

RESEARCH ARTICLE

Validation of a patient-reported outcomes symptom measure for patients with nontransfusion-dependent thalassemia (NTDT-PRO[®])

Ali Taher¹  | Maria Domenica Cappellini²  | Vip Viprakasit³ | Pranee Sutcharitchan⁴ | Dalia Mahmoud⁵ | Abderrahmane Laadem⁵ | Anzalee Khan^{6,7} | Chad Gwaltney⁸ | Gale Harding⁹ | Kenneth Attie¹⁰ | Xiaosha Zhang¹⁰ | Jun Zou⁵ | Joseph Pariseau⁵ | X. Henry Hu^{5†} | Antonis Kattamis¹¹

¹American University of Beirut Medical Center, Beirut, Lebanon

²Ospedale Maggiore Policlinico IRCCS, University of Milan, Milan, Italy

³Siriraj Hospital, Mahidol University, Bangkok, Thailand

⁴Chulalongkorn University, Bangkok, Thailand

⁵Celgene Corporation, Summit, New Jersey

⁶Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York

⁷NeuroCog Trials, Durham, North Carolina

⁸ERT, Philadelphia, Pennsylvania

⁹Evidera, Bethesda, Maryland

¹⁰Accelaron Pharma, Cambridge, Massachusetts

¹¹Aghia Sophia Children's Hospital, Athens, Greece

Correspondence

Ali Taher, American University of Beirut Medical Center, PO Box 11-0236, Riad El Solh, Beirut 1107 2020, Lebanon.
Email: ataher@aub.edu.lb

Funding information

Celgene

Abstract

This study demonstrates the quantitative characteristics of the first patient-reported outcome (PRO) tool developed for patients with nontransfusion-dependent β -thalassemia (NTDT), the NTDT-PRO[®]. A multicenter validation study was performed over 24 weeks, involving 48 patients from Italy, Lebanon, Greece, and Thailand. Most patients were female (68.8%), with a median age of 34.5 years (range, 18–52); 66.7% were diagnosed with β -thalassemia intermedia, and median time since diagnosis was 22 years (range, 0–43). The NTDT-PRO comprises 6 items across 2 domains (Tiredness/Weakness and Shortness of Breath [SoB]), and was valid and reliable, with good consistency. At baseline, most patients reported symptoms as present via the NTDT-PRO, and were highly compliant, $\geq 90\%$ completing the NTDT-PRO tool. In a pairwise correlation analysis, all items were positively correlated. Correlations between NTDT-PRO and existing tools—36-Item Short Form Health Survey version 2 (SF-36v2) and Functional Assessment of Cancer Therapy-Anemia (FACT-An)—were assessed at weeks 1, 3, and 12; robust correlations were seen between SoB and SF-36v2-Vitality ($r_s = -0.53$), and between SoB and Fact-An-Fatigue Experience ($r_s = -0.66$) at week 1. Internal consistency was high for both Tiredness/Weakness (Cronbach alpha, 0.91) and SoB (Spearman-Brown coefficient, 0.78); intraclass correlation coefficients were high (Tiredness/Weakness, 0.88 and 0.97; SoB, 0.92 and 0.98), demonstrating stability. Further studies are required to fully support the validity of this tool, this study demonstrated the usefulness of the NTDT-PRO in the clinical setting and for longitudinal clinical research, particularly in trials where patient health-related quality of life is expected to change.

1 | INTRODUCTION

The hereditary blood disorder β -thalassemia affects approximately 1.5% of the global population.¹ In 2008, the World Health Organization estimated that more than 40 000 children are born annually with β -thalassemia, of which over 20 000 are in Southeast Asia.² More recent estimates from Europe suggest there are approximately 7000

patients in Italy (0.012% of a total population of approximately 60 million) and 3241 patients in Greece (0.029% of a total population of 11 million) with thalassemias.³ Modern patterns of increased migration have resulted in thalassemias becoming widespread in other regions,⁴ as illustrated by a report from the Centers for Disease Control that reported that 27% of people with nontransfusion-dependent β -thalassemia (NTDT) in North America were born overseas.⁵

Increasing global access to health care and the resultant increase in life expectancy of patients with thalassemias has intensified efforts

[†]Former employee of Celgene Corporation.

to understand how this condition affects health-related quality of life (HRQoL). Although disease burden varies according to whether a patient has NTDT or transfusion-dependent thalassemia (TDT), the impact on HRQoL and general well-being is significant.^{6,7}

At present, the main HRQoL patient-reported outcome (PRO) tools used in thalassemias are the 36-Item Short Form Health Survey version 2 (SF-36v2) and the Functional Assessment of Cancer Therapy-Anemia (FACT-An) tools. These tools, particularly the SF-36v2, are not disease specific and may include symptoms beyond those that thalassemia patients would typically experience.

The majority of studies that investigate HRQoL using the SF-36v2 in thalassemias focus on TDT, for example, in Italy⁸ and Greece.⁹ However, the SF-36v2 was used in a longitudinal study of 264 patients with TDT and NTDT from the United States, Canada, and the United Kingdom, and found that HRQoL was significantly lower in this population compared with US norms, and interestingly, saw a higher HRQoL associated with being transfused,¹⁰ indicating important differences are present between TDT and NTDT patients. However, two studies that used the Pediatric Quality of Life Inventory (PedsQL) 4.0 Generic Core scale to investigate HRQoL in Thai children with thalassemia found that this scale gave conflicting results, with the first finding a negative association between HRQoL and transfusion therapy,¹¹ and the second reporting no difference in HRQoL score associated with type of thalassemia or clinical severity,¹² emphasizing the need for specific tools that are developed and validated for this patient population.

The FACT-An was developed to assess HRQoL in relation to anemia experienced by patients with cancer who are undergoing chemotherapy,¹³ and so, although the patient population it was developed for may have a similar experience in some respects, there are aspects of the NTDT patient experience that may not be adequately captured.

In addition, while there are a number of PRO tools that have been developed specifically for TDT patients, such as the Specific Thalassemia Quality of Life Instrument (STQOLI)¹⁴ and the TranQoL,¹⁵ these focus on the impact of hospitalization associated with repeated transfusion and iron chelation therapy. As patients with NTDT receive far fewer transfusions and may experience a different symptom profile, PRO tools to measure HRQoL in this patient group must be developed with this in mind.

Regulatory agencies, such as the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA), are interested in the development and validation of disease-specific PROs as part of clinical outcome assessment tools for clinical research. To this end, the NTDT-PRO[®] tool was developed following a rigorous, qualitative, multicenter study, involving NTDT patients from four countries (Taher et al., manuscript submitted for publication). The NTDT-PRO assesses the presence or severity of the following symptoms using a numerical rating scale ranging from 0 ("absent/minimal") to 10 ("extreme"): Tiredness with or without Physical Activity, Weakness with or without Physical Activity, and Shortness of Breath (SoB) with or without Physical Activity. These items were selected as a basis for the NTDT-PRO following expert panel discussion and patient interviews.

The objective of the present study was to validate the NTDT-PRO against existing HRQoL tools (SF-36v2 and FACT-An) in patients

with NTDT, and demonstrate the sensitivity and score reliability over time of the NTDT-PRO when used in patients with NTDT.

2 | METHODS

2.1 | Evaluation of the NTDT-PRO tool

2.1.1 | Study design

The NTDT-PRO was evaluated in a prospective observational study with no protocol-driven treatment intervention (NCT02626689). Although the observational study included both TDT ($n = 52$) and NTDT ($n = 50$) patients, only NTDT patients were included in this evaluation. Patients were enrolled at five centers across four countries (Italy, Lebanon, Greece, and two centers in Thailand); as the centers in Lebanon and Greece were also involved in the development of the NTDT-PRO tool (Taher et al., manuscript submitted for publication), there was some overlap between patient populations. Eligible patients were ≥ 18 years of age, had a documented diagnosis of β -thalassemia intermedia or hemoglobin (Hb) E/ β -thalassemia, Hb ≤ 10 g/dL, and had received ≤ 5 red blood cell (RBC) units during the 24-week period prior to study start. All patients were asked to complete the NTDT-PRO daily via a handheld device, as well as the SF-36v2¹⁶ and FACT-An¹³ HRQoL questionnaires at baseline and every 3 weeks thereafter for a total of ≥ 24 weeks.

Changes in patient health status since the start of the study were also assessed using the Patient Global Impression of Change (PGIC) scale once every 3 weeks.¹⁷

Three weeks was selected as a testing interval for this study to match the dosing regimen for luspatercept, a novel therapy under development for TDT and NTDT.

2.1.2 | Measures and procedures

The NTDT-PRO was designed to capture thalassemia-related symptoms over the previous 24 hr using six symptom-specific items (Appendix in Supporting Information). An additional item was included in the NTDT-PRO for the purposes of validation, to assess the overall severity of the patients' thalassemia symptoms in the previous 24 hr on a numerical scale ranging from 0 ("no symptoms") to 10 ("very severe symptoms"). Item scores were summarized as the mean of nonmissing values over a 7-day period; if entries were missing for ≥ 4 days in any 7-day period, the weekly average was considered missing. For the two domains, Tiredness/Weakness and SoB, the mean and sum of items in the domain were reported.

All HRQoL questionnaires (NTDT-PRO, SF-36v2, and FACT-An) were administered in the local language (English, Thai, Greek, Arabic, or Italian, as appropriate). For the SF-36v2 and FACT-An, the most recently validated local language versions were used, available from Optum and FACIT.org, respectively. For the NTDT-PRO, the original English version was reviewed for translation accuracy by local thalassemia experts who were involved in the study (MDC for Italian, AT for Arabic, AK for Greek, and VV for Thai; translations available on request).

Data from patients who reported "no change" (ie, a score of 4 on a 0-10 scale) on the PGIC during weeks 3 and 12 of the study were used to assess test-retest reliability.

Patients with both a valid week of daily data collection at baseline (defined as ≥ 4 complete diary entries in the 7-day period) and data for ≥ 1 questionnaire (SF-36v2, FACT-An) at baseline were included in the analysis.

2.2 | Statistical analysis

Quantitative data, including weekly averaged NTDT-PRO items, were summarized using descriptive statistics.

2.2.1 | Internal consistency and validity

Internal consistency was assessed by calculating the Spearman rank correlation coefficient to evaluate the strength and direction of a relationship between all NTDT-PRO items.

To examine internal validity, reliability was tested using the Cronbach alpha coefficient for the Tiredness/Weakness domain score. As SoB contained only two variables, the Spearman-Brown coefficient was calculated to determine the consistency of this measure.

2.2.2 | Validity

To investigate validity of the NTDT-PRO, correlations between the 2 NTDT-PRO domain scores (Tiredness/Weakness, and SoB) and the domain subscale scores of existing HRQoL tools—the SF-36v2 and the Fatigue subscale (FACIT-F) of FACT-An—were evaluated by the Spearman-Brown correlation coefficient. Known group validity compared eDiary scores between groups defined by the number of comorbidities.

2.2.3 | Stability over time and sensitivity to change

To evaluate the stability of weekly averaged NTDT-PRO items and domain scores, test-retest reliability was assessed by calculating the intraclass correlation coefficient (ICC) over 2 intervals (weeks 1-3 and weeks 9-12) in patients who reported no change in their overall β -thalassemia symptoms at week 3 and 12, as measured by PGIC score.

Sensitivity to change was assessed by examining changes in domain scores between weeks 1 and 24 in three groups identified as improved, no change, or worse based on PGIC score at week 24. However, as this study was observational and did not involve an intervention, significant changes were not anticipated.

Correlations between 0.20 and 0.39 were defined as “weak,” correlations between 0.40 and 0.59 as “moderate,” correlations between 0.60 and 0.79 as “strong,” and those ≥ 0.80 were defined as “very strong.” Significance was defined as $P < .05$.

3 | RESULTS

3.1 | Measurement evaluation

3.1.1 | Patients

A total of 50 patients from five centers in four countries were enrolled in this prospective observational study; 48 patients completed ≥ 1 post-baseline NTDT-PRO entry and were included in the NTDT-PRO validation analysis. A total of 15 (31.3%), 13 (27.1%),

TABLE 1 Patient characteristics

Patient characteristic	Analysis cohort (N = 48)
Median age (range), years	34.5 (18-52)
Female, n (%)	33 (68.8)
Thalassemia diagnosis, n (%)	
β -Thalassemia intermedia	32 (66.7)
HbE/ β -thalassemia	8 (16.7)
β -Thalassemia with α -triplication	8 (16.7)
Median time since diagnosis (range), years	22 (0-43)
Hospitalization in last 3 months, n (%)	NA
Median Hb (range), g/dL	8.0 (6.0-10.0)
Highest education level, n (%)	
Primary/elementary	3 (6.3)
Secondary/high school	4 (8.3)
College/university	23 (47.9)
Unknown/other	18 (37.5)

Abbreviations: Hb, hemoglobin; HbE, hemoglobin E; NA, not available.

11 (22.9%), and 9 (18.8%) patients were enrolled in Italy, Lebanon, Greece, and Thailand, respectively. Baseline characteristics are shown in Table 1.

The majority of patients were female (68.8%), with a median age of 34.5 years (range, 18-52).

During the course of the study, 7 patients (15.0%) received RBC transfusions.

3.1.2 | NTDT-PRO distribution characteristics

NTDT-PRO item and domain scores are detailed in Table S1 in the Supporting Information. Of the 48 patients included in the study, a median of 37 patients (range, 30-48) completed the NTDT-PRO and contributed data each week ($\geq 4/7$ days). The individual items with the highest scores across the 24 weeks of testing were Tiredness with Physical Activity (TiredPA) (range, 0-10.00) and Weakness with Physical Activity (WeakPA) (range, 0-10.00); the lowest scores were seen for SoB without Physical Activity (SoBNA) (range, 0-7.83).

3.1.3 | Internal consistency of NTDT-PRO

Spearman rank correlations between NTDT-PRO items are shown in Table S2 in the Supporting Information. All pairwise correlations were positive, and very strong correlations were found between the items WeakPA and TiredPA ($r_s = 0.96$), and between Weakness without Physical Activity (WeakNA) and Tiredness without Physical Activity (TiredNA) ($r_s = 0.89$). Strong correlations were also found between different constructs measuring the same level of activity, for example, for physical activity, SoB, and both Tiredness ($r_s = 0.68$) and Weakness ($r_s = 0.72$); similar correlations were seen for no physical activity. Strong correlations were also observed between items measuring the same construct but a different activity level (eg, Tiredness, $r_s = 0.70$, Weakness, $r_s = 0.69$, and SoB, $r_s = 0.69$).

The raw variable value of the Cronbach alpha (0.91) indicates high internal consistency and reliability across the Tiredness/Weakness domain (Table 2). The correlation within the SoB domain was also strong (Spearman-Brown coefficient, 0.78).

TABLE 2 Internal consistency of the NTDT-PRO[®] Tiredness/Weakness domain scores at baseline

	Raw variables		Standardized variables	
	Correlation with total	Cronbach alpha	Correlation with total	Cronbach alpha
Tiredness/Weakness domain ^a	-	0.91	-	0.92
TiredNA	0.79	0.89	0.82	0.89
TiredPA	0.85	0.87	0.82	0.89
WeakNA	0.76	0.90	0.79	0.90
WeakPA	0.85	0.87	0.81	0.89

Correlations >0.60 were considered "strong."

Abbreviations: NTDT-PRO[®], nontransfusion-dependent thalassemia patient-reported outcomes measure; TiredNA, Tiredness with no Physical Activity; TiredPA, Tiredness with Physical Activity; WeakNA, Weakness with no Physical Activity; WeakPA, Weakness with Physical Activity.

^a Spearman-Brown coefficient for the Shortness of Breath domain was 0.78 ($n = 48$).

3.1.4 | General validity of the NTDT-PRO as compared with SF-36v2 and FACT-an

Correlation between NTDT-PRO domain scores and the SF-36v2 and FACT-An was assessed at weeks 1, 3, and 12. As the NTDT-PRO measures symptom severity (ie, high scores indicate worse symptoms) and the SF-36v2 and FACT-An measure function (ie, high scores indicate better function), all correlations seen were negative (Table 3). An exploratory factor analysis was used to

explore item groupings, which led to the combining of data from the Tiredness and Weakness domains. The structural coefficients for the NTDT-PRO suggested that TiredNA and WeakNA are correlated 0.892 and 0.898, respectively, with Factor 1, with loadings >0.5. TiredPA and WeakPA are correlated 0.908 and 0.918, respectively, with Factor 2.

The strongest correlations were seen during week 1 between the Tiredness/Weakness domain of the NTDT-PRO and the Vitality domain of the SF-36v2 ($r_s = -0.53$) and the Fatigue Experience subscale on the FACT-An ($r_s = -0.66$). This pattern was repeated across all time points tested, with the exception of week 12 of the FACT-An score correlations, where the highest correlations were between the Fatigue scale and Tiredness/Weakness ($r_s = -0.80$), and between the FACT-General and SoB ($r_s = -0.61$).

At all time points investigated, the weakest correlations observed were between the Social/Family Well-being domain of Fact-An and the SoB domain of NTDT-PRO; the Social/Family Well-being, Functional Well-being, and Anemia Symptoms domains correlated poorly with the Tiredness/Weakness domain of the NTDT-PRO. Low correlations were also observed between the SF-36v2 item Bodily Pain and SoB, and also between General Mental Health and both SoB and Tiredness/Weakness.

Known groups validity analysis with comorbidities showed that as the number of comorbidities increased from 1 to 4, the ratings of

TABLE 3 Convergent correlations between NTDT-PRO[®] domain scores and the SF-36v2 and FACT-An scales

Spearman correlation coefficient	NTDT-PRO domain					
	Tiredness/Weakness			Shortness of Breath		
	Week 1 ($n = 48$)	Week 3 ($n = 46$)	Week 12 ($n = 35$)	Week 1 ($n = 48$)	Week 3 ($n = 46$)	Week 12 ($n = 35$)
SF-36v2						
Physical functioning	-0.41*	-0.41*	-0.43*	-0.40*	-0.42*	-0.42*
Role-physical functioning	-0.29	-0.30	-0.21	-0.22	-0.21	-0.21
Bodily pain	-0.20	-0.20	-0.10	-0.10	-0.10	-0.10
General health perceptions	-0.39*	-0.40*	-0.37*	-0.37*	-0.37*	-0.37*
Vitality	-0.53*	-0.53*	-0.45*	-0.45*	-0.46*	-0.45*
Social functioning	-0.32*	-0.32	-0.14	-0.14	-0.14	-0.14
Role-emotional functioning	-0.17	-0.18	-0.19	-0.19	-0.19	-0.19
General mental health	-0.10	-0.12	-0.12	-0.10	-0.12	-0.10
Physical component score	-0.43*	-0.42	-0.28	-0.26	-0.26	-0.27
Mental component score	-0.21	-0.22	-0.19	-0.19	-0.19	-0.19
FACT-An						
Physical well-being	-0.33	-0.62*	-0.63	-0.24	-0.42*	-0.60*
Social/family well-being	-0.19	-0.33*	-0.39*	-0.08	-0.18	-0.32*
Emotional well-being	-0.20	-0.35*	-0.54*	-0.14	-0.34*	-0.48*
Functional well-being	-0.16	-0.49*	-0.44*	-0.21	-0.37	-0.42*
Anemia symptoms ($k = 7$)	-0.36*	-0.43*	-0.34*	-0.24	-0.35*	-0.35*
Fatigue scale	-0.55	-0.70	-0.80*	-0.38*	-0.51*	-0.58*
Fatigue experience ($k = 5$)	-0.66*	-0.75*	-0.78*	-0.54*	-0.53*	-0.54*
Fatigue impact ($k = 8$)	-0.38*	-0.61*	-0.78*	-0.21	-0.47*	-0.60*
FACT-general	-0.29	-0.59*	-0.64*	-0.20	-0.45*	-0.61*

Correlations >0.60 were considered "strong." "k" indicates degrees of freedom. Abbreviations: FACT, Functional Assessment of Cancer Therapy; FACT-An, FACT-Anemia; NTDT-PRO[®], nontransfusion-dependent thalassemia patient-reported outcomes measure; SF-36v2, 36-Item Short Form Health Survey version 2.

* $P < .05$.

TABLE 4 Test-retest reliability of the NTDT-PRO[®] domain scores

Domain score	Weeks 1-3 (n = 32)		Weeks 9-12 (n = 21)	
	Tiredness/Weakness	Shortness of Breath	Tiredness/Weakness	Shortness of Breath
Mean score in first week of time period (SD) ^a	2.84 (1.52)	1.87 (1.50)	2.51 (1.81)	1.77 (1.68)
Mean score in last week of time period (SD) ^b	2.76 (1.84)	1.93 (1.73)	2.50 (1.92)	1.82 (1.73)
Difference in means (SD)	0.08 (1.69)	-0.06 (1.62)	0.01 (1.86)	-0.05 (1.71)
P value ^c	.84	.88	.96	.93
ICC	0.88	0.92	0.97	0.98

Correlations >0.60 were considered "strong." Significance was defined as $P < .005$.

Abbreviations: ICC, intraclass correlation coefficient; NTDT-PRO[®], nontransfusion-dependent thalassemia patient-reported outcomes measure; SD, standard deviation.

^a Week 1 or 9.

^b Week 3 or 12.

^c P value by t-test.

severity for the Tiredness/Weakness, SoB items and both domain scores increased. When the number of comorbidities was greater than 4, the severity decreased, and in some cases, below the level of severity for 0 comorbidities (eg, TiredPA, TiredNA, WeakNA, SoB domain) (data not shown).

3.1.5 | Stability over time and sensitivity to change

The stability of the NTDT-PRO was assessed by looking at test-retest reliability over weeks 1-3 ($n = 32$) and weeks 9-12 ($n = 21$) for the 2 NTDT-PRO domain scores in patients who reported no change by PGIC (Table 4). ICCs were high for Tiredness/Weakness (0.88 and 0.97) and SoB (0.92 and 0.98) at all time points studied, and the average difference in scores between weeks 1 and 3 (Tiredness/Weakness: 0.08; SoB: -0.06), and weeks 9 and 12 (Tiredness/Weakness: 0.01; SoB: -0.05) was small.

To assess the ability of the NTDT-PRO to detect change in health status, PGIC responses were examined. Across the 24 weeks of the study, the majority of patients (17 patients; 63%) reported no change in PGIC score; 5 patients (19%) reported a worsening, and 5 patients (19%) an improvement of the PGIC score. However, as expected in this observational, noninterventional study, no significant change in NTDT-PRO domain score from weeks 1 to 24 was observed, regardless of change in PGIC score (data not shown).

4 | DISCUSSION

Patients with NTDT experience symptoms that negatively impact their HRQoL, and there is a paucity of specific tools to quantify this in these patients. This study reports the multicenter validation of a novel disease-specific PRO tool for patients with NTDT, the NTDT-PRO. The NTDT-PRO was developed using concept-elicitation interviews, interview discussions, cognitive interviews, and clinical input. Concept elicitation identified 11 symptoms from which 9 were selected for inclusion in the questionnaire. An 11-point numerical rating scale score was given to each item. Cognitive interviews were conducted to confirm that the selected symptoms were relevant and understood by patients. Based on patient interviews, 6 items were selected for inclusion in the final tool; TiredPA or TiredNA, WeakPA or WeakNA, and SoB with Physical Activity (SoBPA) or SoBNA (Taher et al., manuscript

submitted for publication). Validation involved 48 patients from five centers across four countries. It comprises 6 items across 2 domains (Tiredness/Weakness and SoB), and an overall severity score (from 0 to 5). Our evaluation found the NTDT-PRO to be valid, reliable, and consistent across different countries. At baseline, all items in the NTDT-PRO were identified as present by most patients, and compliance with the NTDT-PRO was generally high; the proportion of patients able to complete the daily diary ≥ 4 times per week was in the range of 60%-100%.

The items with the highest score in the NTDT-PRO were TiredPA and WeakPA, highlighting that physical activity is often most severely impacted for patients with NTDT.

SoBNA was the item with the lowest scores in the NTDT-PRO. This may reflect differences in the way patients experience or report shortness of breath, or the relative level of discomfort associated with this symptom. Alternatively, it is possible that the levels of SoBNA are generally low in this population. Providing patients with a clear understanding of the definition of shortness of breath, for example, by using a short video clip, and how to differentiate this symptom from tiredness and weakness, as well as clarification on what constitutes physical activity, may therefore be beneficial.

All pairwise inter-item correlations of the NTDT-PRO were positive. Very strong correlations were found between physical-activity-congruent Weakness and Tiredness items in particular. This indicates a level of redundancy across Weakness and Tiredness and, in combination with the fact that these outcomes also registered the highest scores overall, highlights the importance of including both measures in the NTDT-PRO.

Measures of Cronbach reliability for the NTDT-PRO were encouraging, indicating a high degree of precision with low measurement error. High ICC values for Tiredness/Weakness in weeks 1-3, and across all domains in weeks 9-12, suggested good reproducibility of this measurement.

Strong correlations were observed between the NTDT-PRO and relevant domains of the SF-36v2 and FACT-An tools, supporting the validity of the NTDT-PRO. The strongest correlations were observed between the Tiredness/Weakness domain of the NTDT-PRO, and the Vitality domain of the SF-36v2 and the Fatigue Experience subscale of the Fact-An, confirming both the ubiquitous nature of tiredness and weakness in patients with β -thalassemia and the likelihood that

the NTDT-PRO measures that same concept in the corresponding domains/scales of the SF-36v2 and FACT-An tools.

The strength of these correlations suggests that tiredness and weakness, and shortness of breath, as assessed by the NTDT-PRO, represent proximal aspects of the NTDT disease experience rather than HRQoL per se.

Patient PGIC scores were assessed at baseline and throughout the study to provide an anchor for the data set and allow identification of clinically important responders. While the majority (63%) of patients reported no change in PGIC, several patients reported either improvement (19%) or worsening (19%) of disease by PGIC. As this study was designed as an observational, noninterventional study, significant changes were not expected in PGIC or NTDT-PRO scores.

The NTDT-PRO tool was developed in a population of patients with a diagnosis of nontransfusion-dependent β -thalassemia confirmed by chart review; it is likely that it would yield similar results in patients with nontransfusion-dependent α -thalassemia; however, further research would be required before this tool can be recommended for use in other NTDT populations. This study did not assess the responsiveness of the NTDT-PRO to treatment, and future studies to validate this tool in an interventional setting will provide further insight into the sensitivity to change of the NTDT-PRO. The ecological validity, that is, the generalizability of the tool to a real-world setting, such as community-based centers versus academic centers, has also not been assessed. Further validation of the NTDT-PRO in these real-world settings will be important to increase the utility of the tool.

This study supports the validity of the NTDT-PRO tool as appropriate and stable for assessment of key symptoms in patients with NTDT. The NTDT-PRO tool takes under 5 min to complete, less than the time required both for the SF-36v2 and for the FACT-An (around 10 min for each). Moreover, this score provides a novel, disease-specific tool focusing on symptoms that were identified as significant to patients with NTDT, both according to physician insight and from self-report of patient experience. The NTDT-PRO could, therefore, be used in routine clinical evaluation and to support clinical trials investigating new pharmacologic agents that might ameliorate clinical anemia and improve quality of life.

For the first time in the area of clinical hematology, both the development (Taher et al., manuscript submitted for publication) and validation of the NTDT-PRO were conducted across multiple countries and in multiple languages, with a high level of reliability and reproducibility. As thalassemia, particularly NTDT, is prevalent across the world, this NTDT-PRO tool could be of clinical use worldwide. Further interventional studies to confirm the psychometric properties of the NTDT-PRO, validate the sensitivity to change, and determine what constitutes the minimal clinically important difference and a clinically important responder in this population are ongoing.

ACKNOWLEDGMENTS

The authors received editorial and writing support provided by Rosie Morland, PhD, from Excerpta Medica, funded by Celgene. The authors had full access to the data and are fully responsible for content and editorial decisions for this manuscript. The authors wish to

acknowledge the generous participation of all patients who were involved in this study.

COPYRIGHT STATEMENT

All rights reserved. When using this set of outcomes, or quoting therefrom, in any way, we require reference to this article. Funding was provided by Celgene Corporation.

CONFLICT OF INTEREST

Taher, Ali: Novartis: Honoraria and research funding; Celgene Corporation: Research funding; Roche: Research funding.

Cappellini Domenica, Maria: Sanofi Genzyme, Celgene Corporation and Roche: Advisory board member.

Viprakit, Vip: Celgene Corporation, Novartis, SEBIA, Bio-Rad and Roche: Research support.

Sutchartitchan, Pranee: Celgene Corporation, Novartis: Research support.

Mahmoud, Dalia: Celgene Corporation: Employment.

Laadem, Abderrahmane: Celgene Corporation: Employment.

Khan, Anzalee: Manhattan Psychiatric Center, Nathan S. Kline Institute for Psychiatric Research and NeuroCog Trials: Employment.

Gwaltney, Chad: Gwaltney Consulting: Employment; ERT Inc. and Celgene Corporation: Consulting fees.

Harding, Gale: Evidera: Employment; Celgene Corporation: Consulting fees.

Attie, Kenneth: Acceleron Pharma: Employment.

Zhang, Xiaosha: Acceleron Pharma: Employment.

Zou, Jun: Celgene Corporation: Employment.

Pariseau, Joseph: Celgene Corporation: Employment.

Hu, X. Henry: Celgene Corporation: Ex-employee; equity ownership.

Kattamis, Antonis: Celgene Corporation, Novartis, ApoPharma: Research support, advisory and educational board member.

ORCID

Ali Taher  <https://orcid.org/0000-0001-8515-2238>

Maria Domenica Cappellini  <https://orcid.org/0000-0001-8676-6864>

REFERENCES

- Colah R, Gorakshakar A, Nadkarni A. Global burden, distribution and prevention of β -thalassemias and hemoglobin E disorders. *Expert Rev Hematol*. 2010;3:103-117.
- Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ*. 2008;86:480-487.
- Angastiniotis M, Vives Corrons J-L, Soteriades ES, Eleftheriou A. The impact of migrations on the health services for rare diseases in Europe: the example of haemoglobin disorders. *Sci World J*. 2013;2013:727905.
- Vichinsky EP. Changing patterns of thalassemia worldwide. *Ann N Y Acad Sci*. 2005;1054:18-24.
- Vichinsky E, Cohen A, Thompson AA, et al. Epidemiologic and clinical characteristics of nontransfusion-dependent thalassemia in the United States. *Pediatr Blood Cancer*. 2018;65:e27067.

6. Amid A, Leroux R, Merelles-Pulcini M, et al. Factors impacting quality of life in thalassemia patients; results from the Intercontinenthal collaborative study. *Blood*. 2016;128:3633.
7. Taher A, Vichinsky E, Musallam K, Cappellini MD, Viprakasit V. Blood Transfusion. In: Weatherall D, ed. *Guidelines for the Management of Non Transfusion Dependent Thalassaemia (NTDT)*. Nicosia, Cyprus: Thalassaemia International Federation; 2013.
8. Gollo G, Savioli G, Balocco M, et al. Changes in the quality of life of people with thalassemia major between 2001 and 2009. *Patient Prefer Adherence*. 2013;7:231-236.
9. Goulas V, Kourakli-Symeonidis A, Camoutsis C. Comparative effects of three iron chelation therapies on the quality of life of Greek patients with homozygous transfusion-dependent beta-thalassemia. *ISRN Hematol*. 2012;2012:139862.
10. Sobota A, Yamashita R, Xu Y, et al. Quality of life in thalassemia: a comparison of SF-36 results from the thalassemia longitudinal cohort to reported literature and the US norms. *Am J Hematol*. 2011;86:92-95.
11. Surapolchai P, Satayasai W, Sinlapamongkolkul P, Udomsubpayakul U. Biopsychosocial predictors of health-related quality of life in children with thalassemia in Thammasat University Hospital. *J Med Assoc Thai*. 2010;93(Suppl 7):S65-S75.
12. Torcharus K, Pankaew T. Health-related quality of life in Thai thalassaemic children treated with iron chelation. *Southeast Asian J Trop Med Public Health*. 2011;42:951-959.
13. Cella D. The functional assessment of cancer therapy-anemia (FACT-an) scale: a new tool for the assessment of outcomes in cancer anemia and fatigue. *Semin Hematol*. 1997;34(3 Suppl 2):13-19.
14. Lyrakos GN, Vini D, Aslani H, Drosou-Servou M. Psychometric properties of the specific thalassemia quality of life instrument for adults. *Patient Prefer Adherence*. 2012;6:477-497.
15. Klaassen RJ, Barrowman N, Merelles-Pulcini M, et al. Validation and reliability of a disease-specific quality of life measure (the TranQol) in adults and children with thalassaemia major. *Br J Haematol*. 2014;164:431-437.
16. Ware JE, Kosinski MR, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. *User's Guide for the SF-36v2 Health Survey*. Lincoln, RI: QualityMetric Incorporated; 2007.
17. Guy W, ed. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: US Department of Health, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration; 1976.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Taher A, Cappellini MD, Viprakasit V, et al. Validation of a patient-reported outcomes symptom measure for patients with nontransfusion-dependent thalassemia (NTDT-PRO[®]). *Am J Hematol*. 2018;1-7. <https://doi.org/10.1002/ajh.25344>