, respectively.[27]

According to Koo's criteria[28], ICC values below 0.50 indicate poor agreement, 0.50-0.75 moderate, 0.75-0.90 good, and above 0.90 excellent agreement. For dimension-level agreement, weighted kappa values were interpreted using Cohen's criteria,[29] where values below 0.40 represent poor agreement, 0.40-0.75 moderate, and above 0.75 excellent agreement.

Floor or ceiling effects were considered present when more than 15% of respondents reported the lowest or highest possible score, respectively, indicating potential limitations in content validity, reduced reliability and limited responsiveness, as recommended by Terwee, et al.[27] for both the EQ-5D-Y-3L and EQ-VAS scores, floor and ceiling effects were determined based on this threshold. A significance level of $\alpha = 0.05$ was used. All analyses were performed using IBM SPSS Statistics, Version 29.0.2.

Results

Of the 115 eligible patients, six were excluded because they were unable to communicate either verbally or in writing (n = 2), their caregivers were absent on the data collection day (n = 2), or they declined to participate (n = 2). This results in 109 patient-caregiver pairs included in the analysis.

Table 1 Patient Characteristics (n = 109)

Characteristic	Value
Age (years)	
Mean±SD	11±3
Sex; n (%)	
Male	59 (54.1%)
Healthcare Coverage; n (%)	
Civil servant	14 (12.8%)
Universal Coverage Scheme	95 (87.2%)
Patient education; n (%)	
No formal education	1 (0.9%)
Kindergarten	1 (0.9%)
Primary school (Grades 1-4)	47 (43.1%)
Primary school (Grades 5-6)	20 (18.3%)
Lower secondary school	34 (31.2%)
Upper secondary school	6 (5.5%)
Caregiver education; n (%)	
No formal education	2 (1.8%)
Primary school	12 (11.0%)
Lower secondary school	15 (13.8%)
Upper secondary school	16 (14.7%)
Diploma	11 (10.1%)
Bachelor's degree	44 (40.4%)
Graduate degree	9 (8.3%)
Time since diagnosis (days)	
Mean±SD	1,153±926
Types of pediatric can; n (%)	
Malignant neoplasms, benign neoplasms	3 (2.8%)
Malignant neoplasms, eye, brain and other parts of CNS	26 (23.9%)
Malignant neoplasms, stated or presumed to be primary, of lymphoid, hematopoietic and related tissues	61 (56.0%)
Malignant neoplasms, mesothelial and soft tissue	7 (6.4%)
Malignant neoplasms, thyroid endocrine and other endocrine glands	2 (1.8%)
Neoplasms of uncertain or unknown behavior	1 (0.9%)
Malignant neoplasms, bone and articular cartilage	7 (6.4%)
Malignant neoplasms, urinary tract	2 (1.8%)
On/Off treatment status; n (%)	
On-treatment	73 (67.0%)
Off-treatment	36 (33.0%)
Patient setting; n (%)	
OPD	95 (87.2%)
IPD	14 (12.8%)
ECOG score; n (%)	
ECOG 0	83 (76.1%)
ECOG 1	15 (13.8%)
ECOG 2	4 (3.7%)
ECOG 3	4 (3.7%)
ECOG 4	3 (2.8%)

Note: On-treatment was defined as being under active cancer treatment and follow-up for disease progression, whereas off-treatment was defined as having completed cancer treatment but remaining under follow-up for disease progression. Lower secondary school refers to grades 7 to 9 in the Thai education system (typically ages 12 to 14), while upper secondary school refers to grades 10 to 12 (typically ages 15 to 17)

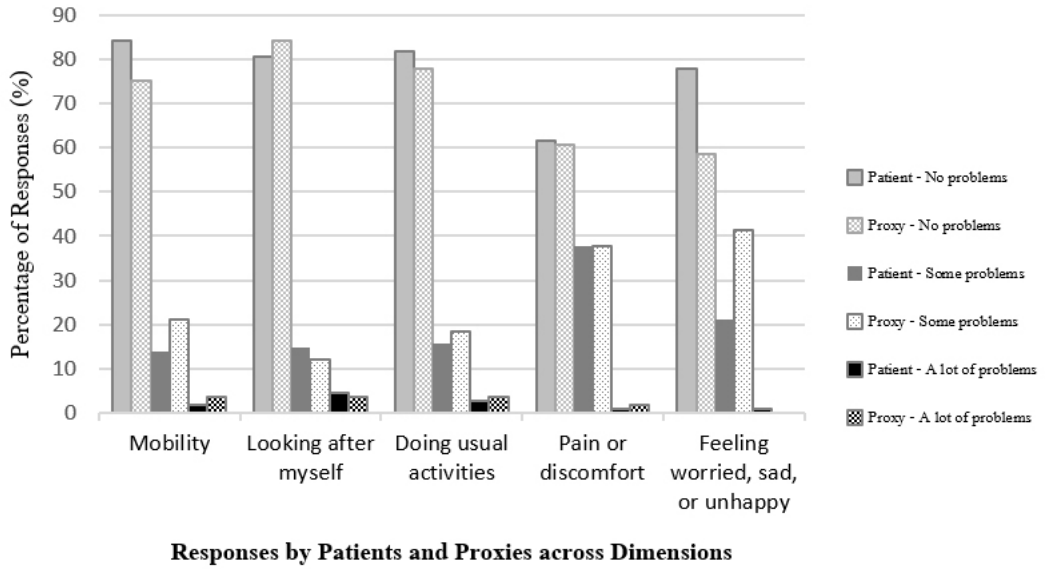

Figure 1 Distribution of Response Levels in Each EQ-5D-Y-3L Dimension Reported by Patients and Proxies (n = 109)

Table 2a Comparison of EQ-5D-Y-3L and EQ-VAS Scores between Patient and Proxy Reports

Measure	Mean±SD	Median	IQR	% Floor[a]	% Ceiling[a]	*p*-value[b]
EQ-5D-Y-3L scores (patient-report)						
All patients (n = 109)	0.94±0.10	0.98	0.07	0.9%	42.2%	-
Aged 8-11 years (n = 58)	0.95±0.1	1.00	0.06	1.7%	51.7%	-
Aged 12-17 years (n = 51)	0.93±0.1	0.98	0.08	2.0%	31.4%	-
EQ-5D-Y-3L scores (proxy-report)						
All patients (n = 109)	0.92±0.13	0.98	0.08	0.9%	32.1%	0.03
Aged 8-11 years (n = 58)	0.93±0.12	0.98	0.08	3.4%	36.2%	0.02
Aged 12-17 years (n = 51)	0.91±0.14	0.98	0.16	2.0%	27.5%	0.44
EQ-VAS scores (patient-report)						
All patients (n = 109)	81.78±20.77	90	29.5	0.9%	24.8%	-
Aged 8-11 years (n = 58)	83.33±22.36	95	25.0	1.7%	36.2%	-
Aged 12-17 years (n = 51)	80.01±18.86	85	25.0	2.0%	11.8%	-
EQ-VAS scores (proxy-report)						
All patients (n = 109)	85.23±13.19	90	15.0	3.7%	12.8%	0.60
Aged 8-11 years (n = 58)	87.00±11.17	90	15.0	1.7%	13.8%	0.66
Aged 12-17 years (n = 51)	83.22±15.03	85	25.0	5.9%	11.8%	0.80

Note: Descriptive statistics are presented as mean (standard deviation; SD), median, interquartile range (IQR) and range; [a]Floor and ceiling effects were descriptively reported without statistical testing; [b]*p*-values represent the results of Wilcoxon signed-rank tests comparing patient-reported and proxy-reported scores.

The mean age of patients was 11±3 years, with a nearly equal distribution of male and female participants. Based on the ECOG performance status, 89.9% of patients were classified as ECOG < 2 and 10.1% as ECOG ≥ 2.

Figure 1 shows that patients more frequently reported "no problems" in most dimensions, particularly mobility and doing usual activities. In contrast, proxies tended to report more problems across all dimensions, with the greatest discrepancy observed in the emotional domain, feeling worried, sad or unhappy, where 41.3% of proxies reported "some problems" compared to 21.1% of patients.

As shown in Table 2a, the mean EQ-5D-Y-3L scores were marginally higher among patient reports across all age groups, with statistically significant differences observed in the overall sample and among patients aged 8 to 11 years ($p < 0.05$). In contrast, no statistically

significant differences were found between patient and proxy reports for EQ-VAS scores. Notably, the ceiling effects were consistently lower in proxy reports than in patient reports across all age groups for both EQ-5D-Y-3L and EQ-VAS scores, with values exceeding 15% in most subgroups. Floor effects were minimal, with values below 15% for all measures in both reporting groups.

As shown in Table 2b, comparisons between patient-reported and proxy-reported scores within the on-treatment and off-treatment groups revealed that off-treatment patients tended to report higher EQ-5D-Y-3L and EQ-VAS scores, particularly in proxy reports. In the on-treatment group, proxy-reported EQ-5D-Y-3L scores were significantly lower than patient-reported scores ($p < 0.01$), whereas no significant differences were observed in the off-treatment group or for EQ-VAS scores in either group. Floor effects were below 15% in all subgroups, while ceiling effects exceeded 15% in most subgroups, except for EQ-VAS scores from proxy reports.

Table 3 presents the level of agreement between patients and proxy reports across EQ-5D-Y-3L dimen-

Table 2b Comparison of EQ-5D-Y-3L Scores and EQ-VAS Scores between On-treatment and Off-treatment Groups

Measure	Mean±SD	Median	IQR	*p*-value[a]	Range	% Floor	% Ceiling
EQ-5D-Y-3L scores (patient-report)							
On-treatment (n = 73)	0.93±0.11	0.98	0.08	-	0.43-1.00	1.37%	41.10%
Off-treatment (n = 36)	0.96±0.06	0.98	0.07	-	0.75-1.00	2.78%	44.44%
EQ-5D-Y-3L scores (proxy-report)							
On-treatment (n = 73)	0.89±0.15	0.93	0.14	< 0.01	0.35-1.00	1.37%	23.29%
Off-treatment (n = 36)	0.97±0.06	1.00	0.02	0.30	0.73-1.00	2.78%	50.00%
EQ-VAS scores (patient-report)							
On-treatment (n = 73)	80.69±18.81	85.00	25.0	-	5-100	1.37%	17.80%
Off-treatment (n = 36)	83.99±24.40	95.00	23.8	-	10-100	5.56%	38.89%
EQ-VAS scores (proxy-report)							
On-treatment (n = 73)	83.13±13.71	85.00	20.5	0.70	50-100	4.11%	12.33%
Off-treatment (n = 36)	89.49±11.07	95.00	16.0	0.70	50-100	2.78%	13.89%

Note: Descriptive statistics are presented as mean (standard deviation; SD), median, interquartile range (IQR) and range. Floor and ceiling effects were descriptively reported without statistical testing. [a]*P*-values represent the results of Wilcoxon signed-rank tests comparing patient-reported and proxy-reported scores within each treatment group.

Table 4 Agreement between Pediatric Patients and Proxies on EQ-5D-Y-3L and EQ-VAS Scores

Measure	ICC (95%CI)	*p*-value	Level of Agreement
EQ-VAS scores			
All patients (n=109)	0.28 (0.10-0.44)	< 0.01	Poor
Aged 8–11 years (n=58)	0.27 (0.02-0.49)	0.02	Poor
Aged 12–17 years (n=51)	0.28 (0.01-0.51)	0.02	Poor
EQ-5D-Y-3L scores			
All patients (n=109)	0.52 (0.37-0.64)	< 0.01	Moderate
Aged 8–11 years (n=58)	0.68 (0.50-0.80)	< 0.01	Moderate
Aged 12–17 years (n=51)	0.35 (0.09-0.57)	< 0.01	Poor

Note : Intraclass correlation coefficients (ICCs) were calculated using a two-way mixed-effects model, absolute agreement, single measurement, to assess agreement between pediatric patients and proxies on EQ-VAS and EQ-5D-Y-3L scores.

Table 3 Agreement between Pediatric Cancer Patients and their Caregivers on EQ-5D-Y-3L Dimensions, Stratified by Age Group

Dimension	All patients aged 8-17 years (n = 109)			Aged 8-11 years (n = 58)			Aged 12-17 years (n = 51)		
	Weighted kappa (95%CI)	*p*-value	Level of Agreement	Weighted kappa (95%CI)	*p*-value	Level of Agreement	Weighted kappa (95%CI)	*p*-value	Level of Agreement
Mobility	0.52 (0.32-0.71)	< 0.01	Moderate	0.32 (0.01-0.63)	< 0.01	Poor	0.71 (0.49-0.93)	< 0.01	Moderate
Looking after myself	0.58 (0.38-0.77)	< 0.01	Moderate	0.60 (0.37-0.84)	< 0.01	Moderate	0.53 (0.18-0.88)	< 0.01	Moderate
Doing usual activities	0.36 (0.16-0.55)	< 0.01	Poor	0.39 (0.12-0.66)	< 0.01	Poor	0.33 (0.06-0.59)	< 0.01	Poor
Pain or discomfort	0.31 (0.14-0.48)	< 0.01	Poor	0.34 (0.09-0.60)	< 0.01	Poor	0.23 (-0.01-0.47)	0.06	Poor
Feeling worried, sad or unhappy	0.28 (0.12-0.45)	< 0.01	Poor	0.19 (-0.03-0.41)	0.07	Poor	0.33 (0.09-0.57)	0.01	Poor

Note: 95%CI, 95% confidence interval

sions, as assessed by weighted kappa. Agreement was moderate in the dimensions of mobility and looking after myself, except for mobility among patients aged 8 to 11 years, which showed poor agreement. All other dimensions showed poor agreement across all age groups.

Table 4 presents the level of agreement between patient- and proxy-reported outcomes on EQ-VAS and EQ-5D-Y-3L scores. ICCs for EQ-VAS scores indicated poor agreement across all age groups. In contrast, agreement on EQ-5D-Y-3L scores was moderate overall and among patients aged 8 to 11 years, except in the 12 to 17-year age group, where poor agreement was observed.

Table 5 presents EQ-5D-Y-3L scores by ECOG performance status and the observed effect sizes. Across all age groups, mean scores were consistently higher among patients with ECOG < 2 compared to those with ECOG ≥ 2, in both patient- and proxy-reported data. These differences were statistically significant in all comparisons ($p < 0.05$), with p-values less than 0.01 in most subgroups, except for the patient-reported data among children aged 8 to 11 years ($p = 0.03$). The comparisons were analyzed using the Mann-Whitney U test due to non-normal distribution of the data, as assessed by the Shapiro-Wilk test. Large effect sizes (Cohen's d ≥ 0.8)[30] were observed across all subgroups, particularly among children aged 8 to 11 years in both reporting versions. Notably, the effect size from proxy reports in the overall sample was greater than that from patient reports, indicating stronger known-groups discrimination.

Discussion

This is the first study to assess agreement between self- and proxy-reported health utility using the EQ-5D-Y-3L among Thai pediatric cancer patients. The instrument was found to be practical and straightforward to administer, with most participants able to complete it within a few minutes and without difficulty.

Table 5 Mean EQ-5D-Y-3L Scores by ECOG Performance Status and Effect Size *(Cohen's d)*

Age Group	ECOG < 2 (n = 98) Mean ± SD	ECOG ≥ 2 (n = 11) Mean ± SD	*P*-value[a]	Effect Size *(Cohen's d)* (95%CI)
Patient Report				
All patients (n = 109)	0.96 ± 0.07	0.79 ± 0.17	<0.01	2.06 (0.78-3.80)
Aged 8-11 years (n = 58)	0.97 ± 0.05	0.73 ± 0.25	0.03	3.28 (2.08-4.45)
Aged 12-17 years (n = 51)	0.95 ± 0.08	0.83 ± 0.10	<0.01	1.36 (0.51-2.20)
Proxy Report				
All patients (n = 109)	0.95 ± 0.09	0.68 ± 0.22	<0.01	2.50 (-0.94-1.93)
Aged 8-11 years (n = 58)	0.95 ± 0.08	0.64 ± 0.24	<0.01	3.27 (2.08-4.44)
Aged 12-17 years (n = 51)	0.94 ± 0.09	0.71 ± 0.22	<0.01	2.01 (1.11-2.89)

Note: ECOG < 2 refers to patients with ECOG scores of 0 or 1, while ECOG ≥ 2 includes scores of 2, 3 or 4. Effect sizes were calculated using *Cohen's d*, with pooled standard deviations computed from group-specific SDs. *P*-value[a] represent the results of Mann-Whitney *U* tests comparing ECOG < 2 and ECOG ≥ 2 within each age group, due to non-normal distribution as assessed by the Shapiro-Wilk test.

Agreement varied by dimension and age group.[29] Moderate agreement was observed in mobility and looking after myself, especially among older children, while other dimensions showed consistently low agreement. EQ-VAS scores also showed poor agreement,[28] with low ICCs across age groups. This unexpectedly low agreement may be due to the unstructured nature of the VAS scale, which allows differing interpretations between children and caregivers. In particular, younger children may base their ratings on immediate feelings, whereas proxies may rely more on their clinical observations.

For known-groups validity, both patient- and proxy-reported scores were significantly lower among children with poorer ECOG status. Both patient- and proxy-reported scores were significantly lower among children with poorer ECOG status (ECOG ≥ 2). Overall, proxy reports yielded slightly higher effect sizes than patient reports (2.50 vs. 2.06), suggesting stronger known-groups discrimination. The largest differences occurred among children aged 8 to 11 years, where patient reports showed the highest effect size (Cohen's d = 3.28), followed closely by proxy reports (Cohen's d = 3.27), indicating strong discriminatory power in this age group.

Ceiling effects above 15% were found in patient reports for children aged 8 to 11 years (51.7%) and proxy reports in the same group (36.2%). No floor effects exceeded 15% in any subgroup. These high ceiling effects suggest that the EQ-5D-Y-3L may have limited sensitivity in detecting differences in health status among individuals with relatively good perceived health. Consequently, the EQ-5D-Y-3L may underestimate subtle impairments in quality of life among pediatric oncology patients, especially those with better functional status. These concerns have also been noted in adult populations, leading the EuroQol Group to develop the EQ-5D-5L as an improved version to address high ceiling effects.[5] For children, the EQ-5D-Y-5L was later introduced to improve sensitivity. However, to date, no official Thai version of the EQ-5D-Y-5L is available, and it has not yet been widely implemented in pediatric research in Thailand. [5]

These results are consistent with previous studies that reported limited agreement between children and proxies. Shiroiwa, et al.[14] found low adjusted agreement in the pain and emotional dimensions among healthy Japanese children, with PABAK values of 0.15 and 0.12, respectively. In our study, the weighted kappa values in these dimensions were higher but still remained low.

Similar to the study by Sun, et al.[11] on Chinese pediatric inpatients with hematologic malignancies, our findings supported the known-groups validity of the EQ-5D-Y-3L based on ECOG performance status. Sun,

et al. reported significantly lower utility and EQ-VAS scores among patients with ECOG scores of 2 or 3. In our study, a similar pattern was observed, with large effect sizes indicating meaningful differences in EQ-5D-Y-3L scores between patients with ECOG 0-1 and those with ECOG ≥ 2 in both self- and proxy-reported data.

The study by Sun, et al.[11] focused on pediatric inpatients, whereas our study included both inpatients and outpatients, with a broader representation of patients who were either on or off treatment. This wider inclusion resulted in greater variation in health utility. In particular, patients who were off-treatment had higher health utility, contributing to the higher mean health utility scores observed in our study. In Sun, et al.'s study, the mean health utility score was 0.88±0.10, and the EQ-VAS score was 85.8±15.1. In comparison, our study yielded a slightly higher mean health utility score (0.94±0.10 from patient-reported data and 0.92±0.13 from proxy-reported data), but a lower mean EQ-VAS score (81.78±20.77 from patient-reported data and 85.23±13.19 from proxy-reported data). These differences reflect variations in patient composition and treatment status across study settings. These findings were further supported by subgroup comparisons in our study. Within both on-treatment and off-treatment groups, proxy-reported EQ-5D-Y-3L scores tended to be lower than patient-reported scores, with a statistically significant difference observed only in the on-treatment group ($p < 0.01$). The lower proxy-reported scores in the on-treatment group may capture aspects of distress or impairment not fully perceived by the children themselves, underscoring the value of proxy reports during active treatment.

When examining the distribution of responses, floor effects were minimal across all subgroups, suggesting that the instrument adequately captured poor health states. In contrast, ceiling effects exceeded 15% in most subgroups, particularly among off-treatment patients. The only exception was the EQ-VAS scores from proxy reports, where the ceiling effect remained below the threshold. This suggests that the EQ-5D-Y-3L may have limited ability to reflect small differences in health-related quality of life among children with relatively good health.

Consistent with our findings on EQ-VAS scores, Sousa, et al.[12] studied children with cerebral palsy and mild to moderate functional impairment and found ICCs below 0.60, reflecting poor to moderate agreement. For EQ-5D-Y dimensions, they used Cohen's kappa and found uniformly poor agreement (kappa < 0.20). In contrast, our study employed weighted kappa and showed higher agreement in several dimensions, particularly in the subgroup analysis: mobility among older children and looking after myself among younger children. This may reflect that most participants in our study were cognitively intact and able to self-report their health status.

In contrast, Bray, et al.,[13] who conducted their study among mobility-impaired children, reported a strong correlation between child and proxy EQ-5D-Y scores (rs = 0.665, $p = 0.026$) and a similarly strong but non-significant correlation for EQ-VAS scores (rs = 0.545, $p = 0.054$), suggesting a discrepancy between correlation strength and statistical significance. By comparison, although agreement levels in our study were lower, both EQ-5D-Y and EQ-VAS scores showed consistent patterns with statistical significance.

This study includes several strengths, such as the use of both self- and proxy-reported EQ-5D-Y-3L data, age stratification and analysis of known-groups validity and ceiling/floor effects, all of which enhance the methodological rigor of our research approach. Limitations include its cross-sectional design, which restricts conclusions about changes over time or treatment responsiveness. Additionally, as the sample was drawn from tertiary hospitals, generalizability may be limited. Future longitudinal studies are recommended to better assess temporal changes and intervention effects.

Conclusion

Proxy-report versions of the EQ-5D-Y-3L are recommended for assessing health utility scores among Thai pediatric cancer patients. Proxy reports are appropriate

across both age groups. Among children aged 8 to 11 years, proxy reports showed large effect sizes comparable to those of patient reports, with moderate agreement and lower ceiling effects. Among those aged 12 to 17 years, proxy reports demonstrated better known-groups discrimination based on ECOG status. Regarding EQ-VAS scores, poor agreement across all subgroups supports the use of proxy reporting, as it showed lower ceiling effects than those of patient-reported scores.

Acknowledgments

This study was supported by the Asia Pacific Oncology Pharmacy Association (APOPA). The authors thank all participants for their time and cooperation in completing the health-related quality of life questionnaires. The authors also appreciate the assistance and support provided by the hospital staff during data collection.

References

1. *World Health Organization. CureAll framework: Who global initiative for childhood cancer: increasing access, advancing quality, saving lives. Geneva: World Health Organization; 2021.*
2. *Brinkman TM, Recklitis CJ, Michel G, Grootenhuis MA, Klosky JL. Psychological Symptoms, Social Outcomes, Socioeconomic Attainment, and Health Behaviors Among Survivors of Childhood Cancer: Current state of the literature. J Clin Oncol. 2018;36:2190-7.*
3. *Age-standardized rate (world) per 100,000, incidence and mortality, both sexes, age (0-19), in 2022 Thailand [Internet]. 2022 [cited 1 August 2024]. Available from: https://gco.iarc.who.int/en.*
4. *Sakthong P. Health-related quality of life. 2nd rev. ed: Chulalongkorn University Press; 2019.*
5. *EuroQol Research Foundation. EQ-5D-Y user guide [Internet]. EuroQol Research Foundation; 2020 [Available from: https://euroqol.org/publications/user-guides.*
6. *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;19:875-86.*
7. *Sakthong P, Sonsa-Ardjit N, Sukarnjanaset P, Munpan W. Psychometric properties of the EQ-5D-5L in Thai patients with chronic diseases. Qual Life Res. 2015;24:3015-22.*
8. *Kim TH, Jo MW, Lee SI, Kim SH, Chung SM. Psychometric properties of the EQ-5D-5L in the general population of South Korea. Qual Life Res. 2013;22:2245-53.*
9. *Tran BX, Ohinmaa A, Nguyen LT. Quality of life profile and psychometric properties of the EQ-5D-5L in HIV/AIDS patients. Health Qual Life Outcomes. 2012;10:132.*
10. *Guideline Development Working group. Guideline for health technology assessment in Thailand updated edition: 2019.*
11. *Sun Y, Zhou HJ, Shen A, Wu B, Wang W, Luo N, et al. A cross-sectional study evaluating health-related quality of life of Chinese pediatric patients with hematological malignancies using EQ-5D-Y. Front Public Health. 2022;10:1050835.*
12. *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;115:541-6.*
13. *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;10:377.*
14. *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;28:3093-105.*
15. *Rosner B. Fundamentals of biostatistics. 7th ed. Boston: Brooks/Cole, Cengage Learning; 2011.*
16. *Guideline for health technology assessment in Thailand. 2nd edition, 2013.: Wacharin Printing; 2014.*
17. *Chaiyakledkaew U. Guideline for health technology assessment in Thailand. 1st ed: The Graphico Systems Co., Ltd.; 2009.*
18. *Dewilde S, Roudijk B, Tollenaar NH, Ramos-Goñi JM. An EQ-5D-Y-3L value set for Belgium. Pharmacoeconomics. 2022;40:169-80.*
19. *Yang Z, Jiang J, Wang P, Jin X, Wu J, Fang Y, et al. Estimating an EQ-5D-Y-3L value set for China. Pharmacoeconomics. 2022;40:147-55.*
20. *Kreimeier S, Mott D, Ludwig K, Greiner W. EQ-5D-Y value set for Germany. Pharmacoeconomics. 2022;40:217-29.*
21. *Rencz F, Ruzsa G, Bató A, Yang Z, Finch AP, Brodszky V. Value set for the EQ-5D-Y-3L in Hungary. Pharmacoeconomics. 2022;40:205-15.*
22. *Fitriana TS, Roudijk B, Purba FD, Busschbach JJV, Stolk E. Estimating an EQ-5D-Y-3L value set for Indonesia by mapping the DCE onto TTO Values. Pharmacoeconomics. 2022;40:157-67.*
23. *Shiroiwa T, Ikeda S, Noto S, Fukuda T, Stolk E. Valuation survey of EQ-5D-Y based on the international common protocol: development of a value set in Japan. Med Decis Making. 2021;41:597-606.*
24. *Prevolnik Rupel V, Ogorevc M, Greiner W, Kreimeier S, Ludwig K, Ramos-Goni JM, et al. EQ-5D-Y value set for Slovenia. Pharmacoeconomics. 2021;39:463-71.*
25. *Ramos-Goñi JM, Oppe M, Estévez-Carrillo A, Rivero-Arias O, Wolfgang G, Simone K, et al. Accounting for unobservable preference heterogeneity and evaluating alternative anchoring approaches to estimate country-specific EQ-5D-Y value sets: A case study using Spanish preference data. Value Health. 2022;25:835-43.*

26. *Roudijk B, Sajjad A, Essers B, Lipman S, Stalmeier P, Finch AP. A value set for the EQ-5D-Y-3L in the Netherlands. Pharmacoeconomics. 2022;40:193-203.*
27. *Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34-42.*
28. *Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155-63.*
29. *Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. 3rd ed. Hoboken, NJ: Wiley; 2013.*
30. *Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988. p 384.*