Systematic Reviews

RESEARCH Open Access

Measurement properties of the Traumatic Brain Injury Quality of Life (TBI-QoL) and Spinal Cord Injury Quality of Life (SCI-QoL) measurement systems: a systematic review

Rebecca Ataman[1,2,3], Rehab Alhasani[4], Line Auneau-Enjalbert[1,2], Adria Quigley[5,6], Henry Ukachukwu Michael[7,8] and Sara Ahmed[1,2,8*]

Abstract

Purpose Traumatic brain injury and spinal cord injury impact all areas of individuals' quality of life. A synthesis of available evidence for the Traumatic Brain Injury Quality of Life (TBI-QoL) and Spinal Cord Injury Quality of Life (SCI-QoL) measurement systems could inform evidence-based clinical practice and research. Thus, we aimed to systematically review the literature of existing evidence on the measurement properties of SCI-QoL and TBI-QoL among rehabilitation populations.

Methods We used the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) framework for evaluating measures to guide this systematic review. We searched nine electronic databases and registries, and hand-searched reference lists of included articles. Two independent reviewers screened selected articles and extracted the data. We used COSMIN's thresholds to synthesize measurement properties evidence (insufficient, sufficient), and the modified GRADE approach to synthesize evidence quality (very-low, low, moderate, high).

Results We included 16 studies for SCI-QoL and 14 studies for TBI-QoL. Both measurement systems have sufficient content validity, structural validity, internal consistency and construct validity across nearly all domains (GRADE: high). Most SCI-QoL domains and some TBI-QoL domains have sufficient evidence of cross-cultural validity and test–retest reliability (GRADE: moderate-high). Besides the cognition domains of TBI-QoL, which have indeterminate evidence for measurement error and sufficient evidence for responsiveness (GRADE: high), there is no additional evidence available for these measurement properties.

Conclusion Rehabilitation researchers and clinicians can use SCI-QoL and TBI-QoL to describe and evaluate patients. Further evidence of measurement error, responsiveness, and predictive validity would advance the use and interpretation of SCI-QoL and TBI-QoL in rehabilitation.

Keywords Systematic review, Patient-reported outcome measures, Traumatic brain injury, Spinal cord injury, Psychometrics, COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN)

*Correspondence:
Sara Ahmed
sara.ahmed@mcgill.ca
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Patient-reported outcome measures play a key role in the delivery of patient-centered care in rehabilitation [1, 2] and have become common, or even a requirement, in clinical trials research [3]. The Patient Reported Outcomes Measurement Information System (PROMIS) [4] and the Neurological Quality of Life (Neuro-QoL) [5] measurement system are two of the most commonly used patient-reported outcome measures. They provide health-related quality-of-life data collected directly from both the general and neurological populations. However, neither PROMIS nor Neuro-QoL included traumatic brain injury patients in their development [6–8]. While PROMIS did include spinal cord injury patients, they were grouped with other patient populations [9]. There is evidence that spinal cord injury and traumatic brain injury are specialized patient populations with their own unique concerns relative to other neurological populations in rehabilitation [10–13]. The Traumatic Brain Injury Quality of Life (TBI-QoL) and Spinal Cord Injury Quality of Life (SCI-QoL) measurement systems were developed for traumatic brain injury and spinal cord injury populations using items verbatim from PROMIS and Neuro-QoL, as well as including new items or developing new scales where relevant. The TBI-QoL and SCI-QoL item banks were calibrated with individuals with traumatic brain injury and spinal cord injury to assess their physical health, emotional health, and social participation while maintaining the metrics of respective PROMIS and Neuro-QoL items banks to aid in cross-population comparison [6, 7]. Thus, TBI-QoL and SCI-QoL are optimized for use in traumatic brain injury and spinal cord injury populations, while providing easy comparison to all populations for whom PROMIS and Neuro-QoL can be administered.

TBI-QoL was first published in 2016 [8] and SCI-QoL in 2015 [6]. Since publication TBI-QoL and SCI-QoL have been used in clinical trials [14, 15] and clinical practice [16, 17]. Despite the increasing use of TBI-QoL and SCI-QoL, there has been no systematic synthesis of the measurement properties of these measurement systems to inform their use in evidence-based rehabilitation. Beyond commentaries published by the developers of these measures [6, 7], the only synthesis available for these measures is of a single domain of SCI-QoL (bowel management difficulties) [18]. There is a need for a synthesis of the measurement properties of TBI-QoL and SCI-QoL to inform their use and interpretation by rehabilitation clinicians and researchers.

Thus, the objective of this review was to systematically review the literature of existing evidence on the measurement properties of TBI-QoL and SCI-QoL measures among rehabilitation populations.

Methods

In this systematic review, we followed the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) 2018 guidelines [19]. We originally intended to synthesize and report information on PROMIS, Neuro-QoL, SCI-QoL, and TBI-QoL in one review. However, the volume of included articles across these four measurement systems led us to report the results across multiple manuscripts. Our methods are reported accordingly, with all four measurement systems present in our literature search strategy, followed by a focus solely on SCI-QoL and TBI-QoL beginning with data extraction. Our results for Neuro-QoL and PROMIS are available elsewhere [20]. We conducted this review of SCI-QoL and TBI-QoL in accordance with the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) 2018 guidelines [19] and report the methods in accordance with PRISMA guidelines for systematic reviews [21].

Literature search and eligibility

We searched electronic databases (MEDLINE, EMBASE, PsycINFO, and HaPI (Ovid), CINAHL (EBSCO), Cochrane Library, and Web of Science) and clinical trials registries (ISRCTN Registry and ClinicalTrials.gov) from inception to March 23rd, 2024. The search strategy (Psychometric properties AND (Neuro-QoL OR PROMIS OR TBI-QoL OR SCI-QoL) AND Rehabilitation Conditions; MEDLINE search strategy in Supplementary material 1) was developed using a measurement properties search filter validated by COSMIN [22].

EndNote X9 [23] was used to deduplicate articles, after which two independent reviewers screened titles, abstracts, and full texts. As part of the broad scope of our initial review, we included peer-reviewed articles in English or French that provided original data on TBI-QoL, SCI-QoL, Neuro-QoL, or PROMIS measurement properties among rehabilitation populations (Table 1). For this specific manuscript, we included TBI-QoL and SCI-QoL articles only. We excluded articles that (1) did not investigate the measurement properties of these measurement systems (e.g., used as an outcome measure only); (2) used these measurement systems to validate another measure [19]; (3) were published before 2004 (this being the year of the first PROMIS publication); (4) were posters or abstracts or (5) pediatric or (6) non-rehabilitation populations (e.g., mental health, focus on surgical modality such as for orthopedic injuries, etc.). We resolved disagreements by consensus or with another member of the research team. We hand-searched reference lists of included articles for possible inclusion.

Table 1 COSMIN guidelines and extended Rasch criteria for evaluated measurement properties

Measurement property	Definition	Data management and interpretation[a]	Sufficient (+) Measurement property rating criteria[b]
Content validity	The degree to which the content of a measure is an adequate reflection of the construct to be measured	COSMIN synthesis: adequate if the development paper reported clear descriptions of the measurement aim, target population, dimensions measured, and item selection process	Measure should be comprehensive, comprehensible, and relevant according to clinicians/researchers, caregivers, and patients
Structural validity	The degree to which the scores of a measure are an adequate reflection of the dimensionality of the construct to be measured	COSMIN synthesis: exploratory or confirmatory factor analysis with adequate model fit (e.g., Kaiser–Meyer–Olkin test 0.8–1.0, Bartlett's test significant) COSMIN synthesis IRT: unidimensional, locally independent, monotonic with adequate model fit	Confirmatory factor analysis: CFI or TLI or comparable measure > 0.95 OR RMSEA < 0.08 IRT: CFI or TLI or comparable measure > 0.95 OR RMSEA < 0.20 OR Q3's < 0.37 AND residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND adequate looking graphs OR item scalability > 0.30 AND $\chi 2 > 0.01$
Internal consistency	The degree of interrelatedness among the items	COSMIN synthesis: adequate analysis if completed for each unidimensional scale or subscale Meta-analysis: weighted mean and range of results calculated for Cronbach's alpha where possible	At least low evidence for sufficient structural validity AND Cronbach's alpha(s) ≥ 0.70
Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted measure is an adequate reflection of the performance of the items of the original version	COSMIN synthesis: differential item functioning analysis	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important differential item functioning for group factors (McFadden's $R^2 < 0.02$)
Reliability	The proportion of the total variance in the measurements which is due to 'true' differences between patients	COSMIN synthesis: inter-rater, intra-rater, or test–retest reliability Meta-analysis: weighted mean and range of results calculated for intra-class correlation coefficient where possible	Intra-class correlation coefficient or Spearman's correlation ≥ 0.70 OR Rater reliability: > 0.8 AND Rater separation: < 0.2
Measurement error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	COSMIN synthesis: standard error of measurement, smallest detectable change, and limits of agreement	Smallest detectable change or limit of agreement < minimal important change
Criterion validity	The degree to which the scores of a measure are an adequate reflection of a 'gold standard'	COSMIN synthesis: for predictive validity, the prediction should be clinically meaningful	Correlation with gold standard ≥ 0.70 OR area under the curve ≥ 0.70
Construct validity	The degree to which the scores of a measure are consistent with hypotheses based on the assumption that the measure validly assesses the construct to be measured	COSMIN synthesis: the result is in accordance with the hypothesis Meta-analysis: weighted mean and range of results calculated for correlations to a measure	1. Correlations with (changes in) instruments measuring similar constructs should be ≥ 0.50 2. Correlations with (changes in) instruments measuring related, but dissimilar constructs should be lower, i.e., 0.30–0.50 3. Correlations with (changes in) instruments measuring unrelated constructs should be < 0.30 4. Correlations defined under 1, 2, and 3 should differ by a minimum of 0.10
Responsiveness	The degree to which a measure can detect change over time in the construct to be measured	COSMIN synthesis: the result is in accordance with the hypothesis	Area under the curve ≥ 0.70 1. Meaningful changes between relevant (sub)groups (e.g., patients with expected high vs low levels of the construct of interest)

COSMIN COnsensus-based Standards for the selection of health Measurement Instruments, *CTT* Classical Test Theory, *SD* Standard deviation, *ICC* Intra-class correlation, *TCC* Test characteristic curve, *RMSE* Root mean square error of approximation, *CFI* comparative fit index, *TLI* Tucker–Lewis index

[a] According to the COSMIN systematic review methodology

[b] According to Terwee et al. 2009 (as recommended by COSMIN)

Data extraction

Two independent reviewers piloted and extracted the methods and results of the estimated measurement properties, study characteristics, and study population data. Extraction was structured according to COSMIN guidance [19]. We consulted a third reviewer in the case of any disagreement during extraction.

Data analysis

Two reviewers independently assessed the measurement properties in each study. They rated content validity against COSMIN criteria [19, 24]. All other measurement properties were rated using Terwee and colleagues' standards [25] as "sufficient" (+), "insufficient" (–), or "indeterminate" (?) (Table 1). When these standards did not provide cut-offs for the statistical test in the included study, we summarized and reported the measurement properties narratively. A third reviewer was consulted when there were discrepancies.

We decided a priori that there is no gold standard measure that could be used to assess the criterion validity of TBI-QoL or SCI-QoL. We set a priori hypotheses based on recommendations by de Vet and colleagues [26] for testing construct validity and responsiveness (Supplementary material 2).

Data synthesis

Within the TBI-QoL and SCI-QoL measurement systems, the research team subdivided results for synthesis by domain only. No further subdivision by diagnosis, setting, or respondent was necessary due to uniformity among these characteristics.

Two independent reviewers summarized the results for each measurement property across studies (i.e., + / – / ± / ?). Per COSMIN guidelines, gave an overall "sufficient" (+) or "insufficient" (–) rating if > 75% of measurement property results across studies were concurrent. We assigned an "inconsistent" (±) rating when no rating surpassed 75% and no adequate explanation for the inconsistency was provided. An "indeterminate" (?) rating was given when the results neither qualified as sufficient nor insufficient, meaning they had more than 25% but less than 75% sufficient ratings.

Quality assessment

Two independent reviewers assessed the methodological quality of individual studies using the COSMIN risk of bias checklist [19, 27]. Each checklist item was rated as "very good", "adequate", "doubtful", or "inadequate". The overall rating of the methodological quality for a measurement property was based on the worst item rating [19, 27]. The reviewers then graded the quality of evidence for each property per subgroup using the COSMIN-modified Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [19, 28]. The quality of evidence was rated "high", "moderate", "low", or "very low" after considering the risk of bias, inconsistency, imprecision, and indirectness.

Results

We retrieved a total of 6289 articles and 4957 articles remained following deduplication. Title and abstract screening resulted in 381 included articles. The full-text screen resulted in 146 included articles and reference checks resulted in an additional 52 included articles for a total of 198 included articles. Of these, 14 were TBI-QoL and 16 SCI-QoL. Only one article (SCI-QoL) was identified via hand-search (Fig. 1).

All collected measured patient-reported outcomes at the individual level and all but one study was conducted in English in the United States, namely Brouwers and colleagues 2022 translation of SCI-QoL into Dutch and Flemish in the Netherlands [29]. The full extraction table with all study characteristics can be found in Supplementary material 3. In the text, we present the synthesis for each domain of TBI-QoL (Table 2) and SCI-QoL below (Table 3)—the analysis for each included study can be found in Supplementary material 4.

TBI-QoL content validity evidence

TBI-QoL was conceptually developed via a series of qualitative studies [10, 11] described elsewhere [7]. Since the initial report of its psychometric properties in 2016 [8, 30], the results of several studies [8, 31–36] have indicated sufficient content validity based on high-quality evidence for all TBI-QoL domains identified in this study with two exceptions—evidence for comprehensibility for cognition general concerns and executive function are both rated as low quality. The one study reporting this information did not include all necessary methodological information.

TBI-QoL evidence for all other measurement properties by domain

A total TBI-QoL score has been reported in a variety of ways, including as a global composite score [37], a combination of 14 item banks [38], and all 20 item banks [39, 40]. In all cases, there is evidence of sufficient construct validity (18/24, 1/1, and 1/1 tested hypotheses met, respectively). However, the evidence quality for the 20-item bank scale is low, because the authors did not provide evidence that the subgroups were the same on key variables other than that being tested. The 20-item bank scale also has evidence of sufficient responsiveness (23/27 hypotheses met [40]) and structural validity. The exploratory (EFA) and confirmatory factor analyses

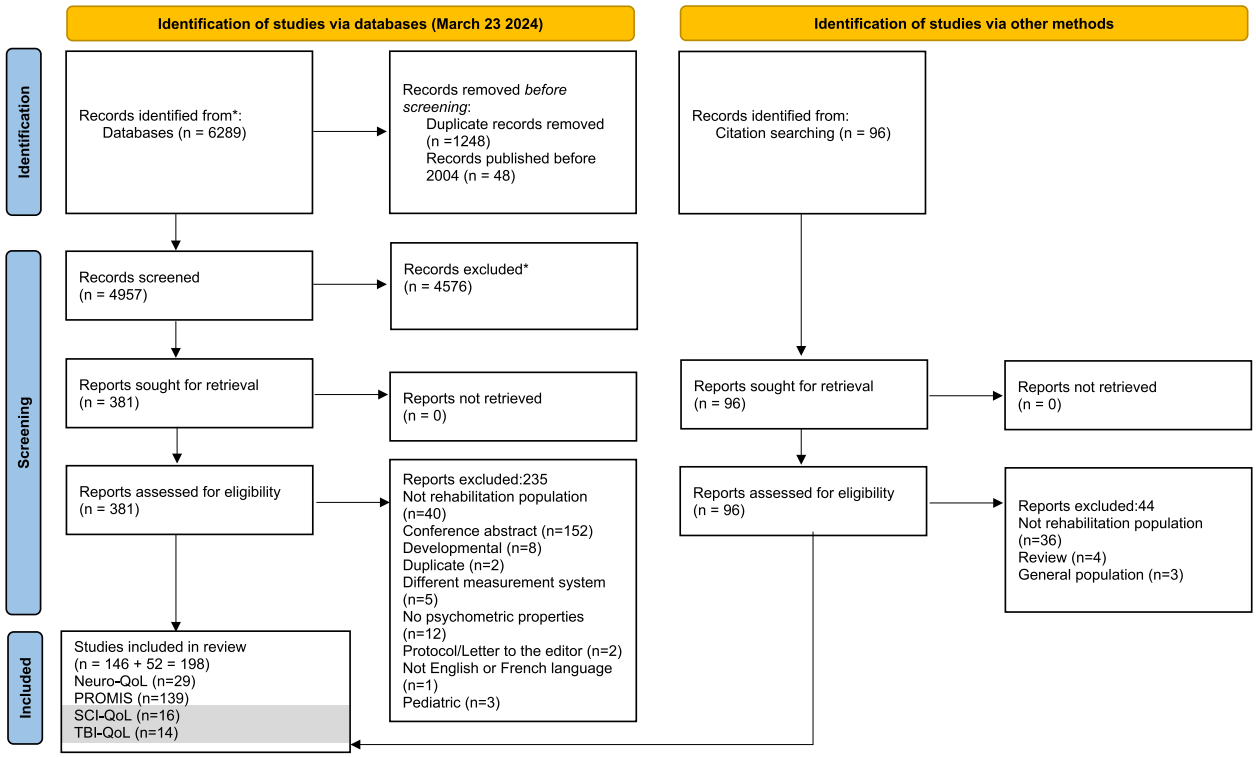

Fig. 1 PRISMA flow diagram

(CFA) suggest a 7-factor model rather than the 4-factor model found during TBI-QoL development (cognitive, emotional, physical, and social health) [8, 39].

The cognitive, emotional, physical, and social health composite scores all have sufficient construct validity (19–22/24 hypotheses met) based on high-quality evidence [37]. Furthermore, the cognitive health composite score also has sufficient responsiveness (13/14 hypotheses met) based on high-quality evidence.

The cognition–executive function and cognition–general cognitive concerns short forms both have sufficient structural validity (CFA and IRT) [8, 36], internal consistency [8, 30, 36], test–retest reliability [36], cross-cultural validity [36], construct validity (28/29 hypotheses met) [30] and responsiveness (12/14, 14/14 hypotheses met) [41]. Measurement error is currently undetermined because while the standardized error of measurement (SEM) has been calculated (2.7–3.2) [41], minimal important change (MIC) has not, and thus it cannot be determined whether the SEM is less than the MIC. Theta values are reported for the cognition–executive function (−3.61– to −0.49) and general cognitive concerns (−4.61–0.01) short forms. The computer adaptive tests (CATs) have sufficient structural validity (IRT) [8] based on high-quality evidence.

For all remaining domains (ability to participate in social activities, anger, anxiety, communication, depression, emotional and behavioral dyscontrol, fatigue, grief loss, headache pain, independence, mobility, pain interference, positive affect and wellbeing, resilience, satisfaction with social roles, self-esteem, stigma, and upper extremity) there is sufficient structural validity [8, 32–35, 42] and internal consistency [8, 30, 32–35, 42] for the short form and sufficient structural validity for the CAT [8, 33]. Theta values vary widely for all remaining domains, from −3.87 to 0.11– and −2.02 to 4.78).

There is evidence of cross-cultural validity for the ability to participate in social activities, asking for help, communication, fatigue, headache pain, independence, and pain interference short forms only. All have sufficient cross-cultural validity except for the ability to participate in social activities and pain interference. The ability to participate in social activities in short form is currently rated as having insufficient evidence for cross-cultural validity (McFadden's $r^2 = 0.393$) based on moderate quality evidence because the authors did not provide all necessary evidence that the subgroups were the same on key variables [31]. The pain interference short form is rated as insufficient due to differential item functioning (DIF) for 6 items in a low-quality study where the authors did

Table 2 TBI-QoL result by domain

			COSMIN											IRT	
Domain	**Type**	**# articles**	**Relevance**	**Comprehen-siveness**	**Compre-hensibility**	**Structural validity**	**Internal consistency**	**Cross-cultural validity**	**Reliability**	**Measure-ment error**	**Criterion validity**	**Construct validity**	**Respon-siveness**	**Structural validity**	**IRT precision estimate (theta)**
Total scale	com-posite	1										+; H			
	14 banks	1										+; H			
	20 banks	2				+; H						+; L	+; H		
Cognitive health composite score	N/A	2										+; H	+; H		
Emotional health composite score	N/A	1										+; H			
Physical health composite score	N/A	1										+; H			
Social health composite score	N/A	1										+; H			
Ability to participate in social activities	SF	3	+; H	+; H	+; H	+; H	+; H	-; M				?; H		+; H	−2.02–0.24
	CAT	2										+; H			
Anger	SF	2			+; H	+; H	+; H					+; H		+; H	−1.74–4.78
	CAT	1													
Anxiety	SF	2			+; H	+; H	+; H					+; H		+; H	−0.94–3.22
	CAT	1													
Asking for help	scale	2	+; H	+; H	+; H	-; H	-; H	+; H						+; H	
Cognition-general concerns	SF	4	+; L		+; H	+; H	+; H	+; H	+; H	?; H		+; H	+; H	+; H	−4.61–0.01
	CAT	2													
Cognition-executive function	SF	4	+; L		+; H	+; H	+; H	+; H	+; H	?; H		+; H	+; H	+; H	−3.61–(−0.49)
	CAT	2													
Communica-tion	SF	2	+; H	+; H	+; H	+; H	+; H	+; H						+; H	−3.87–1.06
	CAT	2													

Table 2 (continued)

Domain	Type	# articles	COSMIN											IRT	
			Relevance	Comprehensiveness	Comprehensibility	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness	Structural validity	IRT precision estimate (theta)
Depression	SF	2			+; H	+; H	+; H					+; H		+; H	−0.50–3.89
	CAT	1													
Emotional and behavioral dyscontrol	SF	2			+; H	+; H	+; H					+; H		+; H	−0.94–3.39
	CAT	1													
Fatigue	SF	3	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	−1.37–3.69
	CAT	1													
Grief loss	SF	2			+; H	+; H	+; H					+; H		+; H	−0.93–2.67
	CAT	1													
Headache pain	SF	3	+; H		+; H	+; H	+; H	+; H	+; H			+; H		+; H	−0.23–1.86
	CAT	1													
Independence	SF	2	+; H	+; H	+; H	+; H	+; H	+; H						+; H	−3.21–0.83
	CAT	1													
Mobility	SF	1			+; H	+; H	+; H							+; H	−3.33–0.6
	CAT	1													
Pain Interference (item bank exactly same as PROMIS)	SF	3			+; H	+; H	+; H	-; L				+; H		+; H	0.11–2.56
	CAT	1										+; H			
Positive Affect and Well-being	SF	2			+; H	+; H	+; H					?; H		+; H	−3.33–0.62
	CAT	1													
Resilience	SF	2			+; H	+; H	+; H					+; H		+; H	−3.47–2.00
	CAT	1													
Satisfaction with Social Roles	SF	1	+; H	+; H	+; H	+; H	+; H							+; H	−1.92–0.21
	CAT	1													
Self-esteem	SF	2			+; H	+; H	+; H					+; H		+; H	−2.67–0.58
	CAT	1													
Self-evaluation	SF	1			+; H										
Stigma	SF	1			+; H	+; H	+; H							+; H	−2.86–0.82
	CAT	1													

Table 2 (continued)

Domain	Type	# articles	COSMIN											IRT	
			Relevance	Comprehen-siveness	Compre-hensibility	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measure-ment error	Criterion validity	Construct validity	Respon-siveness	Structural validity	IRT precision estimate (theta)
Upper Extremity	SF	1			+; H	+; H	+; H							+; H	−3.66-(−0.69)
	CAT	1													

Table 3 SCI-QoL results by domain

Domain	Type	# articles	COSMIN											IRT	
			Relevance	Comprehensiveness	Comprehensibility	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness	Structural validity	IRT precision estimate (theta)
Ability to participate in social activities	SF	2	+; H	+; H	+; H	+; H	+; H	+; M	+; H			+; H		+; H	−2.3–1.2
	CAT	1										+; H			
Ambulation	SF	1	+; H	+; H	+; H	+; H	+; H					+; H		+; H	0.48–3.35
	CAT	1										+; H			
Anxiety	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; M			+; H		+; H	−0.5–3.0
	CAT	2							+; H			+; H			
Basic mobility	SF	1	+; H	+; H	+; H	+; H	+; H		+; M			+; H		+; H	−2.15–1.73
	CAT	1										+; H			
Bladder management difficulties	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	0.3–2.4
	CAT	1										+; H			
Bowel management difficulties	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	−0.3–2.5
	CAT	1										+; H			
Bladder complication	Scale	2	+; H	+; H	+; H	+; H	+; H	+; H	?; H					+; H	−0.33–3.34
Depression	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	−0.7–2.8
	CAT	1										+; H			
Fine motor function	SF	1	+; H	+; H	+; H	+; H	+; H		+; M			+; H		+; H	−1.82–1.05
	CAT	1										+; H			
Grief loss	SF	2	+; H	+; H	+; H	+; H	+; H	+; L	+; H			+; H		+; H	−0.8–1.8
	CAT	1										+; H			
Independence	SF	1	+; H	+; H	+; H	+; H	+; H		+; M			+; H		+; H	−2.59–1.05
	CAT	1										+; H			
Pain behavior	Scale	2	+; H	+; H	+; H	+; H	+; H	+; M						+; H	−0.62–2.31
Pain interference	SF	2	+; H	+; H	+; H	+; H	+; H	+; M	+; H			+; H		+; H	0.00–2.66
	CAT	2										+; H			
Positive affect and well-being	SF	2	+; H	+; H	+; H	+; H	+; H	+; M	+; H			+; H		+; H	−2.9–1.2
	CAT	1										+; H			
Pressure ulcers	Scale	2	+; H	+; H	+; H	+; M	+; H		+; H			+; H		+; M	0.7–1.8
Psychological trauma	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	0.6–2.5
	CAT	1										+; H			

Table 3 (continued)

Domain	Type	# articles	COSMIN											IRT	
			Relevance	Comprehen-siveness	Compre-hensibility	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measure-ment error	Criterion validity	Construct validity	Respon-siveness	Structural validity	IRT precision estimate (theta)
Resilience	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	−3.1–0.9
	CAT	2										+; H			
Satisfaction with social roles and activi-ties	SF	2	+; H	+; H	+; H	+; H		+; H	+; H			+; H		+; H	−2.3–1.6
	CAT	1										+; H			
Self-care	SF	1	+; H	+; H	+; H	+; H	+; H		+; M			+; H		+; H	−3.94–1.38
	CAT	1										+; H			
Self-esteem	SF	2	+; H	+; H	+; H	-; H	-; H	+; H	+; M			+; H		+; H	−2.7–0.7
	CAT	2										+; H			
Stigma	SF	2	+; H	+; H	+; H	+; H	+; H	+; H	+; H			+; H		+; H	–0.5–2.4
	CAT	1										+; H			
Wheelchair mobility (power and manual)	SF-power	2	+; H	+; H	+; H	+; H	+; H		+; M			+; H		+; H	−3.79–1.64
	SF-manual	2				+; H	+; H		+; M			+; H			
	CAT	1										+; H			

not report evidence that the subgroups were the same on key variables other than that being tested [33].

Fatigue [35] and headache pain [34] short forms have sufficient test–retest reliability based on high-quality evidence. The short forms for various domains—including anger, anxiety, depression, emotional and behavioral dyscontrol, fatigue, grief loss, headache pain, pain interference, resilience, and self-esteem—predominantly met their respective hypotheses, with all achieving at least 83% of hypotheses met [30, 31, 41] (add references). The CAT for the ability to participate in social activities [31] (4/4 hypotheses met) and pain interference (4/4 hypotheses met) have sufficient construct validity. The ability to participate in social roles and positive affect and well-being construct validity is currently indeterminant (21/30 and 21/29 hypotheses met, respectively) [30, 31, 41].

SCI-QoL content validity evidence

SCI-QoL was conceptually developed via a series of extensive qualitative studies [12, 13] described elsewhere [6, 12]. The articles that first published SCI-QoL included results indicating sufficient content validity based on high-quality evidence across 21 domains in 2015 [43–53]. Since then, a manuscript focused on the development of the pain interference and pain behavior domains has been published using the same dataset in the original development of the measures [54]. One article confirming sufficient content validity (comprehensibility) has been published outside of these development studies, in which the authors conduct cognitive debriefing as part of translating SCI-QoL from English to Dutch/Flemish [29].

SCI-QoL evidence for all other measurement properties by domain

For nearly all SCI-QoL domains there is evidence of sufficient structural validity (CFA and IRT) [44–51, 54], internal consistency [43–46, 48–50, 54, 55] and construct validity (all 100% hypotheses met) [43–46, 48, 49] for the short forms and construct validity for the CATs (all 100% hypotheses met except resilience (6/8 hypotheses met [55], 13/15 hypotheses met [56]). The exceptions are insufficient structural validity for the positive affect and wellbeing, and self-esteem short forms because neither CFI (0.947, 0.946) nor RMSEA (0.094, 0.087) met the cut-offs ($>0.95, <0.06$) which in turn leads to insufficient internal consistency for these domains [43]. Similarly, the independence domain has a small number of items (8) and a high RMSEA (0.111) indicating possible multidimensionality; however, we maintained its rating of sufficient due to CFI meeting cut-off values (0.980). The pressure ulcers short form currently has only moderate level evidence for structural validity (CFA and IRT) due to a borderline sample size ($n=189$) [45]. Two domains do not yet have evidence of construct validity, namely bladder complications and pain behavior.

Most SCI-QoL short forms have evidence of sufficient cross-cultural validity and test–retest reliability. Among the domains for which there is evidence of these properties, only the test–retest reliability of the bladder complication scale is currently rated as indeterminant (ICC = 0.69) [49]. There are several instances across domains in which the quality of evidence is rated as moderate as opposed to high for both measurement properties. The reason is similar—the authors did not provide all necessary evidence that the subgroups were the same on key variables other than being tested or that the individuals were stable over the retest period.

There is sufficient structural validity for all SCI-QoL CATs based on IRT methods. In all cases, the evidence is of high quality. Theta values vary widely across SCI-QoL domains, from − 3.1 to 0.70 to 0.7 to 3.0.

Discussion

We conducted a systematic review to assess the strength and quality of the measurement properties of TBI-QoL and SCI-QoL, as guided by COSMIN's systematic review framework. COSMIN's framework does differ from others in the field (e.g., ISOQOL [3], ISPOR [57, 58], HealthMeasures reporting guidelines [59]) as it provides specific requirements for strength (e.g., Cronbach's alpha > 0.7) and quality (e.g., minimum sample sizes for certain statistical tests). By applying COSMIN's requirements, this manuscript serves as a comprehensive resource for researchers and clinicians, offering recommendations for the evidence-based application of TBI-QoL and SCI-QoL based on current evidence. It should be noted that these recommendations may evolve as new evidence emerges.

In this review, we found that all TBI-QoL and SCI-QoL short forms have sufficient content validity, structural validity (CFA and IRT), internal consistency, and construct validity. However, neither TBI-QoL nor SCI-QoL has evidence for criterion validity or statistical estimates from Rasch analysis, and both have limited cross-cultural validity evidence and variation in theta values across domains. TBI-QoL has evidence for sufficient test–retest reliability across domains. In contrast, SCI-QoL has evidence of sufficient test–retest reliability only for pain behavior and ambulation domains. Regarding the CAT, both TBI-QoL [8, 31, 33] and SCI-QoL [46, 55] have some limited evidence for construct validity which suggests that it measures the same construct as the short form and item bank. Finally, all available evidence of responsiveness is sufficient; however, since there is no reported MIC (only MID reported), the available measurement error evidence is currently rated as indeterminant.

Taken together, these results indicate that clinicians could integrate TBI-QoL and SCI-QoL into routine assessments and incorporate findings into multidisciplinary care strategies, including personalized care planning that focuses on areas most relevant to each patient's quality of life (i.e., describe and assess patients at a single time point). However, clinicians should be aware that they may encounter limitations when using TBI-QoL or SCI-QoL for certain patient populations, and that there is limited evidence to support using these measures to evaluate change over time (e.g., between admission and discharge) and no evidence to support them in predicting patient outcomes (e.g., using the admission score to predict a likely outcome at discharge).

Based on the results of this review, researchers can use TBI-QoL and SCI-QoL for outcome measurement and evaluation in studies, although they should apply caution when using domains for which there is not yet evidence of sufficient cross-cultural validity in diverse populations. The evidence for TBI-QoL is slightly weaker than SCI-QoL due to a lack of evidence of reliability which is required to understand if the measure yields consistent scores. The standard test interval is 7 days for PROMIS instruments [60]. Future research contributing to this measurement property should adhere to this standard. Other fruitful directions for future research include exploring better-calibrated item sets or improving the measurement efficiency of fixed-length short forms and/or CATs. Further research is also required to demonstrate that SEM is less than MIC to result in a sufficient rating for measurement error for both measurement systems [19]. Together, an estimated MIC along with sufficient evidence of measurement error and responsiveness would aid in interpreting changes in TBI-QoL and SCI-QoL scores.

While the included articles usually provided evidence of high quality, there were recurring methodological choices that did not fit COSMIN's standard, and thus we downgraded the quality scores. The most common was a lack of hypotheses for both the magnitude and direction of expected values for construct validity or responsiveness. Per COSMIN guidelines, the authors of this review had to assign testable hypotheses for the evidence from these studies to be interpretable [26]. Due to the extensive list of comparisons between measures, subgroup analyses, and pre-post-tests that we have compiled, we recommend that researchers investigating construct validity in TBI-QoL and SCI-QoL refer to the a priori hypotheses that we developed for this review (Supplementary material 2). We also noted minor issues with reporting all necessary details to confirm the stability of patients between administrations when estimating test–retest reliability and to confirm that subgroups are the same other than for the variable being tested for cross-cultural validity. Ensuring this information is reported in future manuscripts concerning the measurement properties of TBI-QoL and SCI-QoL domains would advance our understanding of this measurement system.

This TBI-QoL and SCI-QoL review was part of a larger review that included Neuro-QoL [20] and PROMIS measurement systems. Similar findings regarding evidence-based use were found across many domains and diagnoses in which these measurement systems are used. There is limited but strong psychometric evidence of the use of anger, anxiety, depression, emotional support, fatigue, pain interference, physical function, sleep-related impairments, and satisfaction with social roles PROMIS domains by SCI patients [61–64] and the same for all domains but sleep-related impairments and satisfaction with social roles [63–65]. In all cases, there is more psychometric information available for TBI and SCI populations for the SCI-QoL and TBI-QoL domains rather than PROMIS domains.

The TBI-QoL and SCI-QoL measurement systems are versatile, providing options to use short forms, CATs, or the profile, as well as the opportunity to compare across populations given the use of IRT and calibrations of items. Among the studies to date, these measurement systems perform comparably or sometimes better than other measures (e.g., when compared to other measures to generate evidence of construct validity and responsiveness). Thus, TBI-QoL and SCI-QoL are ideal for measuring patient-reported outcomes in clinical practice and in research.

Strengths and limitations

For both TBI-QoL and SCI-QoL, there were only 1–3 articles per domain. Although the existing research is rigorous, there is a need for future studies to replicate the current evidence in varied contexts and to expand upon it. This would result in stronger recommendations for the use of TBI-QoL and SCI-QoL.

We had to go beyond the guidance in COSMIN's systematic review protocol to provide ratings for some evidence for this review. This included deciding to apply a score of 'sufficient' for structural validity when at least one criterion for CFA was met (e.g., CFI) even if others were not (e.g., RMSEA), and for cross-cultural validity when DIF was found between group factors, but they were found to be non-significant. The ratings for these two measurement properties may be more optimistic because of this rating strategy.

Conclusion

Both TBI-QoL and SCI-QoL have sufficient content validity, structural validity, internal consistency, and construct validity, but with a few exceptions for some TBI-QoL domains, only SCI-QoL has sufficient evidence of cross-cultural validity and reliability. Based on the current evidence, rehabilitation researchers and clinicians can apply TBI-QoL and SCI-QoL to describe and evaluate patients. The results of this review highlight that future research investigating TBI-QoL and SCI-QoL's measurement error, responsiveness, criterion validity, and TBI-QoL's reliability is required to use TBI-QoL or SCI-QoL to evaluate change over time or to predict patient outcomes. As gaps in the evidence base are addressed, more widespread use of TBI-QoL and SCI-QoL could be possible, which would contribute to a patient-centered model of care.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13643-024-02722-x.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Acknowledgements

We gratefully acknowledge the contributions of Alia Osman for their support of this review, including in article piloting and screening.

During this work, Rebecca Ataman was supported by a doctoral scholarship from the Centre de recherche interdisciplinaire en Réadaptation (CRIR) and from the Fonds de Recherche Quebec—Santé (FRQ-S). Dr. Sara Ahmed is the principal investigator for the Biomedical Research and Informatics Living Laboratory for Innovative Advances of New Technologies (BRILLIANT) in Community Mobility Rehabilitation program which is funded by the Canadian Foundation of Innovation and the Ministry of Health of Quebec (#36053), the Initiatives for the Development of New Technologies and Innovative Practices in Rehabilitation – INSPIRE, and the Fonds de Recherche du Québec-Santé. Henry is supported by a doctoral scholarship from the Division of Experimental Medicine, Department of Medicine, McGill University. Dr. Adria Quigley was supported by a Gilead/CIHR Canadian HIV Trials Network Postdoctoral Fellowship Award.

Authors' contributions

RA2 and SA conceptualized this review. All authors contributed to the study acquisition, analysis, synthesis, and/or interpretation. RA1 drafted the manuscript and all authors critically reviewed and approved the final version for publication.

Funding

There is no funding to report in support of this manuscript.

Data availability

The data supporting the findings of this study are available within the article and its supplementary materials.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

[1]School of Physical and Occupational Therapy, McGill University, Montréal, Québec, Canada. [2]Centre for Interdisciplinary Research in Rehabilitation of Greater Montréal, Montréal, Québec, Canada. [3]Quality Division, Clinical Institutes and Quality Programs, Ontario Health, Toronto, Canada. [4]Department of Rehabilitation, Princess Nora Bint Abdul Rahman University, Riyadh, Saudi Arabia. [5]School of Physiotherapy, Dalhousie University, Forrest Building, 5869 University Avenue, PO Box 15000, Halifax, NS B3H 4R2, Canada. [6]Nova Scotia Health Authority, Nova Scotia Rehabilitation and Arthritis Centre, 1341 Summer St, Halifax, NS B3H 4K4, Canada. [7]Division of Experimental Medicine, McGill University, Montreal, QC, Canada. [8]Center for Outcomes Research and Evaluation (CORE), Research Institute of the McGill University Health Center, Montreal, QC, Canada.

Received: 13 July 2024 Accepted: 25 November 2024
Published online: 21 January 2025

References

1. Abrams D, Davidson M, Harrick J, Harcourt P, Zylinski M, Clancy J. Monitoring the change: current trends in outcome measure usage in physiotherapy. Man Ther. 2006;11:46–53.
2. Law M, Baum C, Dunn W. Measuring occupational performance: Supporting best practice in occupational therapy. Thorofare, NJ: Slack Incorporated ;2005.
3. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, Lyons JC. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Quality of life research. 2013;22(8):1889–905. https://doi.org/10.1007/s11136-012-0344-y.
4. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Cook K. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol. 2010;63(11):1179–94. https://doi.org/10.1016/j.jclinepi.2010.04.011.
5. Gershon RC, Lai JS, Bode R, Choi S, Moy C, Bleck T, Cella D. Neuro-QOL: Quality of life item banks for adults with neurological disorders: Item development and calibrations based upon clinical and general population testing. Qual Life Res. 2012;21(3):475–86. https://doi.org/10.1007/s11136-011-9958-8.
6. Tulsky DS, Kisala PA, Victorson D, Tate DG, Heinemann AW, Charlifue S, Bombardier CH. Overview of the spinal cord injury–quality of life (SCI-QOL) measurement system. J Spinal Cord Med. 2015;38(3):257–69.
7. Tulsky DS, Kisala PA. An overview of the traumatic brain injury–quality of life (TBI-QoL) measurement system. J Head Trauma Rehabil. 2019;34(5):281–8.
8. Tulsky DS, Kisala PA, Victorson D, Carlozzi N, Bushnik T, Sherer M, Englander J. TBI-QOL: Development and calibration of item banks to measure patient reported outcomes following traumatic brain injury. J Head Trauma Rehabil. 2016;31(1):40–51. https://doi.org/10.1097/HTR.0000000000000131.
9. Amtmann D, Cook K, Johnson K, Cella D. The PROMIS initiative: involvement of rehabilitation stakeholders in development and examples of applications in rehabilitation research. Arch Phys Med Rehabil. 2011;29(10):S12–9. https://doi.org/10.1016/j.apmr.2011.04.025.
10. Carlozzi N, Tulsky D, Kisala PA. Traumatic brain injury patient-reported outcome measure: identification of health-related quality-of-life issues relevant to individuals with traumatic brain injury. Arch Phys Med Rehabil. 2011;92(10(suppl)):S52–60.
11. Carlozzi NE, Kratz AL, Sander AM, Chiaravalloti ND, Brickell TA, Lange RT, Tulsky DS. Health-related quality of life in caregivers of individuals with traumatic brain injury: development of a conceptual model. Arch Phys Med Rehabil. 2015;96(1):105–13.
12. Tulsky DS, Kisala PA, Victorson D, Tate D, Heinemann AW, Amtmann D, Cella D. Developing a contemporary patient-reported outcomes measure for spinal cord injury. Arch Phys Med Rehabil. 2011;92(10):S44–51.
13. Kisala PA, Tulsky D. Opportunities for CAT applications in medical rehabilitation: development of targeted item banks. J Appl Meas. 2010;11(3):315–30.

14. Bergquist TF, Moessner AM, Mandrekar J, Ransom JE, Dernbach NL, Kendall KS, Brown AW. CONNECT: a pragmatic clinical trial testing a remotely provided linkage to service coordination after hospitalization for TBI. Brain Inj. 2022;36(2):147–55.
15. Spungen AM, Bauman WA, Biswas K, Jones KM, Snodgrass AJ, Goetz LL, Asselin PK. The design of a randomized control trial of exoskeletal-assisted walking in the home and community on quality of life in persons with chronic spinal cord injury. Contemp Clin Trials. 2020;96:106102. https://doi.org/10.1016/j.cct.2020.106102.
16. Basch, E. Patient-reported outcomes—harnessing patients' voices to improve clinical care. New England J Med. 2017;376(2):105–8.
17. Toyinbo PA, Vanderploeg RD, Donnell AJ, Mutolo SA, Cook KF, Kisala PA, Tulsky DS. Development and initial validation of military deployment-related TBI quality-of-life item banks. J Head Trauma Rehabil. 2016;31(1):52–61. https://doi.org/10.1097/HTR.0000000000000089.
18. Choukou MA, Best KL, Craven BC, Hitzig SL. Identifying and classifying quality of life tools for assessing neurogenic bowel dysfunction after spinal cord injury. Topics Spinal Cord Injury Rehabil. 2019;25(1):1–22. https://doi.org/10.1310/sci18-00019.
19. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147–57. https://doi.org/10.1007/s11136-018-1798-3.
20. Ataman R, Alhasani R, Auneau-Enjalbert L, Quigley A, Michael HU, Ahmed S. The psychometric properties of the Quality of Life in Neurological Disorders (Neuro-QoL) measurement system in neurorehabilitation populations: a systematic review. J Patient-Reported Outcomes. 2024;8(1):106. https://doi.org/10.1186/s41687-024-00743-7.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, Tugwell P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Ann Internal Med. 2009;151(4):264–9. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.
22. Terwee C, Jansma E, Riphagen I, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Health Res. 2009;18(8):1115–23. https://doi.org/10.1007/s11136-009-9528-5.
23. Bramer WM, Giustini D, De Jong GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in endnote. J Med Libr Assoc. 2016;104(3):240–3. https://doi.org/10.3163/1536-5050.104.3.014.
24. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, De Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010;10(1):22. https://doi.org/10.1186/1471-2288-10-22.
25. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.
26. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge, UK: Cambridge University Press; 2011.
27. Mokkink LB, Boers M, Van Der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, Terwee CB. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments : a Delphi study. BMC Med Res Methodol. 2020;20(1):1–13. https://doi.org/10.1186/s12874-020-01179-5.
28. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6. https://doi.org/10.1136/bmj.39489.470347.ad.
29. Brouwers EMJR, Terwee CB, Roorda LD, Hosman AF, van de Meent H, Bartels RHMA. Dutch-Flemish translation and cross-cultural adaption of the Spinal Cord Injury-Quality of Life (SCI-QOL) short forms. Eur Spine J. 2022;31(1):1349–57.
30. Lange RT, Brickell TA, Bailie JM, Tulsky DS, French LM. Clinical utility and psychometric properties of the Traumatic Brain Injury Quality of Life Scale (TBI-QOL) in US military service members. J Head Trauma Rehabil. 2016;31(1):62–78.
31. Heinemann AW, Kisala PA, Boulton AJ, Sherer M, Sander AM, Chiaravalloti N, Tulsky DS. Development and calibration of the TBI-QOL ability to participate in social roles and activities and TBI-QOL satisfaction with social roles and activities item banks and short forms. Arch Phys Med Rehabil. 2020;101(1):20–32.
32. Kisala PA, Tulsky DS, Boulton AJ, Heinemann AW, Victorson D, Sherer M, Hanks R. Development and psychometric characteristics of the TBI-QOL Independence item bank and short form and the TBI-QOL Asking for Help Scale. Arch Phys Med Rehabil. 2020;101(1):33–42.
33. Carlozzi NE, Kisala PA, Boulton AJ, Roth E, Kratz AL, Sherer M, Tulsky DS. Measuring pain in TBI: development of the TBI-QOL pain interference item bank and short form. Arch Phys Med Rehabil. 2020;101(1):11–9.
34. Tulsky DS, Tyner CE, Boulton AJ, Kisala PA, Heinemann AW, Roth EJ, Carlozzi NE. Development of the TBI-QOL headache pain item bank and short form. J Head Trauma Rehabil. 2019;34(5):298–307.
35. Kisala PA, Bushnik T, Boulton AJ, Hanks RA, Kolakowsky-Hayner SA, Tulsky DS. Measuring fatigue in TBI: development of the TBI-QOL fatigue item bank and short form. J Head Trauma Rehabil. 2019;34(5):289–97.
36. Carlozzi NE, Tyner CE, Kisala PA, Boulton AJ, Sherer M, Chiaravalloti N, Tulsky DS. Measuring self-reported cognitive function following TBI: development of the TBI-QOL executive function and cognition-general concerns item banks. J Head Trauma Rehabil. 2019;34(5):308–25.
37. Tyner CE, Boulton AJ, Sherer M, Kisala PA, Glutting JJ, Tulsky DS. Development of composite scores for the TBI-QOL. Arch Phys Med Rehabil. 2020;101(1):43–53.
38. Lange RT, French LM, Lippa SM, Bailie JM, Brickell TA. Posttraumatic stress disorder is a stronger predictor of long-term neurobehavioral outcomes than traumatic brain injury severity. J Trauma Stress. 2020;33(3):318–29.
39. Sherer M, Poritz JM, Tulsky D, Kisala P, Leon-Novelo L, Ngan E. Conceptual structure of health-related quality of life for persons with traumatic brain injury: confirmatory factor analysis of the TBI-QOL. Arch Phys Med Rehabil. 2020;101(1):62–71.
40. Poritz JM, Sherer M, Kisala PA, Tulsky D, Leon-Novelo L, Ngan E. Responsiveness of the traumatic brain injury–quality of life (TBI-QOL) measurement system. Arch Phys Med Rehabil. 2020;101(1):54–61.
41. Tyner CE, Kisala PA, Boulton AJ, Sherer M, Chiaravalloti ND, Sander AM, Tulsky DS. Responsiveness of the traumatic brain injury quality of life cognition banks in recent brain injury. Front Hum Neurosci. 2022;4(16):763311. https://doi.org/10.3389/fnhum.2022.763311.
42. Cohen ML, Kisala PA, Boulton AJ, Carlozzi NE, Cook CV, Tulsky DS. Development and psychometric characteristics of the TBI-QOL communication item bank. J Head Trauma Rehabil. 2019;34(5):326–39.
43. Bertisch H, Kalpakjian CZ, Kisala PA, Tulsky DS. Measuring positive affect and well-being after spinal cord injury: development and psychometric characteristics of the SCI-QOL Positive Affect and Well-being bank and short form. J Spinal Cord Med. 2015;38(3):356–65.
44. Kalpakjian CZ, Tulsky DS, Kisala PA, Bombardier CH. Measuring grief and loss after spinal cord injury: development, validation and psychometric characteristics of the SCI-QOL Grief and Loss item bank and short form. J Spinal Cord Med. 2015;38(3):347–55.
45. Kisala PA, Tulsky DS, Choi SW, Kirshblum SC. Development and psychometric characteristics of the SCI-QOL pressure ulcers scale and short form. J Spinal Cord Med. 2015;38(3):303–14.
46. Tulsky DS, Kisala PA, Victorson D, Choi SW, Gershon R, Heinemann AW, Cella D. Methodology for the development and calibration of the SCI-QOL item banks. J Spinal Cord Med. 2015;38(3):270–87.
47. Victorson D, Tulsky DS, Kisala PA, Kalpakjian CZ, Weiland B, Choi SW. Measuring resilience after spinal cord injury: development, validation and psychometric characteristics of the SCI-QOL Resilience item bank and short form. J Spinal Cord Med. 2015;38(3):366–76.
48. Kisala PA, Tulsky DS, Kalpakjian CZ, Heinemann AW, Pohlig RT, Carle A, Choi SW. Measuring anxiety after spinal cord injury: Development and psychometric characteristics of the SCI-QOL anxiety item bank and linkage with GAD-7. J Spinal Cord Med. 2015;38(3):315–25.
49. Tulsky DS, Kisala PA, Tate DG, Spungen AM, Kirshblum SC. Development and psychometric characteristics of the SCI-QOL bladder management difficulties and bowel management difficulties item banks and short forms and the SCI-QOL bladder complications scale. J Spinal Cord Med. 2015;38(3):288–302.
50. Tulsky DS, Kisala PA, Kalpakjian CZ, Bombardier CH, Pohlig RT, Heinemann AW, Choi SW. Measuring depression after spinal cord injury: Development and psychometric characteristics of the SCI-QOL depression item bank and linkage with PHQ-9. J Spinal Cord Med. 2015;38(3):335–46.

51. Kisala PA, Victorson D, Pace N, Heinemann AW, Choi SW, Tulsky DS. Measuring psychological trauma after spinal cord injury: development and psychometric characteristics of the SCI-QOL psychological trauma item bank and short form. J Spinal Cord Med. 2015;38(3):326–34.
52. Kisala PA, Tulsky DS, Pace N, Victorson D, Choi SW, Heinemann AW. Measuring stigma after spinal cord injury: development and psychometric characteristics of the SCI-QOL Stigma item bank and short form. J Spinal Cord Med. 2015;38(3):386–96.
53. Kalpakjian CZ, Tate DG, Kisala PA, Tulsky DS. Measuring self-esteem after spinal cord injury: development, validation and psychometric characteristics of the SCI-QOL Self-esteem item bank and short form. J Spinal Cord Med. 2015;38(3):377–85.
54. Cohen ML, Kisala PA, Dyson-Hudson TA, Tulsky DS. Measuring pain phenomena after spinal cord injury: development and psychometric properties of the SCI-QOL Pain Interference and Pain Behavior assessment tools. J Spinal Cord Med. 2018;41(3):267–80.
55. Kuzu D, Kallen MA, Kalpakjian CZ, Kratz AL. Psychometric properties of the spinal cord injury-quality of life (SCI-QOL) Resilience item bank in a sample with spinal cord injury and chronic pain. Qual Life Res. 2022;31(3):927–36.
56. Kuzu D, Kallen MA, Kratz AL. Psychometric properties of the spinal cord injury-quality of life (SCI-QOL) resilience short form in a sample with spinal cord injury. Arch Phys Med Rehabil. 2024;105(1):59–66. https://doi.org/10.1016/j.apmr.2023.10.002.
57. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of Existing Patient-Reported Outcome (PRO) Instruments and their modification: the ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. Value in Health. 2009;12(8):1075–83. https://doi.org/10.1111/j.1524-4733.2009.00603.x.
58. O'Donohoe P, Reasner DS, Kovacs SM, Byrom B, Eremenco S, Barsdorf AI, Coons SJ. Updated recommendations on evidence needed to support measurement comparability among modes of data collection for patient-reported outcome measures: a good practices report of an ISPOR task force. Value Health. 2023;26(5):623–33. https://doi.org/10.1016/j.jval.2023.01.001.
59. Hanmer J, Jensen RE, Rothrock N. A reporting checklist for HealthMeasures' patient-reported outcomes: ASCQ-Me, Neuro-QoL, NIH Toolbox, and PROMIS. J Patient-Reported Outcomes. 2020;4(1):21. https://doi.org/10.1186/s41687-020-0176-4.
60. PROMIS. PROMIS Adult Profile Instruments. 2021. Retrieved from https://staging.healthmeasures.net/images/PROMIS/manuals/PROMIS_Adult_Profile_Scoring_Manual.pdf.
61. Cook KF, Bamer AM, Amtmann D, Molton IR, Jensen MP. Six patient-reported outcome measurement information system short form measures have negligible age- or diagnosis-related differential item functioning in individuals with disabilities. Arch Phys Med Rehabil. 2012;93(7):1289–91. https://doi.org/10.1016/j.apmr.2011.11.022.
62. Cook KF, Kallen MA, Bombardier C, Bamer AM, Choi SW, Kim J, Amtmann D. Do measures of depressive symptoms function differently in people with spinal cord injury versus primary care patients: the CES-D, PHQ-9, and PROMIS®-D. Qual Life Res. 2017;26(1):139–48. https://doi.org/10.1007/s11136-016-1363-x.
63. Patrick SD, Sanders G, Boulton AJ, Tulsky DS. Measurement invariance of physical, mental, and social health PROMIS measures across individuals with spinal cord injury and traumatic brain injury. Qual Life Res. 2022;31(7):2223–33. https://doi.org/10.1007/s11136-022-03114-1.
64. Kisala PA, Boulton AJ, Cohen ML, Slavin MD, Jette AM, Charlifue S, Tulsky DS. Interviewer- versus self-administration of PROMIS® measures for adults with traumatic injury. Health Psychol. 2019;38(5):435–44. https://doi.org/10.1037/hea0000685.
65. Heinemann AW, Miskovic A, Semik P, Wong A, Dashner J, Baum C, Gray DB. Measuring environmental factors: unique and overlapping international classification of functioning, disability and health coverage of 5 instruments. Arch Phys Med Rehabil. 2016;97(12):2113–22. https://doi.org/10.1016/j.apmr.2016.05.021.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.