bjh research paper

Validation and reliability of a disease-specific quality of life measure (the TranQol) in adults and children with thalassaemia major

Robert J. Klaassen,[1] Nicholas Barrowman,[1] Manuela Merelles-Pulcini,[2] Elliott P. Vichinsky,[3] Nancy Sweeters,[3] Melanie Kirby-Allen,[4] Ellis J. Neufeld,[4] Janet L. Kwiatkowski,[5] John Wu,[6] Linda Vickars,[7] Victor S. Blanchette,[2] Melissa Forgie,[8] Robert Yamashita,[9] Durhane Wong-Rieger[10] and Nancy L. Young[11]

[1]Department of Pediatrics, Children's Hospital of Eastern Ontario, Ottawa, [2]Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada, [3]Department of Hematology/Oncology, Children's Hospital & Research Center Oakland, Oakland, CA, [4]Department of Pediatrics, Boston Children's Hospital, Boston, MA, [5]Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA, [6]Department of Pediatrics, British Columbia Children's Hospital, [7]Department of Internal Medicine, St. Paul's Hospital, Vancouver, BC, [8]Department of Internal Medicine, Ottawa Hospital, Ottawa, ON, Canada, [9]Department of Liberal Studies, California State University, Long Beach, CA, USA, [10]Anemia Institute for Research & Education, Toronto, and [11]School of Rural and Northern Health, Laurentian University, Sudbury, ON, Canada

Received 21 January 2013; accepted for publication 17 September 2013
Correspondence: Dr Robert Klaassen, Division of Hematology/Oncology, Department of Pediatrics, Children's Hospital of Eastern Ontario, 401 Smyth, Ottawa K1H 8L1, ON, Canada.
E-mail: rklaassen@cheo.on.ca

Summary

This study aimed to demonstrate the validity, reliability and responsiveness of a new disease-specific quality of life (QoL) questionnaire for children and adults with thalassaemia major, the Transfusion-dependent QoL questionnaire (TranQol). 106 participants (51 adults and 55 children) were recruited from six North American thalassaemia treatment centres with a mean age of 20·7 years (standard deviation [SD] 9, range 7–51 years). The mean total TranQol score was 71 (SD 17, 32–97) on a scale of 0–100. Patients with co-morbidities had significantly lower scores (63 vs. 75, $P = 0{\cdot}001$). TranQol scores showed substantial agreement ($P < 0{\cdot}001$) with the Health Utilities Index Mark 3 (all patients, $r = 0{\cdot}65$), the Pediatric QoL (children, $r = 0{\cdot}77$) and the Short Form (36) physical (adults, $r = 0{\cdot}69$) and mental summary scores ($r = 0{\cdot}76$). In the subgroup who rated their QoL as better, there was a 4·0 point (SD 9·0) improvement in TranQol scores, from baseline of 67·1–71·1 one week later ($P = 0{\cdot}008$). Test–retest reliability was excellent (intra-class correlation coefficient, 0·93). The TranQol was valid, with acceptable correlation for all administered measures and was reliable and responsive to change. The TranQol can be incorporated into future studies of thalassaemia major.

Keywords: thalassaemia, quality of life, validation studies, reproducibility of results, questionnaires.

Beta thalassaemia is an inherited haemoglobin disorder with a carrier frequency as high as 20% in the Mediterranean region (Weatherall & Clegg, 2001). Thalassaemia major is characterized by impaired production of haemoglobin, and affected individuals need red blood cell transfusions every 2–5 weeks for survival (Olivieri *et al*, 1994). Chronic transfusions result in an accumulation of total body iron, which is lethal if not treated. Patients develop complications due to both the disease process and the subsequent iron overload. Examples include: (i) severe anaemia with haemolysis and extramedullary haematopoesis, (ii) organomegaly, (iii) cardiac disease, including heart failure, arrhythmias and pericarditis, (iv) endocrinopathies, including bone disease, diabetes, abnormal growth and development, and (v) liver cirrhosis

First published online 3 November 2013
doi:10.1111/bjh.12631

(Rachmilewitz & Giardina, 2011). All of these issues can adversely impact on patient quality of life (QoL).

Quality of life has been defined by the World Health Organization (WHO) as the 'net consequence of life characteristics on a person's perception of their position in life, in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards and concerns.'(WHO, 1993, 1995) Health-related quality of life (HRQL) refers to the specific impact a disease or illness may have on an individual's overall QoL. Very few studies have assessed HRQL in thalassaemia, and those studies used generic measures (not specific to thalassaemia) yet were still able to demonstrate a significant decrease in HRQL in this patient population (Jansen *et al*, 2003; Pakbaz *et al*, 2005; Ismail *et al*, 2006). Disease-specific QoL measures allow for a more relevant, sensitive and complete assessment of the impact of an intervention in both the clinical and research settings, and can facilitate informed choices among patients (Eiser & Morse, 2001). Thus, disease-specific QoL measures provide valuable additional information about both the clinical condition and the associated treatment in conditions such as thalassaemia.

Currently, the only validated disease-specific tool available for transfused patients is the 'Satisfaction with Iron Chelation Therapy' (SICT). The SICT was developed to assess patient satisfaction with iron chelation for those with thalassaemia, sickle cell disease, and myelodysplastic syndrome (MDS) who required chelation (Rofail *et al*, 2009). It was found to have acceptable validity and reliability, and has been used in two large multi-centre clinical trials that included more than 400 patients with thalassaemia major (Rofail *et al*, 2010; Porter *et al*, 2012). In the landmark EPIC (evaluation of patient's iron chelation) study, an open label study of deferasirox, the SICT demonstrated a significant improvement in side effects, acceptance, and decreased burden of iron chelation therapy compared to baseline (Porter *et al*, 2012). However, the SICT is limited to assessing satisfaction and adherence, and was not designed as a tool for measuring QoL.

Due to the lack of instruments designed to measure QoL for thalassaemia patients, our team developed the TranQol (transfusion-dependent quality of life) questionnaire. The TranQol is a disease-specific QoL measure developed in Canada using rigorous methodology (Guyatt *et al*, 1986; Juniper *et al*, 1996). The details of the tool's development, which involved 120 participants: 16 healthcare workers, 31 children and 30 adults with thalassaemia and 43 parents, have been published (Klaassen *et al*, 2013). Initially, we included patients with MDS, but it became apparent during the item generation phase of development that there were important differences in this patient population when compared to the thalassaemia patients. Therefore a decision was made to develop a separate tool for the MDS patients, and this tool was called the Quality of Life in Myelodysplasia Scale (QUALMS-1). Thus the TranQol evolved into a tool developed specifically for patients with thalassaemia major.

A valid, reliable, and responsive disease-specific HRQL measure is needed to evaluate thalassaemia major patients in both the research and clinical setting. The purpose of this study was to test the psychometric properties (how well it measures QoL) of the TranQol in a North American multicentred study of adults and children with thalassaemia major.

Methods

Recruitment of participants

Participants were recruited from six tertiary health care centres in five cities across North America, including: Boston, Oakland, Philadelphia, Toronto, and Vancouver. These sites were chosen as they provided a broad spectrum of patients and treatment approaches. Each institution obtained approval from their respective research ethics board or institutional review board before enrolling patients at their centre. Study recruitment included both paediatric (2–17 years of age, with parent proxy report for the younger children) and adult (18 years of age or older) patients. This report focusses on the patients aged 7 years or older who were able to read English at a grade 2 level and able to self-report. Patients with co-morbidity were included. The definition of co-morbidity was left to the discretion of the site investigators and included diabetes mellitus, splenectomy, osteoporosis, hypogonadism, and hypothyroidism, with the majority of these patients having multiple co-morbidities.

Participants completed a baseline (the time of enrolment) questionnaire package, including both the TranQol and the Health Utilities Index Mark 3 (HUI3), and either the Peadiatric QoL Inventory generic core module version 4.0 (PedsQL) for the paediatric participants or the Medical Outcomes Study short-form 36 version 2 (SF-36) for adult participants. Questionnaires were completed in the clinic during a routine blood transfusion. One week later, the participants were asked to complete a second set of the same questionnaires. This was done to assess responsiveness, as the impression of previously interviewed patients was that persons who had just received a transfusion tended to report an improved HRQL.

Two to five weeks later, a third set of the same questionnaires were completed to assess reliability. This aligned with the next routine blood transfusion for each patient. This time interval was selected so that the transient benefit of the transfusion would no longer be in effect, and the patients HRQL should have returned to baseline. Each patient was also asked to rate any change in their health since enrolment using a 7-point Likert-type scale, ranging from 'much better', to 'same' to 'much worse' (rating of change). All respondents were asked at the end of each survey to record the time that they spent completing the questionnaire and details regarding the involvement/help of others. Ten percent of the sample was double-entered to assess the rate and impact of data entry errors. A data entry error rate of <1% was considered acceptable.

Questionnaires

The TranQol is a disease-specific QoL measure for children and adults with thalassaemia major. It has four versions: (i) a child self-report, (ii) an adult self-report, (iii) a parent self-report (measuring the impact of the disease on the parent), and (iv) a parent proxy-report (measuring the child's QoL). The questionnaire length ranges from 29 (child) to 39 items (parent). The questions are grouped into four domains: physical health, emotional health, family functioning, and school and career functioning. The adult and parent self-report questionnaires include a fifth category on sexual activity which is only one item (Klaassen *et al*, 2013). The TranQol cannot be published in BJH due to copyright issues, but can be obtained on request from the authors.

In this study we used three questionnaires in addition to the TranQol: HUI3, PedsQL and SF-36. The HUI3, a generic multi-attribute preference-based tool for assessing HRQL, has gained widespread acceptance in both paediatric and adult patients (Torrance *et al*, 1996). The psychometric properties of this measure have been established in both children and adults with very good test-retest reliability, as well as good predictive and discriminant validity (Rizzo *et al*, 1998; Nixon *et al*, 1999; Sung *et al*, 2003). The PedsQL is a generic QoL measure for children that was initially derived for paediatric oncology patients and consists of a core module that can be used in both healthy and ill children. The PedsQL is feasible: it has no floor effects (scoring clustered at the lower limit) and low to moderate ceiling effects (scoring at the upper limit). There is acceptable consistency and good discriminant, clinical and construct validity (Varni *et al*, 2002, 2003; Connelly & Rapoff, 2005). The SF-36 is a generic, multipurpose, short form health survey with 36 questions. It yields an eight-scale profile of functional health and well-being scores, as well as a psychometric-based physical and mental health summary and a preference-based health utilities index. The survey has good internal consistency, inter-rater and test-retest reliability. Content and discriminant validity are good with reasonable responsiveness to change (Ware, 2004).

A priori *hypothesis and statistical analysis*

Statistical analysis was performed with SPSS version 20 (IBM Corporation, Software Group, Somers, NY, USA), and two-sided *P*-values <0·05 were deemed to be statistically significant. Data checks and cleaning were performed on all data. Summary and domain scores of all questionnaires were calculated as per each questionnaire's owner's manual, and represent the mean of all applicable responses. Summary scores for the TranQol were calculated by reverse scoring the positive response items and then using this equation: $100 \times (1 - [(\text{Sum of all scores} - \text{number of valid responses}) / (\text{valid responses} \times 4)])$.

The overall sample size was estimated to be 50 patients in each age group, based on the numbers required to provide adequate power for the reliability and validity testing. The assessment of reliability was based on the concordance between measurements, and required a minimum of 40 patients with stable disease to prove an intra-class correlation coefficient (ICC) that is statistically different from 0·6 if the true ICC for the measure is 0·8 (Donner & Eliasziw, 1987).

Construct validity was assessed using Pearson's correlation (or Spearman's if the distribution was found to deviate markedly from bivariate normality) between the TranQol scores and the HUI3, PedsQL and SF-36 scores using the data collected at baseline. Our a-priori hypothesis was that there would be a moderately strong correlation (0·4–0·6) between the TranQol and (i) HUI3 in all patients, (ii) PedsQL in the children, and (iii) SF-36 in the adults. Agreement was classified as recommended by Landis and Koch (1977), with a correlation of 0·0 indicating poor agreement, 0·0–0·20 indicating slight agreement, 0·21–0·40 indicating fair agreement, 0·41–0·60 indicating moderate agreement, 0·61–0·80 indicating substantial agreement and 0·81–1·0 indicating almost perfect agreement. Known groups validity was assessed by comparing clinical groups using *t*-tests (for two groups) or analysis of variance (for more than two groups). The hypothesis was that patients chelated using subcutaneous desferal would have significantly lower TranQol scores than patients receiving oral chelation or no chelation.

Internal consistency was determined using Cronbach's alpha. Test-retest reliability of the patient self-report was assessed using random effects ICC limited to those who rated their QoL as 'same' at the time of the third administration. The test-retest hypothesis was that scores in stable subjects (rating of change recorded as 'same') would be consistent within the same rater (ICC >0·7) at two time points (compare scores at subsequent transfusion to baseline).

To assess responsiveness, a *t*-test was used to determine if the mean change score was significantly different from 0 in those whose rating of change was 'a little better'. The responsiveness hypothesis was considered to be exploratory due to concerns about sample size: the change in scores of the TranQol at 1 week post-transfusion to baseline would be significantly higher for those whose rating of change was 'much better, better' or 'much worse, worse' compared to those patients rating their health as 'the same'. We compared the TranQol to the PedsQL, SF-36 and HUI3 to determine if there is improved responsiveness.

Results

One hundred and six patients self-completed the questionnaires, including 55 children (aged 7–17 years) and 51 adults enrolled between March 2010 and November 2011. A detailed description of the demographic characteristics is shown in Table I. Of particular note, the majority of both adults and children self-identified with Asian ethnic origin (53%) and the vast majority of the patients spoke English as

Table I. Sample characteristics by subgroup.

	Child $N = 55$	Adult $N = 51$	Total $N = 106$
Male sex, *N* (%)	27 (51)	16 (31)	43 (41)
Age, years (range)	13·4 (7–18)	28·6 (19–51)	20·7 (7–51)
Adult marital status, *N* (%)			
Divorced		1 (2)	1 (2)
Married		9 (18)	9 (18)
Never married		38 (75)	38 (75)
Other		3 (6)	3 (6)
Ethnicity, *N* (%)			
Caucasian	20 (36)	13 (26)	33 (31)
Black	0 (0)	3 (6)	3 (3)
Asian	26 (47)	30 (59)	56 (53)
Other	10 (18)	6 (12)	16 (15)
Transfusion frequency			
Every 3 weeks	16 (29)	35 (6)	51 (48)
Every 4 weeks	38 (69)	9 (18)	47 (44)
Other	1 (2)	6 (12)	7 (7)
Co-morbidities, Y (%)	4 (7)	37 (73)	41 (39)
Chelator, *N* (%)			
Desferal	15 (27)	27 (53)	42 (40)
Exjade	48 (87)	33 (65)	81 (76)
Ferriprox	0 (0)	8 (16)	8 (8)
Ferritin, pmol/l (range)	4017 (45–21, 857)	5186 (921–20,403)	4579 (45–21,857)
Hepatic iron concentration, mg/g dry tissue (range)	5·2 (0·1–27)	6·6 (0·3–42)	5·9 (0·1–42)
TranQol score, mean (range)	77 (38–97)	63 (32–93)	71 (32–97)

their first language (99%). TranQol scores at baseline ranged from a low of 32 to a high of 97 out of 100, indicating no concerns regarding ceiling or floor effect (clustering of scores at the lower or upper limit). Figure 1 shows the distribution of TranQol scores compared to those of the other measures.

Validity

Construct validity: does the tool actually measure quality of life? The TranQol self-report summary scores at baseline showed substantial agreement ($P < 0{\cdot}001$) with all of the comparative generic QoL measures, including: the HUI3 (all patients, $r = 0{\cdot}65$), the PedsQL (children, $r = 0{\cdot}77$) and the SF-36 physical health (adults, $r = 0{\cdot}69$) and mental health scores (adults, $r = 0{\cdot}76$). Table II compares the domains of the paediatric TranQol questionnaire with the corresponding domains of the PedsQL. We found moderate correlation for the school and family functioning domains with substantial agreement of both the emotional and physical health domains (all $P \leq 0{\cdot}001$). We then looked at the adult TranQol domains and compared them to the scales of the SF-36. There was substantial agreement of the physical health domain with all three relevant SF-36 scales (physical functioning, role-physical, bodily pain, all $P < 0{\cdot}001$). The SF-36 mental component summary did not correspond as clearly to the TranQol domains but still showed moderate to substantial correlation (all $P < 0{\cdot}001$) (Table III).

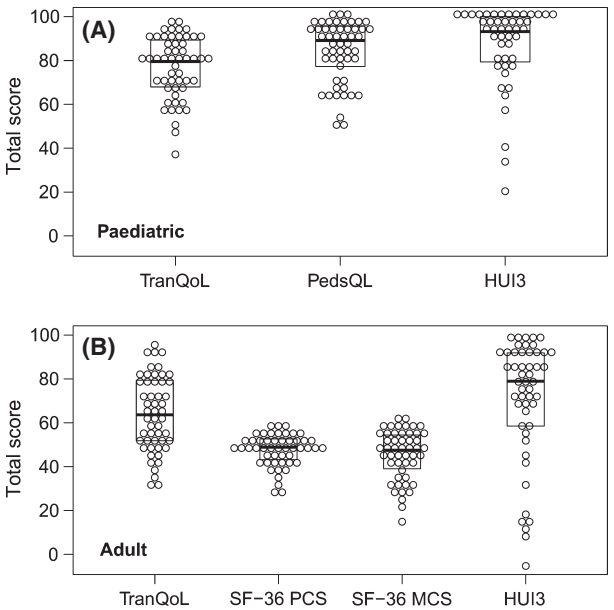

Fig 1. Beeswarm graphs of (**A**) paediatric baseline TranQol, PedsQL and HUI3 scores and (**B**) adult baseline TranQol, SF-36 physical component summary (PCS), SF-36 mental component summary (MCS) and HUI3 scores. The horizontal line represents the median, and the box indicates the 25–75 percentiles.

Known groups validity: does the tool demonstrate different scores for groups known to have different QoL? Patients with co-morbidities had significantly lower TranQol summary scores than patients who did not (63 vs. 75, $P = 0{\cdot}001$). In addition, adults reported significantly lower scores compared to the child report (63 vs. 77, $P < 0{\cdot}001$). We were not able to show a statistical difference between the different types of chelation.

Reliability

Internal consistency: do all the items in the tool measure the same concept (QoL)? Cronbach's alpha of all of the TranQol versions was good (child 0·84, proxy 0·89 and adult 0·96). Correlation of the individual items to the total score revealed a consistent negative correlation of one of the items: 'Thalassaemia positively affected my family…' with the total score in the child, adult and parent versions as well as a poor correlation in the proxy tool (<0·3). Because of this finding, we have deleted this item from all versions of the tool. There was no other consistent pattern noted with the other items.

Test-retest reliability: are the results consistent across time? Thirty-one adults (61%) and 29 children (53%) rated their QoL (rating of change) as 'same' between the baseline

Table II. Correlation of child TranQol and PedsQL domain scores at baseline.

TranQol, mean (SD)	Emotional health 66·4 (25·2)	Family functioning 77·8 (15·7)	School functioning 81·2 (17·8)	Physical health 79·6 (15·9)
PedsQL, mean (SD)	Emotional functioning 85·0 (17·4)	Social Functioning 89·7 (15·8)	School functioning 75·9 (18·6)	Physical functioning 86·2 (15·6)
Domain correlation, *r*	0·65	0·50	0·43	0·69

Table III. Adult TranQol and SF-36 domain scores at baseline.

TranQol, mean (SD)	Emotional health (EH) 62·1 (18·0)	Family functioning (FF) 65 (19·9)	School & career functioning (SF) 61·5 (27·7)		Physical health (PH) 64·7 (18·0)			
SF-36 summary, mean (SD)	Mental component summary 45·5 (11·3)				Physical component summary 47·6 (7·3)			
SF-36 scales, mean (SD)	Role-emotional 77·1 (25·3)	Mental health 69·0 (19·5)	Vitality 52·3 (20·5)	Social functioning 79·0 (21·9)	Physical functioning 82·3 (18·7)	Role-physical 75·4 (22·7)	Bodily pain 67·1 (23·9)	General health 52·2 (23·5)
Domain/scale correlation, *r*	EH 0·64 SF 0·68	EH 0·70		FF 0·62 SF 0·68	PH 0·71	PH 0·71	PH 0·69	

(during their initial blood transfusion) and the third administration (during the subsequent transfusion) of the questionnaires. The test-retest reliability for this group was excellent with almost perfect agreement (ICC 0·94 for adults and 0·89 for the children, both $P < 0·001$). We further analysed the reliability of the individual TranQol domains and found acceptable reliability in all domains (0·74–0·88, $P < 0·001$), with the only exception being the child physical domain, which was 0·63, $P < 0·001$.

Responsiveness: is the tool able to detect a meaningful change in QoL? In the week following their red blood cell transfusion 39 patients (37%) rated their QoL 'a little better', 'better' or 'much better'. Only three patients rated their QoL as a 'little worse' with no patients 'worse' or 'much worse'. Therefore the 'worse' or 'much worse' group was not analysed. The subgroup of patients whose rating of change was at least a 'little better' showed a statistically significant increase in their TranQol scores of 4·0 (standard deviation [SD] 9·0) from baseline (67·1–71·1, $P = 0·008$). When further subdivided, the 18 adult patients had a significant improvement of 6·0 (SD 9·5, $P = 0·016$), which was not seen in the SF-36 physical component summary ($P = 0·54$) or mental component summary ($P = 0·08$). The 21 children did not show a significant change in their scores ($P = 0·22$), however we did not find a statistically significant difference between the responsiveness of the child and adult self-reports ($P = 0·129$). Figure 2 shows the change in self-reported TranQol scores at 1 week post-transfusion according to the patients rating of change.

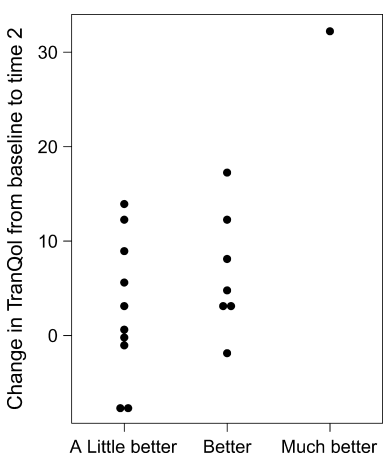

Fig 2. Changes in adult TranQol scores at 1 week post-transfusion by self-reported change in quality of life rating.

Discussion

Thalassaemia major is a unique condition that requires patients to receive life-long transfusions to survive. Frequent transfusions lead to iron overload and daily chelation for patients who, through government support, insurance or personal finances, are able to afford the financial burden of these medications. The cumulative effects of the disease impact the patient's QoL as they age. Multiple organ dysfunction, chronic pain and loss of physical fitness decrease the QoL into adulthood (Pakbaz *et al*, 2005). The distinct aspects of thalassaemia major require a disease-specific tool for assessing the QoL of affected persons. The TranQol (transfusion-dependent quality of life) was developed in a methodologically rigorous manner, using the input of specialized healthcare workers and 104 patients with thalassaemia major. This tool has now been further tested in 106 patients followed in thalassaemia treatment centres across the United States and Canada.

The results are convincing: the TranQol showed substantial agreement with the HUI3, the SF-36 and the PedsQL. This is important to note, because these are well-established generic tools that are familiar to many clinicians. The beeswarm graphs in Fig 1 demonstrate the even distribution of the TranQol scores in both adults and children, whereas the HUI3 and the PedsQL show ceiling effects (a clustering of scores at the upper limit of the instrument). In addition, the TranQol shows excellent reliability with almost perfect agreement between the baseline scores and those reported during their subsequent transfusion 2–5 weeks later in patients who reported that their QoL had not changed. Finally, we were able to show that the TranQol is responsive to change, with an improvement in scores at 1 week post-transfusion that reached statistical significance for patients who rated their QoL at least 'a little better'. This was maintained in a sub-analysis of the adult patients, with a change in score that was superior to both summary scores of the SF-36. We were not able to replicate this in the children. This was probably due to the small sample size and wide standard deviation in this sub-group. Ideally, responsiveness should be assessed in patients who rate a more substantial change in their QoL rather than 'a little bit better' as it would be more clinically meaningful, as shown in Fig 2. However, the number of patients in this study with higher change ratings was limited to eight adults and 13 children and as thalassaemia is a relatively stable illness it will take a large study with a longer follow-up to corroborate these results.

The QoL of thalassaemia patients has been assessed previously using the PedsQL and the SF-36 (Messina *et al*, 2008; Clarke *et al*, 2010; Porter *et al*, 2012). A recent study of 22 children followed in the United Kingdom (UK) (Clarke *et al*, 2010) reported PedsQL domain scores at least 10 points lower than those in our North American (NA) study population (physical: UK 67 vs. NA 86, emotional: UK 74 vs. NA 85, social: UK 78 vs. NA 90 and school UK 61 vs. NA 76). They found all domains to be significantly lower than the UK normative data (Clarke *et al*, 2010). Some of the differences in the domain scores reported by Clarke *et al* (2010) compared to our data probably occurred because they reported parent proxy scores as compared to the child self-reported scores in our study. Another possibility is that the NA paediatric population tends to utilize oral chelation more frequently (oral chelation alone: UK 11% vs. NA 65%). A study of 136 Italian (IT) adults with thalassaemia major (Messina *et al*, 2008) found considerably lower SF-36 scale mean values in two of the emotional scales compared to our patients (social: IT 37 vs. NA 79, role emotional IT 36 vs. NA 77). In contrast, five other scales showed similar results (physical: IT 79 vs. NA 82, role physical: IT 75 vs. NA 69, general health: IT 65 vs. NA 52, vitality: IT 53 vs. NA 52, and mental health: IT 67 vs. NA 69) and bodily pain showed values were higher (IT 83 vs. NA 67) (Messina *et al*, 2008). Messina *et al* (2008) was published before the widespread use of oral chelators, which probably explains why social functioning and role limitations due to emotional problems were so much worse in the Italian cohort. It is difficult to explain why the patients studied by Messina *et al* (2008) reported markedly higher body pain scores.

The main limitation of the present study was the sample size. While adequate for analysis of the entire group, it made sub-group analysis difficult. Studies involving rare populations frequently encounter this problem as the patient pool is often limited. In addition, we enrolled patients only from academic NA thalassaemia treatment centres, which may limit the external validity of these results. To address both of these issues, an international study is needed to cross-culturally translate and validate the TranQol in countries that have a higher prevalence of thalassaemia.

QoL tools provide valuable information about patients' perspectives that are not readily accessible to clinicians and researchers in other ways (WHO, 1995). Equally important is that the QoL tools used by researchers and clinicians have been properly developed and evaluated (Patrick *et al*, 2011a, b). This study clearly shows the strengths and advantages of the TranQol in patients with thalassaemia. When combined with well-established generic tools, such as the PedsQL in children and the SF-36 in adults, a clear picture of the patient's perspective becomes apparent. We recommend thalassaemia investigators combine the TranQol with a generic measure to allow the results to be compared to the general population and other disease groups while being sensitive to meaningful changes in QoL.

Conclusion

The TranQol is a disease-specific QoL measure that we have shown to be valid and reliable, with some evidence of superior characteristics compared to generic tools. We recommend that it be incorporated into future clinical trials involving patients with thalassaemia major.

Acknowledgements

Robert J. Klaassen is the principal investigator and supervised all aspects of the study including protocol development, obtaining funding, Research Ethics Board (REB) approval, study implementation, data analysis and wrote the initial draft of the manuscript. Manuela Merelles-Pulcini is the study coordinator and was involved in study implementation, REB approval, data analysis and entry and reviewed the manuscript. Nicholas Barrowman is the study statistician and was involved in setting up the database, data cleaning and analysis and reviewed the manuscript. Elliott P. Vichinsky, Nancy Sweeters, Melanie Kirby-Allen, Ellis J. Neufeld, Janet L. Kwiatkowski, John Wu, Linda Vickars are site investigators for the study, arranging site contract and REB approval, local study implementation and patient recruitment and reviewed the manuscript. Melissa Forgie, Robert Yamashita helped design the research study and reviewed the manuscript. Durhane Wong-Rieger initiated the project and was essential

for arranging the initial funding and reviewed the manuscript. Nancy L. Young and Victor S. Blanchette are the senior investigators and were involved in protocol development and reviewed the manuscript. Nancy Sweeters was supported by a Children's Hospital and Research Center Oakland CTSI grant (UL1RR024131-01). A special thank you to Marketa Waldrum, Colleen Fitzgerald, Latoya Lashley and Sage Green for their work with patient recruitment and data collection, to Jason Vaillancourt for his editorial comments and to Melissa Morrison for her technical support.

Disclosure

The TranQol was developed with an unrestricted grant from Novartis Pharmaceuticals Canada Inc.

References

Clarke, S.A., Skinner, R., Guest, J., Darbyshire, P., Cooper, J., Shah, F., Roberts, I. & Eiser, C. (2010) Health-related quality of life and financial impact of caring for a child with Thalassaemia Major in the UK. *Child: Care Health & Development*, **36**, 118–122.

Connelly, M. & Rapoff, M.A. (2005) Assessing health-related quality of life in children with recurrent headache: reliability and validity of the PedsQLTM 4.0 in a pediatric headache sample. *Journal of Pediatric Psychology*, **31**, 698–702.

Donner, A. & Eliasziw, M. (1987) Sample size requirements for reliability studies. *Statistics in Medicine*, **6**, 441–448.

Eiser, C. & Morse, R. (2001) A review of measures of quality of life for children with chronic illness. *Archives of Disease in Childhood*, **84**, 205–211.

Guyatt, G.H., Bombardier, C. & Tugwell, P.X. (1986) Measuring disease-specific quality of life in clinical trials. *CMAJ*, **134**, 889–895.

Ismail, A., Campbell, M.J., Ibrahim, H.M. & Jones, G.L. (2006) Health Related Quality of Life in Malaysian children with thalassaemia. *Health and quality of life outcomes*, **4**, 39.

Jansen, A.J., Essink-Bot, M.L., Beckers, E.A., Hop, W.C., Schipperus, M.R. & van Rhenen, D.J. (2003) Quality of life measurement in patients with transfusion-dependent myelodysplastic syndromes. *British Journal of Haematology*, **121**, 270–274.

Juniper, E.F., Guyatt, G.H. & Jaeschke, R. (1996) How to develop and validate a new health-related quality of life instrument. In: Quality of Life and Pharmacoeconomics in Clinical Trials (ed. by B. Spilker), pp. 49–56. Raven Press, Philadelphia.

Klaassen, R., Alibhai, S., Kirby Allen, M., Moreau, K., Merelles-Pulcini, M., Forgie, M., Blanchette, V., Buckstein, R., Odame, I., Quirt, I., Yee, K., Wong Rieger, D. & Young, N. (2013) Introducing the TranQol: a new disease-specific quality of life measure for children and adults with thalassemia major. *Journal of Blood Disorders and Transfusion*, **4**, abstract 150.

Landis, J.R. & Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159–174.

Messina, G., Colombo, E., Cassinerio, E., Ferri, F., Curti, R., Altamura, C. & Cappellini, M.D. (2008) Psychosocial aspects and psychiatric disorders in young adult with thalassemia major. *Internal & Emergency Medicine*, **3**, 339–343.

Nixon, S.K., Maunsell, E., Desmeules, M., Schanzer, D., Landgraf, J.M., Feeny, D.H. & Barrera, M.E. (1999) Mutual concurrent validity of the child health questionnaire and the health utilities index: an exploratory analysis using survivors of childhood cancer. *International Journal of Cancer Supplement*, **12**, 95–105.

Olivieri, N.F., Nathan, D.G., MacMillan, J.H., Wayne, A.S., Liu, P.P., McGee, A., Martin, M., Koren, G. & Cohen, A.R. (1994) Survival in medically treated patients with homozygous beta-thalassemia. *New England Journal of Medicine*, **331**, 574–578.

Pakbaz, Z., Treadwell, M., Yamashita, R., Quirolo, K., Foote, D., Quill, L., Singer, T. & Vichinsky, E.P. (2005) Quality of life in patients with thalassemia intermedia compared to thalassemia major. *Annals of the New York Academy of Sciences*, **1054**, 457–461.

Patrick, D.L., Burke, L.B., Gwaltney, C.J., Leidy, N.K., Martin, M.L., Molsen, E. & Ring, L. (2011a) Content validity–establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1–eliciting concepts for a new PRO instrument. *Value Health*, **14**, 967–977.

Patrick, D.L., Burke, L.B., Gwaltney, C.J., Leidy, N.K., Martin, M.L., Molsen, E. & Ring, L. (2011b) Content validity–establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2–assessing respondent understanding. *Value Health*, **14**, 978–988.

Porter, J., Bowden, D.K., Economou, M., Troncy, J., Ganser, A., Habr, D., Martin, N., Gater, A., Rofail, D., Abetz-Webb, L., Lau, H. & Cappellini, M.D. (2012) Health-Related Quality of Life, Treatment Satisfaction, Adherence and Persistence in beta-Thalassemia and Myelodysplastic Syndrome Patients with Iron Overload Receiving Deferasirox: results from the EPIC Clinical Trial. *Anemia*, **2012**, 297641.

Rachmilewitz, E.A. & Giardina, P.J. (2011) How I treat thalassemia. *Blood*, **118**, 3479–3488.

Rizzo, J.A., Pashko, S., Friedkin, R., Mullahy, J. & Sindelar, J.L. (1998) Linking the health utilities index to National Medical Expenditure Survey data. *Pharmacoeconomics*, **13**, 531–541.

Rofail, D., Abetz, L., Viala, M., Gait, C., Baladi, J.F. & Payne, K. (2009) Satisfaction and adherence in patients with iron overload receiving iron chelation therapy as assessed by a newly developed patient instrument. *Value Health*, **12**, 109–117.

Rofail, D., Viala, M., Gater, A., Abetz-Webb, L., Baladi, J.F. & Cappellini, M.D. (2010) An instrument assessing satisfaction with iron chelation therapy: psychometric testing from an open-label clinical trial. *Advances in therapy*, **27**, 533–546.

Sung, L., Greenberg, M.L., Doyle, J.J., Young, N.L., Ingber, S., Rubenstein, J., Wong, J., Samanta, T., McLimont, M. & Feldman, B.M. (2003) Construct validation of the Health Utilities Index and the Child Health Questionnaire in children undergoing cancer chemotherapy. *British Journal of Cancer*, **88**, 1185–1190.

Torrance, G.W., Feeny, D.H., Furlong, W.J., Barr, R.D., Zhang, Y. & Wang, Q. (1996) Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Medical Care*, **34**, 702–722.

Varni, J.W., Burwinkle, T.M., Katz, E.R., Meeske, K. & Dickinson, P. (2002) The PedsQL in pediatric cancer: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer*, **94**, 2090–2106.

Varni, J.W., Burwinkle, T.M., Seid, M. & Skarr, D. (2003) The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambulatory Pediatrics*, **3**, 329–341.

Ware, J.E. (2004) SF-35 Health survey update. In: The Use of Psychological Testing For Treatment Planning and Outcome Assessment (ed. by M.E. Maruish), pp. 693–718. Lawrence Erlbaum Associates, Mahwah, NJ, USA.

Weatherall, D.J. & Clegg, J.B. (2001) Inherited haemoglobin disorders: an increasing global health problem. *Bulletin of the World Health Organization*, **79**, 704–712.

WHO. (1993) Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). *Quality of Life Research*, **2**, 153–159.

WHO. (1995) The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Social Science and Medicine*, **41**, 1403–1409.