

Cognitive debriefing and the validity of the Malay EQ-5D-3L in transfusion-dependent thalassemia patients from Sabah, Malaysia

AiCh'I Liew¹, Asrul Akmal Shafie^{2*}, Bei Ying Tan²

Article Info

Received date: 20 Mar 2024
Accepted date: 20 June 2024
Published date: 30 June 2024

Keywords: cognitive debriefing, construct validity, EQ-5D-3L, health-related quality of life (HRQOL), reliability, known-group validation, validation

ABSTRACT

Background: Divergences in health-related quality of life (HRQOL) questionnaire responses may be attributed to multi-ethnic variation. **Objective:** The aim of this research is to examine the construct validity of the EQ-5D-3L in the context of Sabah, with Kadazan-Dusun constituting the preponderance of transfusion-dependent thalassemia (TDT) cases. **Method:** Independently or with the help of their careers, TDT patients who had received iron chelators answered a series of questionnaires. Along with the established group validity, test-retest reliability and correlations between HRQOL instruments were ascertained. **Results:** The study included 332 patients in all, 173 of whom completed the reliability analysis. EQ-5D-3L clarity and comprehension were observed in the majority of patients. The EQ-5D-3L is moderately correlated with the SF-36 and PedsQL. Reliability across domains is moderate to strong, with proxy reporting outperforming self-reporting. **Conclusion:** The results provide preliminary evidence of the reliability and validity of the EQ-5D-3L in Sabah population.

INTRODUCTION

Transfusion-dependent thalassaemia (TDT) is a type of thalassemia that requires subtypes [1] of regular blood transfusion, iron chelation therapy and preventive measures of disease-related complications, to ensure their optimal growth and survival [2, 3].

Health-Related Quality of Life (HRQOL) is an important parameter to monitor their psychological and physical health statuses. In the health technology assessment, is indicated by health utilities, which can be used to calculate quality-adjusted life years (QALY) to assist in healthcare management decision-making. EQ-5D-3L is an instrument that captures the HRQOL through five health domains, which can transform responses to weighted health indices or utility values [4]. EQ-5D-3L has been validated in patients with TDT (5) in Malaysia over the past decade. Sabah has the highest number of registered cases

of thalassemia, accounting for 22.72% (n = 1814) of all cases registered in the Malaysia Thalassaemia Registry in 2018 (N =

8684) [6]. Most patients with thalassemia are of Kadazan-Dusun descent and are between 10 and 14.9 years old on average. The Kadazan-Dusun were distinguished as an indigenous people of Borneo, with a documented cultural legacy by the United Nations Educational, Scientific, and Cultural Organisation (UNESCO) since 2004. They have retained their indigenous language, cultural heritage, and customs.

Malay is officially designated as Malaysia's national language. Nevertheless, previous research has documented a lack of mastery of the national language in Sabah. Furthermore, the variability of EQ-5D responses may be attributed to disparities in socioeconomic status and population demographics [7]. Ensuring an instrument's reliability and validity in producing consistent outcomes across diverse health conditions among study populations is of utmost importance, and questionnaire

*Correspondence: aakmal@usm.my; aichiliew@yahoo.com
DOI: <https://doi.org/10.52494/maljpharmv10108>

¹Clinical Research Center, Hospital Seberang Jaya, Penang, Malaysia
²Discipline of Social & Administrative Pharmacy, School of Pharmaceutical, Sciences, Universiti Sains Malaysia, Malaysia

validation is especially critical in a country as multicultural and multiethnic as Malaysia. Thus, the purpose of this study was to evaluate the construct validity and reliability of the EQ-5D-3L in patients with TDT in Sabah and to analyse the comprehension of the EQ-5D-3L items in terms of clarity and relevance.

METHODS

Study design and participants

Patients with TDT or their caregivers were recruited in the thalassemia clinics of five divisions (West Coast division, Kudat division, interior division, Sandakan division and Tawau division) based on their geographical locations. Quota sampling was used in the patient recruitment.

The inclusion criteria for completing the surveys were patients aged >3 years, a diagnosis of TDT, and patients who had received treatment with iron chelation therapy for at least 6 months. Patients who defaulted treatment or regular follow-up for more than 3 months and had impaired cognitive function were excluded from the study. The sample size was determined using the formula for estimating population prevalence [8]. Based on an estimated population of 1272 in Sabah [9], a 95% confidence level, a 5% accuracy rate, and a 15% dropout rate, 340 patients were needed as the final sample size.

Study tool

a. EQ-5D-3L: This survey comprises of five single-item health dimensions, yielding 243 health states (each mapped to a utility score using a utility function reflective of the general Malaysian population), and visual analogue scales (EQ-VAS). The EQ-VAS ranges from 0 to 100, with 0 representing the "worst health condition" and 100 representing the "best health condition."

b. SF-36: A standardized questionnaire used to measure patient health on eight different dimensions: physical functioning (PH), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role-emotional (RE), and mental health (MH). It is comprised of items that provide respondents with options regarding their health perception. For instance, the dimension of physical functioning has ten components. Three responses—"limited a lot," "limited a little," and "not limited at all" are coded 1, 2, and 3, respectively. The ten coded responses are added up to create a final score ranging from 10 to 30. This score was converted to a scale of 0 to 100.

c. PedsQL Inventory™: This questionnaire assesses four health functions: 1) physical functioning (eight items), 2) emotional functioning (five items), 3) social functioning (five

items), and 4) school functioning (4 items). There are two types of reporting: child self-report (5-18 years old) and parent proxy report (2-18 years old). The items are reverse-scored and linearly translated to a 0-100 score (0 = 100, 1 = 75, 2 = 50, 3 = 25 and 4 = 0) using a five-point response scale. Better HRQOL is indicated by a higher score.

Data collection

Patients were selected by reviewing the appointment lists at transfusion treatment centres or haematology clinics in seven hospitals. The goals of the study and the information on the patient information sheet were explained to patients and their carers before the interview (for patients under the age of 18). Upon agreeing to participate in this study, consent and assent forms were signed and dated. The patients were given a questionnaire to answer, and the interview session was conducted after they had completed the questionnaires.

Cognitive debriefing

The Malaysian Malay EQ-5D-3L was used in this investigation. Cognitive debriefing was performed on patients with TDT in Sabah who had varied socio-demographic backgrounds and local dialects to determine the readability and ease of completion of this version of the questionnaire. Willis et al. have suggested that seven to ten interviews are sufficient to confirm patient comprehensibility of an item (10). Thus, ten of seven study sites were chosen to show the range of severity levels in the community across different areas to confirm patient comprehensibility.

The cognitive debriefing techniques employed were adapted from Collins [11] utilising several probes, such as response latency timing, card sorts, vignettes, paraphrase, and confidence ratings, which were initially created by psychologists [12]. Paraphrasing and general probing methods were used in this study. The general probing method [13] was used to question the interviewees as to whether the items were comprehensible and clear and to assess the ease of the completion, relevance, and clarity of the items in the questionnaires. Each item was demonstrated to the subject by the researcher, and the respondents were requested to suggest a better or alternate set of item descriptors in their own words. A second researcher compiled field notes that included details of difficulties and respondent behaviours. The study objectives, contents of the patient information sheet, and contact information of the investigators were provided to the patients and their caregivers (in the case of patients aged <18 years) before the interview. If the subjects were willing to participate, teleconference meetings were held in the transfusion-treatment facility. Parents will assist paediatric patients (under 18 years old) during the interview.

Construct validity and reliability studies

Consenting patients with TDT who visited the thalassemia clinic during their routine transfusion were provided with three questionnaires for HRQOL reporting: EQ-5D-3L, SF-36, and PedsQL Inventory (PedsQL). A set of forms consisting of patient information sheets, informed consent forms, assent forms, Malay version of the HRQOL survey forms (EQ-5D-3L, PedsQL, and SF-36), and a data collection form that captured the participants' sociodemographic data and medical history. The PedsQL was to be completed independently by children, whereas EQ-5D-3L and SF-36 were to be completed independently by adult respondents. PedsQL were obtained from the enrolled children, with caregiver completion if the child was unable to do so without severely compromising data quality. Data collection for assessing reliability was performed prospectively with selected patients (n=170) from the total number of recruited patients, with an interval of 2 weeks.

Statistical Analyses

Cognitive debriefing

The content analysis method was used to analyse cognitive debriefing. A total of ten patients, chosen deliberately to encompass the various levels of severity within the population, will be interrogated using cognitive debriefing techniques. The EQ-5D-3L questionnaire items' relevance, clarity, and simplicity of completion for participants were assessed. The analysis would be conducted via teleconference from the patients' homes or within the hospital facility, under the guidance of a single researcher. The patients' demographics were presented as percentages, standard deviations, and mean values.

Construct validity

Stata (v14.0) was used to analyse the data. The patients' general characteristics were reported as percentages, means, and standard deviations. The total and summary HRQOL scores were reported as the mean and standard deviation for HRQOL. The value set used in the present study was identical to that of adults (14). Based on the sample (paediatric and adult) and survey source, the EQ-5D-3L domain responses, index, and EQ-VAS score were summarised (self-reported or proxy-reported). The chi-square test was used to determine the statistical significance of any reported differences between the groups. The responses were categorised as either 'no problem' or 'problems'.

The hypothesised relationship between the EQ-5D-3L, the SF-36(15) and PedsQL(5) items, that were validated in the thalassemia population, was used to examine construct validity. By comparing the domains in EQ-5D-3L with those in PedsQL and SF-36, convergent and discriminant validity were assessed

[16] using the Spearman rank correlation and Pearson correlation. The hypothesised correlations are shown in Appendices-Table 3 and 4. The Spearman rank correlation was used to test the known-group validity.

Hypotheses: Attributes of the healthcare dimension

The participants would be more likely to report the following problems: mobility if they were older (known-group approach [16]) or had lower physical function (PF) scores (convergent validity [16]); for EQ-5D, self-care and usual activities if they had lower role-physical (RP) scores; and for EQ-5D, anxiety/depression if they were female [17] or had lower mental health, PF, RP, bodily pain, general health, social functioning, role-emotional, mental health scores.

Hypotheses: Overall score concern

The participants had lower overall EQ-5D-3L utility and EQ-VAS scores if they were older, female, and had more side effects [15, 18].

Hypotheses: Correlation between EQ-5D-3 L and PedsQL

A higher score in the PedsQL domains indicates better health conditions, whereas a higher score in EQ-5D-3L indicates poorer health states. EQ-5D-3L and PedsQL correlated negatively and moderately (5, 19).

Hypotheses: Correlation between EQ-5D-3L and SF-36

Higher scores in the SF-36 domains indicate better health conditions, whereas higher scores in the EQ-5D-3L domains indicate poorer health states. EQ-5D-3L negatively and moderately (20) to strongly (21) correlated with SF-36.

Reliability

Test-retest reliability was based on the correlation coefficient (r) and internal consistency by Cronbach's alpha. Cohen's kappa or Pearson correlation was used for the interrater reliability of the self-report and proxy-report. The inter-rater agreement was cross-sectionally analysed at each time point (baseline and 2 weeks). The weighted kappa statistics was used to assess the levels of agreement for the EQ-5D-3L domains. These methods allowed for the measurement of the agreement between two or more raters. The kappa scores ranged from -1 to +1 and can be interpreted as follows: <0, no agreement; 0-0.2, slight; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; \square 0.80, almost perfect agreement. Kappa is frequently employed to evaluate the agreement between conditions and a reference standard, but it has been criticised for its strong reliance on prevalence. In order to address this constraint, this research adopted a prevalence-adjusted and bias-adjusted kappa (PABAK).

RESULTS

Demographic Respondents

A total of 332 people responded—187 were adults and 145 were children—including 173 patients who were part of the reliability test. Table I provides a summary of the socio-demographic characteristics of the 332 respondents who presented to the blood transfusion centre in Sabah in May 2022. The Kadazan-Dusun people composed most of the respondents (64.45%, n = 214).

Cognitive Debriefing (Content validity) of EQ-5D-3L in the TDT Population

In the thalassemia clinic, 10 individuals were interviewed through a video conference. A total of 10 individuals were recruited, of whom 60% were adults (19–34 years old) and 40% were paediatric patients (7–15 years old). Eight responders were women, with the patients' mothers providing the most of the responses. The Kadazan-Dusun respondents composed 50% of the sample, and secondary school graduates composed 60% (Appendix-Table1). The debriefing task required 47 minutes to

Table I. Demographic characteristics by age group

Variables	N(%)	
	Children (n= 187)	Adult (n = 145)
Sex		
Male	113 (60.43)	55 (37.93)
Female	74 (39.57)	90 (62.07)
Ethnicity	4 (2.14)	8 (5.52)
Malay		
Chinese	4 (2.14)	7 (4.83)
Kadazan-Dusun	120 (64.17)	94 (64.83)
Bajau	10 (5.35)	5 (3.45)
Murut	25 (13.37)	4 (2.76)
Suluk	4 (2.14)	3 (2.07)
Rungus	6 (3.20)	4 (2.76)
Others	14 (7.49)	20 (13.79)
Presence of Iron Overload Complications		
Cardiac Disease	2 (1.07)	4 (2.76)
Diabetic	0 (0)	11 (7.59)
Hypogonadism	0 (0)	6 (4.14)
Liver Disease	4 (2.14)	6 (4.14)
Others, Specify	3 (1.60)	10 (6.90)
No Complications	177 (94.65)	106 (73.10)
Hypogonadism & Liver Disease	0 (0)	1 (0.69)
Cardiac Disease & Other Disease	1 (0.53)	0 (0)
Liver Disease & Other Disease	0 (0)	1 (0.69)
Current Iron Therapy		
Deferroxamine, DFO	23 (12.30)	21 (14.48)
Deferrasirox, DFX	74 (39.57)	1 (0.69)
Deferiprone, DFP	61 (32.62)	67 (46.21)
Deferroxamine, DFO + Deferiprone, DFP	15 (9.02)	54 (37.24)
DFO + DFP + DFX	0 (0)	1 (0.69)
DFO+DFX	7 (3.74)	1 (0.69)
DFP+DFX	6 (3.21)	0 (0)
Haematinic (iron, Vitamin B ₁₂ , folate)	1 (0.53)	0(0)

complete, while the patients took 5.3 minutes on average to complete the EQ-5D-3L. No overt signs of cognitive impairments were found throughout the interviews; cognitive impairment was not investigated.

Comprehensibility/Clarity

Most patients who participated in the cognitive interviews understood the domains well by providing their own interpretations in their native tongue and connecting the domains to their therapeutic experiences. Most patients also discovered that their experiences applied to the EQ-5D-3L. However, two respondents were uncertain about the clarity and relevance of certain domains. The paediatric patient's proxy was uncertain about the dimension of anxiety and depression: 'Unsure of what depression is' and 'I believe that anxiety refers to worrying, whereas depression is feeling sad' (Sungai, 9, patient's mother).

A 15-year-old patient was unsure of three dimensions, interpreting 'I have some problems in walking about' as 'tired while walking' and 'having difficulty and not comfortable during shower' as 'I cannot wash or dress by myself.' She presumed that the pain dimension refers to pain while getting a Desferal injection, while 'I have extreme pain or discomfort' refers to 'Feeling uncomfortable while sleeping'. This was mainly because of her age and educational level. The wording used was too formal for a teenager to understand. All her interpretations were basically from her daily experiences.

Construct Validation Study

a. Testing hypotheses concerning levels within particular attributes (Appendix-Table 2b)

Compared with the paediatric patients, the adult patients substantially more frequently experienced the pain/discomfort and anxiety/depression domains. The patients' mobility was greatly impacted by the presence of complications, serious adverse events (SAE), and gender. Furthermore, the patients' usual activities and pain/discomfort were considerably impacted by the presence of complications.

b. Testing hypotheses concerning overall scores (Appendix-Table 2a – Table 2b.)

Priori 8 hypotheses were fulfilled in the EQ-5D-3L index and EQ-VAS, with four reaching statistical significance on the EQ-5D-3L index, while three reached statistical significance in EQ-VAS. The mean (standard deviation [SD]) EQ-5D-3L index and EQ-VAS scores were significantly higher in the children [(0.93 (0.16); 80.75 (88.27)] than in the adults [0.85 (0.20);

Table II. Reliability analysis with test-retest method of categorical variables of EQ-5D-3L

2 nd Visit	EQ-5D-3L descriptive questions-1st visit responses					
	1 st Visit					
		Mobility	Selfcare	Usual Activity	Pain	Depression
Mobility	PABAK	0.757[#]	0.792[#]	0.699[#]	0.422[#]	0.503[#]
	Kappa	0.164[*]	0.081^ρ	0.106	0.054	-0.011
	BI	0.052	0.052	0.023	-0.150	-0.075
	PI	0.844	0.844	0.815	0.642	0.717
Selfcare	PABAK	0.861[#]	0.7919[#]	0.827[#]	0.480[#]	0.607[#]
	Kappa	0.110	0.274[#]	0.175[*]	0.031	0.006
	BI	-0.023	0.093	-0.052	-0.225	-0.150
	PI	0.919	0.884	0.890	0.717	0.792
Usual Activity	PABAK	0.769[#]	0.942[#]	0.757[#]	0.526[#]	0.538[#]
	Kappa	0.232[#]	0.076^ρ	0.298[#]	0.231[*]	0.074
	BI	0.058	0.017	0.029	-0.145	-0.069
	PI	0.838	0.960	0.809	0.636	0.711
Pain	PABAK	0.711[#]	0.780[#]	0.676[#]	0.538[#]	0.549[#]
	Kappa	0.180[*]	0.057	0.179[*]	0.283[*]	0.160^ρ
	BI	0.087	0.098	0.058	-0.116	-0.041
	PI	0.809	0.879	0.780	0.607	0.682
Depression	PABAK	0.838[#]	0.850[#]	0.757[#]	0.503[#]	0.676[#]
	Kappa	0.322[#]	0.050[*]	0.157^ρ	0.147[*]	0.288[#]
	BI	0.023	0.064	-0.006	-0.179	-0.104
	PI	0.809	0.913	0.786	0.671	0.746

Note. PABAK, Prevalence Adjusted Bias Adjusted Kappa; BI, Bias Index; PI, Prevalence Index; Strong (Kappa >0.6), moderate (Kappa = 0.01-0.60), and weak (Kappa <0.01).

Bolded values indicate statistically significant outcomes at *P<.01, -at [#] P<.0001, - at ^ρP<.05 and not significant at P>.05.

79.70 (19.15)]. Besides the attribution of age, a similar trend was also reflected in the patients with SAE and complications. The mean (SD) EQ-5D-3L index was 0.74 (0.25) for the patients with complications and 0.91 (0.17) for those without complications. The mean (SD) EQ-VAS score was 68.80 (27.81) for the patients with complications and 81.46 (19.90) for the patients without complications.

c. Testing the correlational hypothesis between PedsQL and EQ-5D-3L (Appendix-Table 3)

Most EQ-5D-3L domains showed statistically significant moderate correlation with the PedsQL dimension scales except between self-care and the physical health summary score (rs = -0.172) and emotional functioning score (rs = -0.213). The

Table III. Interrater reliability analysis with test-retest method of self-report vs proxy-report for EQ-5D-3L

2 nd Visit	EQ-5D-3L descriptive questions-1st visit responses										
	1 st Visit										
		Mobility		Selfcare		Usual Activity		Pain		Depression	
	SR	PR	SR	PR	SR	PR	SR	PR	SR	PR	
Mobility	PABAK	0.750[*]	0.763[*]	0.775[*]	0.807[*]	0.725[*]	0.677[*]	0.300[*]	0.527[*]	0.325[#]	0.656[*]
	Kappa	0.223^ρ	0.100	0.147^ρ	0.000	0.091	0.121	0.064	0.031	-0.037	0.016
	BI	0.050	0.054	0.088	0.097	0.063	-0.011	-0.225	-0.086	-0.163	0.000
	PI	0.825	0.860	0.863	0.903	0.838	0.796	0.550	0.720	0.613	0.807
Selfcare	PABAK	0.875[*]	0.850[*]	0.950[*]	0.936[*]	0.900[*]	0.763[*]	0.325[*]	0.613[*]	0.450[*]	0.742[*]
	Kappa	0.259[#]	-0.038	0.487[*]	0.000	0.310[#]	0.110	0.024	0.048	0.039	-0.051
	BI	-0.038	-0.011	0.000	0.032	-0.025	-0.075	-0.313	-0.151	-0.250	-0.065
	PI	0.913	0.925	0.950	0.968	0.925	0.860	0.638	0.785	0.700	0.871
Usual Activity	PABAK	0.750[*]	0.785[*]	0.775[*]	0.785[*]	0.725[*]	0.785[*]	0.451	0.634[*]	0.375[*]	0.763[*]
	Kappa	0.223^ρ	0.239[#]	0.147^ρ	0.000	0.091	0.440[*]	0.265[#]	0.272[#]	0.0403	0.355[#]
	BI	0.050	0.065	0.088	0.108	0.063	0.000	-0.225	-0.075	-0.163	0.011
	PI	0.825	0.850	0.863	0.893	0.838	0.785	0.550	0.710	0.613	0.796
Pain	PABAK	0.575[*]	0.828[*]	0.650[*]	0.785[*]	0.600[*]	0.742[*]	0.325[#]	0.720[*]	0.350[*]	0.720[*]
	Kappa	0.015	0.391[*]	0.085	0.000	0.036	0.328[*]	0.144	0.443[*]	0.081	0.238[#]
	BI	0.113	0.065	0.150	0.108	0.125	0.000	-0.163	-0.075	-0.100	0.011
	PI	0.763	0.850	0.800	0.893	0.775	0.785	0.488	0.710	0.550	0.796
Anxiety/ depression	PABAK	0.476[*]	0.893[*]	0.450[*]	0.808[*]	0.451[*]	0.763[*]	0.350[#]	0.634[*]	0.575[*]	0.763[*]
	Kappa	0.134	0.591[*]	0.039	0.000	0.076	0.355[*]	0.112	0.183^ρ	0.324[#]	0.205^ρ
	BI	0.213	0.054	0.250	0.097	0.225	-0.011	-0.250	-0.118	-0.188	-0.032
	PI	0.663	0.860	0.700	0.903	0.675	0.796	0.575	0.753	0.638	0.839

Note. PABAK, Prevalence Adjusted Bias Adjusted Kappa; BI, Bias Index; PI, Prevalence Index; Strong (Kappa >0.6), moderate (Kappa = 0.1-0.6), and weak (Kappa <0.1). SR- self-reported; PR-proxy reported

Bolded values indicate statistically significant outcomes at *P<.0001, -at □ P<.01, - at □P<.05 and not significant at P>.05

correlation between usual activities and the social functioning score ($r_s = -0.181$) was weak.

d. Testing correlational hypothesis between EQ-5D-3L and SF-36 (Appendix- Table 4)

Generally, EQ-5D-3L correlated moderately with SF-36. The self-care domain is weakly correlated to the role-physical domain in SF-36. The usual activities domain is weakly correlated with the social functioning and role-emotional domains. The mobility domain is significantly weakly correlated with bodily pain and general health. The pain/discomfort and anxiety/depression domains in EQ-5D-3L showed a statistically significant moderate correlation to most domains in SF-36.

Reliability

PABAK showed moderate (range, 0.1–0.6) - to - strong reliability with test-retest between the categorical domains, as shown in Table 2. The test-retest reliability(r) was moderate for the EQ-5D-3L and EQ-VAS scores (0.368 vs. 0.469, $p < 0.001$; Appendix Table 5).

Table III shows that proxy reporting was generally stronger (range, 0.527–0.936) than self-reporting (range, 0.300–0.950) in PABAK. The PABAK values of self-reporting for the pain and depression domains were generally lower than those for other domains (range, 0.3–0.6).

DISCUSSIONS

In this study, we assessed the construct validity of the Malay version of the EQ-5D-3L in a TDT population in Sabah. The purpose of this qualitative cognitive debriefing study was to discover EQ-5D-3L interpretation problems in TDT populations in Sabah. Most patients understood the EQ-5D-3L and regarded it to be pertinent to their therapeutic experience.

Our cognitive debriefing findings revealed that because children and adolescents have varied levels of cognitive ability and the concordance of proxy reports, methodological concerns regarding the relevance of the domains, and the appropriateness of the EQ-5D-3L in children and the proxy report might have arisen. A systematic review showed that adults, children, and adolescents perceived and valued health differently (22), consistent with our 15-year-old respondent. Childhood and adolescent health dimensions may vary in adults because of differences in cognitive capacities and life experiences. The utility score might have been influenced by the wording used by the children, adults, and adolescents, despite the parameters remaining constant. To accommodate youngsters, EQ-5D, and EQ-5D-Y were developed.

In addition to cognitive function at different ages, national language intelligibility in Sabah and West Malaysian populations has limited linguistic and cultural commonalities. In various parts of Sabah, people speak distinct dialects. According to research, in Sabah, among those with comparable levels of education, language intelligibility was much lower in developing areas than in well-developed areas [23]. Thus, the comprehension of the questionnaire may be impacted by Malay language proficiency. However, the EQ-5D-3L questions took the participants 5.3 (3.94) minutes to complete. This succinct reply might indicate that the questionnaire was simple to understand. However, it might only suggest a few possibilities: Incorrect responses were submitted because the respondents (a) failed to read the instructions, (b) did not comprehend the questions and required examples to illustrate them clearly, and (c) believed that the pre-determined response options did not apply to them.

Construct validity was confirmed by the statistically significant fulfilment of most a priori hypotheses at the attribute and overall scale levels for the Malay language in the TDT population in Sabah, especially in the context of their own native languages and various sociocultural backgrounds.

Although EQ-5D-3L was not originally intended for use in paediatrics, it has been validated by Brussoni et al. in the paediatric primary injury population based on self-reports and parent reports [24]. Proxy reporting is unavoidable because of the respondents' young ages. However, Rand et al. showed no direct interchangeability with self-reports [25]. Parent proxy has been recognised as a viable option for assessing child health status [26], although some studies suggest discrepancies between parent and child responses where parents' ratings tended to be lower than their children's rating in the short term but converged in the long term [27].

Children and adolescents may be less able to reliably assess their own health, necessitating proxy reporting and self-reporting of their health in accordance with their cognition and age. Adolescents appeared to be more self-aware and self-reflective than pre-pubescent youngsters, indicating a qualitative shift in the nature of thinking. Adolescents develop the ability to remember more complex information and hence think more strategically. The substantial changes in identity, self-consciousness, and cognitive flexibility that occur with the transition from childhood to adulthood [28] posed methodological difficulties in evaluating their HRQOL [29]. Thus, the methodologies of these analyses must be adequate for decision-making for various populations, including children and adolescents with various disorders.

Owing to the lack of EQ-5D-Y utility tariffs in the Malaysian version of the instrument, its use was restricted. As a result, the availability of the Malaysian EQ-5D-3L tariff, which has been

validated in the thalassemia population, has made it more appealing to be utilised in Malaysia. The fact that the EQ-5D-3L has only three levels of response makes it simpler for adolescents.

The magnitude of the kappa coefficient represents the proportion of reliability greater than that expected by chance and is influenced by prevalence, bias, and nonindependence of ratings. The prevalence index is elevated; consequently, kappa is decreased [30]. In light of the substantial occurrence of favourable ratings in this research, which could potentially influence the kappa value, it was determined that PABAK was a suitable approach for generating an unbiased kappa value, independent of the conditions under which the initial ratings were acquired.

However, Hoehler et al. criticised the use of PABAK, given that the effects of bias and prevalence are informative in kappa magnitude and thus should not be adjusted and disregarded [31]. Consequently, the PABAK coefficient is uninformative on its own, as it pertains to a hypothetical scenario in which neither prevalence nor bias effects are present. As suggested by Byrt et al., PABAK and kappa were reported in this study to give an indication of the expected effects of prevalence and bias along with the true value of the kappa [32].

Limitations

This study has certain limitations. Most respondents reported being in good health. This may impair the ability to discern power differentials when hypotheses are not substantiated. In addition, laboratory test results were not accessible for comparison with the patients' clinical outcomes, treatment-related complications, and SAE reports. Comparing the applicability of proxy-reporting and self-reporting is particularly crucial.

However, self-reported and proxy-reported scores were not compared in this study; therefore, the consistency of the scores is unknown. Boonchooduang et al. showed that the proxy-reported HRQOL was lower than the self-reported HRQOL[3]. Lastly, cognitive function was not tested in this study.

CONCLUSIONS

This is the first study to examine the reliability and validity of the EQ-5D-3L Malay version in patients with TDT, encompassing adults and paediatrics. The study results demonstrated that EQ-5D-3L is an appropriate, reliable, and valid instrument for measuring HRQOL in patients with TDT in Sabah. Further elaboration on the clinical groups encompassed within the intricate socio-cultural context of Sabah would strengthen the foundation of the findings of this study for future economic health research.

CONFLICT OF INTEREST

None

ACKNOWLEDGEMENTS

Not available.

REFERENCE

- [1] Cappellini M-D, Cohen A, Porter J, Taher A, Viprakasit V. Guidelines for the management of transfusion dependent thalassaemia (TDT): Thalassaemia International Federation Nicosia, Cyprus; 2014.
- [2] Ibrahim HM, Hassan A, George E, Goh AI. Management of Transfusion Dependent Thalassaemia. Putrajaya: Health Technology Assessment Section; 2009.
- [3] Boonchooduang N, Louthrenoo O, Choeprasert W, Charoenkwan P. Health-related quality of life in adolescents with thalassaemia. *Pediatric Hematology and Oncology*. 2015;32(5):341-8. <https://doi.org/10.3109/08880018.2015.1033795>
- [4] Foundation ER. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument. 2018. p. 34.
- [5] Shafie AA, Chhabra IK, Hui Yi JW, Mohammed NS, Ibrahim HM. Validity of the Malay EQ-5D-3L in the Malaysian Transfusion-Dependent Thalassaemia Population. *Value in Health Regional Issues*. 2021 2021/05/01;24:47-56. <https://doi.org/10.1016/j.vhri.2020.08.003>
- [6] Malaysian Thalassaemia Registry 2018. Malaysian Thalassaemia Registry Committee, Ministry of Health, Malaysia, 2019.
- [7] Buchholz I, Janssen MF, Kohlmann T, Feng Y-S. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *Pharmacoeconomics*. 2018;36(6):645-61. <https://doi.org/10.1007/s40273-018-0642-5>
- [8] Glaziou P. Sample size for a prevalence survey, with finite population correction 2005; Available from: <http://sampsize.sourceforge.net/jface/>.
- [9] Pauzy L, Esa E, Mokhri N, Yusoff Y, Jamaludin N. Thalassaemia Distribution Based on Screening Programs in the Population of the East Malaysian State of Sabah. *J Blood Disord Transfus* 9: 395. 2018;163(87):250. <https://doi.org/10.4172/2155-9864.1000395>
- [10] Willis GB. *Cognitive interviewing: A tool for improving questionnaire design*: sage publications; 2004.
- [11] Collins D. Pretesting survey instruments: an overview of cognitive methods. *Quality of life research*. 2003;12(3):229-38.
- [12] Forsyth BH, Lessler JT. *Cognitive laboratory methods: A taxonomy. Measurement errors in surveys*. 2004:393-418. <https://doi.org/10.1002/9781118150382.ch20>
- [13] Dewolf L, Koller M, Velikova G, Johnson C, Scott N, Bottomley A, et al. EORTC Quality of Life Group translation procedure. 2009. Accessed from link: http://groups.eortc.be/qol/downloads/translation_manual_2009.pdf
- [14] Shafie AA, Thakumar AV. Multiplicative modelling of EQ-5D-3L TTO and VAS values. *The European Journal of Health Economics*. 2020;21(9):1411-20. <http://dx.doi.org/10.1016/j.canlet.2016.08.016>
- [15] Sobota A, Yamashita R, Xu Y, Trachtenberg F, Kohlbray P, Kleinert D, et al. Quality of life in thalassaemia: a comparison of SF-36 results from the thalassaemia longitudinal cohort to reported literature and the US norms. *American journal of hematology*. 2011;86(1):92. <http://doi.org/10.1002/ajh.21896>
- [16] Fayers PM, Machin D. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*: John Wiley & Sons; 2013.
- [17] Maier W, Gänsicke M, Gater R, Rezaki M, Tiemens B, Urzúa RF. Gender differences in the prevalence of depression: a survey in primary care. *Journal of affective disorders*. 1999;53(3):241-52. [https://doi.org/10.1016/S0165-0327\(98\)00131-1](https://doi.org/10.1016/S0165-0327(98)00131-1)
- [18] Seyedifar M, Dorkoosh FA, Hamidieh AA, Naderi M, Karami H, Karimi M, et al. Health-related quality of life and health utility values

- in beta thalassemia major patients receiving different types of iron chelators in Iran. *International journal of hematology-oncology and stem cell research*. 2016;10(4):224.
- [19] López-Bastida J, López-Siguero JP, Oliva-Moreno J, Vázquez LA, Aranda-Reneo I, Reviriego J, et al. Health-related quality of life in type 1 diabetes mellitus pediatric patients and their caregivers in Spain: an observational cross-sectional study. *Current medical research and opinion*. 2019. <https://doi.org/10.1080/03007995.2019.1605158>
- [20] Kularatna S, Senanayake S, Gunawardena N, Graves N. Comparison of the EQ-5D 3L and the SF-6D(SF-36) contemporaneous utility score in patients with chronic kidney diseases in Sri Lanka: a cross-sectional survey. *BMJ Open*. 2019;9:9. <http://dx.doi.org/10.1136/bmjopen-2018-024854>
- [21] García-Gordillo MÁ, del Pozo-Cruz B, Adsuar J, Cordero-Ferrera J, Abellán-Perpiñán J, Sánchez-Martínez F. Validation and comparison of EQ-5D-3L and SF-6D instruments in a Spanish Parkinson's disease population sample. *Nutrición Hospitalaria*. 2015;32(6):2808-21.
- [22] Thorrington D, Eames K. Measuring health utilities in children and adolescents: a systematic review of the literature. *PLoS one*. 2015;10(8):e0135672. <https://doi.org/10.1371/journal.pone.0135672>
- [23] Banker J, Banker E. The Kadazan/Dusun language. *Languages of Sabah: A survey report: Pacific Linguistics*; 1984.
- [24] Brussoni M, Kruse S, Walker K. Validity and reliability of the EQ-5D-3L™ among a paediatric injury population. *Health and quality of life outcomes*. 2013;11:1-9.
- [25] Rand S, Caiels J. Using proxies to assess quality of life: a review of the issues and challenges. 2015.
- [26] MacArthur C, Dougherty G, Pless IB. Reliability and validity of proxy respondent information about childhood injury: an assessment of a Canadian surveillance system. *American journal of epidemiology*. 1997;145(9):834-41.
- [27] Gabbe BJ, Simpson PM, Sutherland AM, Palmer CS, Butt W, Bevan C, et al. Agreement between parent and child report of health-related quality of life: impact of time postinjury. *Journal of Trauma and Acute Care Surgery*. 2010;69(6):1578-82. <https://doi.org/10.1097/TA.0b013e3181f8fd5f>
- [28] Rutter M, Rutter M. *Developing minds: Challenge and continuity across the life span*: Basic books; 1993.
- [29] 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;30:697-712. [https://doi.org/1170-7690/12/0008-0697/\\$49.95/0](https://doi.org/1170-7690/12/0008-0697/$49.95/0)
- [30] Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ: British Medical Journal*. 1992;304(6840):1491. <https://doi.org/10.1136/bmj.304.6840.1491>
- [31] Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *Journal of clinical epidemiology*. 2000;53(5):499-503. [https://doi.org/10.1016/S0895-4356\(99\)00174-2](https://doi.org/10.1016/S0895-4356(99)00174-2)
- [32] Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *Journal of clinical epidemiology*. 1993;46(5):423-9. [https://doi.org/10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)

APPENDIX

Appendix-Table1. Cognitive debriefing – Interpretation, clarity and relevance of the EQ-5D-3L (n=10)

Item	Clarity ¹		Relevance ²		Cognitive debriefing item interpretation and findings
	Yes, n	No/ Not sure, n	Yes, n	No/Not sure, n	
1 Under each heading, please tick the ONE box that best describes your health TODAY	10	0	10	0	All participants have good comprehension on questionnaire instructions.
2 Mobility	10	0	10	0	All participants have good comprehension on mobility.
3 I have no problems in walking about	10	0	10	0	There were some participants had different understanding of the moderate level of mobility.
4 I have some problems in walking about	9	1	10	0	
5 I am confined to bed	10	0	10	0	Mobility questions do not apply to wheelchair patients.
6 Self-care	10	0	10	0	Most of the participants comprehended “self-care” as “self-management” in personal hygiene and some in daily activities such as food and medication taking.
7 I have no problems with self-care	10	0	10	0	
8 I have some problems washing or dressing myself	10	0	10	0	
9 I am unable to wash or dress myself	9	1	10	0	
10 Usual Activities (e.g. work, study, housework, family or leisure activities)	10	0	10	0	The activities of usual activities dimension were generally interpreted as “active activities” and “daily activities”.
11 I have no problems with performing my usual activities.	10	0	10	0	
12 I have some problems with performing my usual activities.	10	0	10	0	
13 I am unable to perform my usual activities	10	0	10	0	
14 Pain/ Discomfort	10	0	10	0	Participants had correct interpretations on “pain or discomfort” dimension.
15 I have no pain or discomfort	10	0	10	0	Patients well aligned their treatment and disease experience (i.e., pain while getting Desferral injection, body ache, headache and poor health condition) to pain and discomfort dimension.
16 I have moderate pain or discomfort	9	1	10	0	
17 I have extreme pain or discomfort	10	0	9	1	
18 Anxiety/ Depression	9	1	10	0	Most of the participants interpreted a range of emotions in relation to their thalassemia in “Anxiety or depression” dimension, e.g., feeling uneasy, worry, stress, and sadness. A patient commented should change “depression” to “sadness” for better understanding and easier for some patients.
19 I am not anxious or depressed	10	0	10	0	
20 I am moderately anxious or depressed	10	0	10	0	
22 We would like to know how good or bad your health is TODAY	10	0	10	0	All participants reported no difficulties with EQ-5D-3L health rating instructions.
23 This scale is numbered from 0 to 100	10	0	10	0	All respondents could use this scale to rate their health status. They could comprehend the best and the worst health condition that they could imagine.
24 100 means the best health you can imagine	10	0	10	0	
25 0 means the worst health you can imagine	10	0	10	0	
26 Mark an X on the scale to indicate how your health is TODAY	10	0	10	0	Self-rating scale was considered relevant to their health condition reflection.
27 Now, please write the number you marked on the scale in the box below.	10	0	10	0	
28 YOUR HEALTH TODAY	10	0	10	0	
29 EQ-VAS (visual log)	10	0	10	0	

1, Do you understand what this means? 2, Do you think the response options are appropriate

Appendix-Table 2a. Known-group validity based on a priori hypothesis (EQ-5D-3L)(Demographic)

Demographic Characteristics	EQ-5D-3L domains										EQ-5D-3L Index Mean (SD)	EQ VAS Mean (SD)
	Mobility n (%)		Self-care n (%)		Usual Activity n (%)		Pain/discomfort n (%)		Anxiety/ depression n (%)			
	No Problem	Problem	No Problem	Problem	No Problem	Problem	No Problem	Problem	No Problem	Problem		
Age												
2-4	13 (92.9)	1 (7.1)	13 (92.9)	1 (7.1)	14 (100.0)	0 (0.0)	13 (92.9)	1 (7.1)	13 (92.9)	1 (7.1)	0.948 (0.14)	85.36 (17.37)
5-7	11 (91.7)	1 (8.3)	12 (100.0)	0 (0.0)	10 (83.3)	2 (16.7)	9 (75.0)	3 (25.0)	10 (83.3)	2 (16.7)	0.883 (0.18)	75.83 (20.21)
8-12	36 (94.7)	2 (5.3)	36 (94.74)	2 (5.3)	35 (92.1)	3 (7.9)	32 (84.2)	6 (15.8)	36 (94.7)	2 (5.3)	0.932 (0.18)	81.29 (22.81)
13-18	27 (96.4)	1 (3.6)	26 (92.9)	2 (7.1)	25 (89.3)	3 (10.7)	25 (89.3)	3 (10.7)	25 (89.3)	3 (10.7)	0.926 (0.16)	79.36 (25.86)
> 18	66 (91.7) ^Ω	6 (8.3) ^Ω	69 (95.8) ^Ω	3 (4.1) ^Ω	63 (87.5) ^Ω	9 (12.5) ^Ω	50 (69.4) ^Ω	22 (30.6) ^Ω	58 (80.6) ^Ω	14 (19.4) ^Ω	0.850 (0.20)	79.70 (19.15)
P-value	0.92		0.88(F)		0.624(F)		0.10		0.26(F)		0.04^Ω	0.65
Age Group												
Pediatric	90 (94.7)	5 (5.3)	90 (94.7)	5 (5.3)	87 (91.6)	8 (8.4)	81 (85.3)	14 (14.7)	87 (91.6)	8 (8.4)	0.926 (0.16)	80.747 (88.27)
Adult	63 (91.3)	6 (8.7)	66 (95.7)	3 (4.3)	60 (87.0)	9 (13.0)	48 (69.6)	21(34.4)	55 (79.7)	14 (20.3)	0.851 (0.20)	79.696 (19.15)
P-value	0.39		0.79		0.34		0.02^Ω		0.03^Ω		0.01	0.00*
Sex												
Female	77 (97.5)	2 (2.5)	79 (100.0)	0 (0.0)	74 (93.7)	5 (6.3)	62 (78.5)	17 (21.5)	67 (84.8)	12 (15.2)	0.906 (0.14)	81.37 (19.21)
Male	76 (89.4)	9 (10.6)	77 (90.6)	8 (9.4)	73 (85.9)	12 (14.1)	67 (78.8)	18 (21.2)	75 (88.2)	10 (11.8)	0.883 (0.22)	79.32 (22.54)
P-value	0.04^Ω		0.01(F)*		0.10		0.96		0.52		0.44	0.53

^Ω, self-report; (F) Fisher Exact Test, Bolded values indicate statistically significant outcomes. *P<.01; [#] P<.001; ^ΩP<.05

Appendix-Table 2b. Known-group validity based on a priori hypothesis (EQ-5D-3L)(Clinical characteristic)

	EQ-5D-3L domains										EQ-5D-3L Index Mean (SD)	EQ VAS Mean (SD)
	Mobility n (%)		Self-care n (%)		Usual Activity n (%)		Pain/discomfort n (%)		Anxiety/ depression n (%)			
	No Problem	Problem	No Problem	Problem	No Problem	Problem	No Problem	Problem	No Problem	Problem		
Presence of Complications												
No	142 (95.3)	7 (4.7)	142 (95.3)	7 (4.7)	138 (92.6)	11 (7.4)	122 (81.9)	27 (18.1)	131 (87.9)	18 (12.1)	0.909 (0.17)	81.46 (19.90)
Yes	11 (73.3)	4 (26.7)	14 (93.3)	1 (6.7)	9 (60.0)	6 (40.0)	7 (46.7)	8 (53.3)	11 (73.3)	4 (26.7)	0.744 (0.25)	68.80 (27.81)
P-value	0.01(F)^Ω		0.54(F)		0.00(F)[#]		0.00(F)*		0.121(F)		0.00[#]	0.03^Ω
No. of Iron Chelation Therapy Agents												
Monotherapy	118 (93.6)	8 (6.4)	120 (95.2)	6 (4.8)	115 (91.3)	11 (8.7)	99 (78.6)	27 (21.4)	112 (88.9)	14 (11.1)	0.903 (0.18)	81.08 (21.19)
Dual Therapy	35 (92.1)	3 (7.9)	36 (94.7)	2 (5.3)	32 (84.2)	6 (15.8)	30 (79.0)	8 (21.1)	30 (79.0)	8 (21.1)	0.862 (0.20)	77.74 (20.23)
P-value	0.49(F)		0.59(F)		0.17(F)		0.96		0.12		0.23	0.39
SAE												
No	144 (94.7)	8 (5.3)	145 (95.4)	7 (4.6)	138 (90.8)	14 (9.2)	121 (79.6)	31 (20.4)	132 (86.8)	20 (13.2)	0.903 (0.17)	80.37 (21.24)
Yes	9 (75.0)	3 (25.0)	11 (91.7)	1 (8.3)	9 (75.0)	3 (25.0)	8 (66.7)	4 (33.3)	10 (83.3)	2 (16.7)	0.786 (0.27)	79.50 (17.79)
P-value	0.04(F)^Ω		0.46(F)		0.11(F)		0.29(F)		0.67(F)		0.00[#]	0.03^Ω
Iron chelation administration route												
SC	18 (94.7)	1 (5.3)	18 (94.7)	1 (5.3)	17 (89.5)	2 (10.5)	17 (89.5)	2 (10.5)	17 (89.5)	2 (10.5)	0.936 (0.17)	76.00 (30.58)
PO	103 (93.6)	7 (6.4)	105 (95.5)	5 (4.6)	101 (91.8)	9 (8.2)	85 (77.3)	25 (22.7)	98 (89.1)	12 (10.9)	0.901 (0.18)	81.11 (19.56)
SC+ PO	32 (91.4)	3 (8.6)	33 (94.3)	2 (5.7)	29 (82.9)	6 (17.1)	27 (77.1)	8 (22.9)	27 (77.1)	8 (22.9)	0.851 (0.20)	80.11 (19.28)
P-value	0.89(F)		0.87(F)		0.33(F)		0.57(F)		0.22(F)		0.08	0.62

^Ω, self-report; (F) Fisher Exact Test, Bolded values indicate statistically significant outcomes. *P<.01; [#] P<.001; ^ΩP<.05

Appendix-Table 3. Hypothesized and actual correlation coefficients among EQ VAS, EQ-5D-3L dimensions and PedsQL Inventory™ dimension scales

Dimension	EQ-5D-3L domains					
	Mobility	Self-care	Usual activities	Pain/ Discomfort	Anxiety/Depression	EQVAS
Hypothesized correlation	Moderate (-)	Moderate (-)	Moderate (-)	Moderate (-)	Weak (-)	Moderate (-)
Physical Health Summary Score	-0.339[#]	-0.172[#]	-0.309[*]	-0.462[#]	-0.358[#]	0.448[#]
Hypothesized correlation	Weak (-)	Weak (-)	Weak (-)	Moderate (-)	Moderate (-)	Moderate (-)
Emotional Functioning Score	-0.333[*]	-0.213[#]	-0.181[#]	-0.435[#]	-0.354[#]	0.499[#]
Hypothesized correlation	Weak (-)	Weak (-)	Weak (-)	Moderate (-)	Moderate (-)	Moderate (-)
Social Functioning Score	-0.303[#]	-0.149	-0.194	-0.419[#]	-0.353[#]	0.426[#]
Hypothesized correlation	Weak (-)	Weak (-)	Moderate (-)	Weak (-)	Weak (-)	Moderate (-)
School Functioning Score	-0.278[#]	-0.142	-0.331[#]	-0.454[#]	-0.302[*]	0.551[#]

Note. Strong (coefficient >0.5), moderate (coefficient 0.3-0.5), and weak (coefficient <0.3). (-) indicated a negative correlation; EQ VAS, visual analogue scale portion of the EQ-5D-3L; n.s., not significant P>.05
 Bolded values indicate statistically significant outcomes. *P<.01; # P<.001; [#]P<.05

Appendix-Table 4. Hypothesized and actual correlation coefficients among EQ VAS, EQ-5D-3L dimensions and SF-36 dimension scales

SF-36 Domain	EQ-5D-3L domains						
	Dimension	Mobility	Self-care	Usual activities	Pain/ Discomfort	Anxiety/Depression	EQVAS
Hypothesized correlation		Strong	Strong	Strong	Moderate	Strong	Strong
Physical Functioning		-0.315[*]	-2.000	-0.351[*]	-0.323[*]	-0.259[#]	0.359[*]
Hypothesized correlation		Moderate	Moderate	Moderate	Moderate	Moderate	Strong
Role-physical		-0.330[*]	-0.271[#]	-0.365[*]	-0.399[#]	-0.405[#]	0.296[#]
Hypothesized correlation		Strong	Strong	Strong	Strong	Strong	Strong
Bodily Pain		-0.248[#]	-0.300[#]	-0.357[*]	-0.485[#]	-0.314[*]	0.352[*]
Hypothesized correlation		Strong	Strong	Strong	Strong	Strong	Moderate
General Health		-0.282[#]	-0.221	-0.319[*]	-0.518[#]	-0.241[#]	0.415[#]
Hypothesized correlation		Strong	Strong	Strong	Strong	Strong	Strong
Vitality		-0.155	-0.118	-0.210	-0.371[*]	-0.393[#]	0.366[*]
Hypothesized correlation		Strong	Strong	Strong	Strong	Strong	Strong
Social Functioning		-0.216	-0.357[*]	-0.254[#]	-0.455[#]	-0.473[#]	0.148
Hypothesized correlation		Moderate	Moderate	Moderate	Moderate	Moderate	Strong
Role-emotional		-0.171	-0.258[#]	-0.270[#]	-0.347[*]	-0.421[#]	0.023
Hypothesized correlation		Moderate	Moderate	Moderate	Moderate	Moderate	Strong
Mental Health		0.013	-0.181	-0.170	-0.225	-0.413[#]	0.179
Repeated Health Transition		0.016	0.049	0.222	0.318[*]	0.155	-0.556[#]

FS indicates Functioning Score; HSS, Health Summary Score, Bolded value indicates significant outcomes *P<.01; # P<.001; [#]P<.05

Appendix-Table 5. Reliability analysis with test-retest method of EQ VAS and EQ-5D-3L Index.

	Mean (SD)	Median	EQ-5D-3L Index		r	P-value
			Min	Max		
1 st Visit	0.644(0.09)	0.632	-0.102	0.756	0.368	<0.001
2 nd Visit	0.691(0.12)	0.740	-0.207	0.740		
			EQ VAS			
1 st Visit	82.075(16.81)	85.000	15.000	100.000	0.469	<0.001
2 nd Visit	83.740(16.95)	90.000	10.000	100.00		

Cronbach's Alpha α=0.43

Note. EQ VAS, visual analogue scale portion of the EQ-5D-3L; Strong (coefficient >0.5), moderate (coefficient 0.3-0.5), and weak (coefficient <0.3). Bolded values indicate statistically significant outcomes at P<.05 by Pairwise Correlation Coefficient; n.s., not significant P>.05