ISCoS The International Spinal Cord Society

ARTICLE

Standardized administration and scoring guidelines for the Spinal Cord Independence Measure Version 3.0 (SCIM-III)

Rachel Y. Kim[1]✉, Christina Calhoun Thielen[1], Giulia Heydeman[2] and M. J. Mulcahey[1]

STUDY DESIGN: Qualitative studies.
OBJECTIVE: To develop clear and specific administration and scoring procedures for the Spinal Cord Independence Measure Version 3.0 as a performance-based and interview assessment.
SETTING: Research lab.
METHODS: Modified Delphi Technique survey methods were used in this study. Previously developed SCIM-III administration and scoring procedures for performance-based and interview versions were presented to clinicians experienced in SCI and SCIM-III using the Qualtrix (Qualtrics, Provo, UT) online survey platform. Summary and descriptive statistics were used to assess the percent agreement survey responses.
RESULTS: Three survey rounds were necessary to achieve 80% agreement or above for the performance-based version. Two survey rounds were necessary to achieve 80% agreement or above on the interview version.
CONCLUSIONS: This study describes the development of standardized administration and scoring procedures for the self-care and mobility sub-scales of the SCIM-III as a performance-based and interview version.

Spinal Cord (2023) 61:296–306; https://doi.org/10.1038/s41393-023-00891-5

INTRODUCTION

The Spinal Cord Independence Measure (SCIM-III) evaluates the level of independence and ability of individuals with spinal cord injury (SCI) to perform routine daily tasks [1, 2]. It is comprised of 19 items in three domains: self-care (6 items), respiration and sphincter management (4 items), and mobility (9 items) [3]. Each domain is scored on an ordinal scale based on the clinical relevance and importance for individuals with SCI [3, 4]. Scoring accounts for the amount of assistance provided by another person and use of adaptive equipment by assigning higher scores for less assistance or aids [5].

Since its introduction in 1997, the SCIM-III has been frequently used in SCI clinical trials. Three large-scale adult studies [6–8] evaluated the validity and reliability of the SCIM-III as a functional recovery instrument and established moderate-to-strong reliability and validity. Additional work examined clinically important differences in SCIM-III scores [9], established target values for SCIM-III for neurological levels in complete SCI [10], and developed and validated a self-report version [11]. Most recently, scores from the ISNCSCI motor examination [12] and SCIM-III scores from select items have been used to develop the Spinal Cord Ability Ruler (SCAR) to enable longitudinal assessment of volitional performance on a linear metric from the time of injury [12].

Despite its widespread use, standardized guidelines for administering and scoring the SCIM-III were never developed, which has led to variability in interpretation of scores [13] and feedback that administration and scoring descriptions are incomplete and unclear [14]. Variability in administering and scoring standardized clinical outcome assessments, such as the SCIM-III [14], has implications for clinical practice and clinical trials. Reimbursement, length of stay, and potential for over-estimating or under-estimating function are at risk. For clinical trials, variability in administration and scoring poses a threat to reliability within a single trial and limits the opportunity for comparison of outcomes and pooling of data across several trials. Lack of standardization can also impact test results and measurement properties [14]. For performance-based measures, such as the SCIM-III, reliance on human raters to translate human motor performance into a metric underscores the importance for clear administration and scoring procedures.

Performance-based assessments involve having the participant perform certain movements or tasks, which allow for a direct assessment of the participant's abilities. Although performance-based assessments may be more objective and potentially less biased than self-report and interview-based assessments, they are time-consuming and may not be feasible within all practice environments and clinical trial settings due to limited time and resources, such as equipment and appropriate environments for observation [15, 16], and transportation challenges associated with returning to the clinical setting. Reliability of SCIM-III scores when obtained through self-report and interview are acceptable, and these modes of administration require less time (10–15 min for interview compared to 30–45 min by observation), allow for rapid data collection, and can be utilized across multiple settings

[1]Center for Outcomes and Measurement, Jefferson College of Rehabilitation Sciences, Thomas Jefferson University, Philadelphia, PA, USA. [2]Main Line Health, Philadelphia, PA, USA. ✉email: rachel.kim.2@jefferson.edu

Received: 15 November 2022 Revised: 22 February 2023 Accepted: 10 March 2023
Published online: 25 March 2023

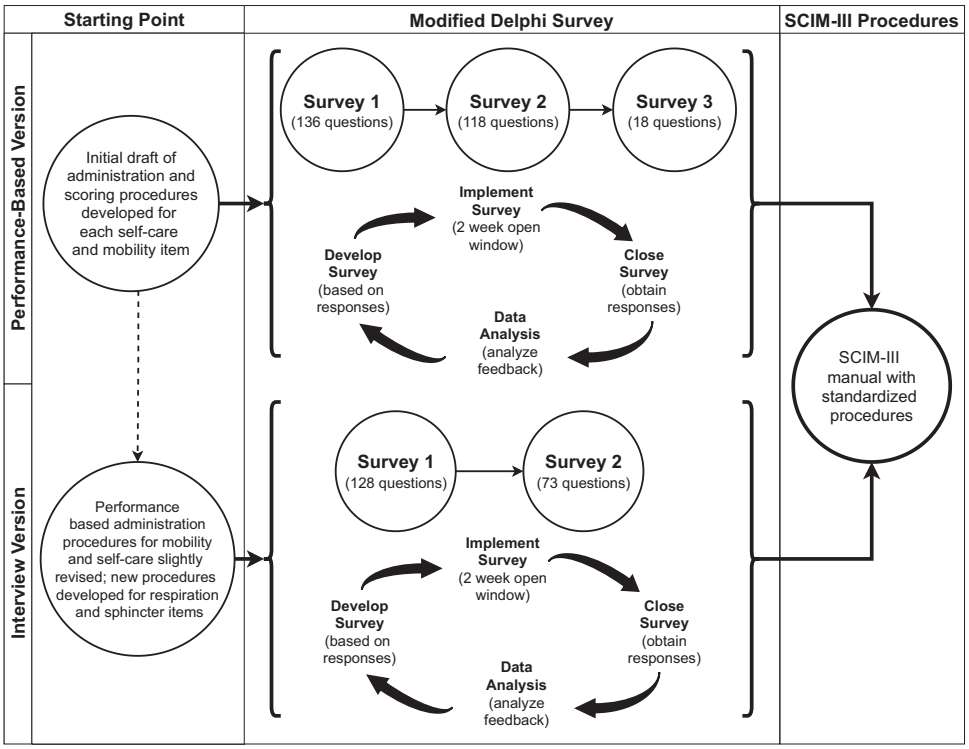

Fig. 1 SCIM-III development of standardized administration and scoring procedures for observation by performance and interview versions.

[16, 17]. Similarly, administration of the SCIM-III by interview may mitigate the amount of missing data in longitudinal studies.

When administrating the SCIM-III as a performance-based measure and as an interview, it is important to have guidelines that standardize administration and scoring of each item. Standardized guidelines would provide a safeguard for variability in item set-up, execution, scoring (performance-based administration), and in the structure and content of questions for each item (interview administration). Previous studies have recognized the need for better training in and standardization of administration and scoring of the SCIM-III [8, 15].

In response to the need for standardization in administration and scoring of the SCIM-III, and in recognition of the serious limitations in SCI clinical outcome measures, we engaged experts in SCI practice and research to create standardized procedures for the SCIM-III when administered and scored as a performance-based measure and standardized procedures for the SCIM-III when administered and scored by interview.

PURPOSE

The purpose of this study was to develop standardized administration and scoring procedures for the SCIM-III items when administered as a performance-based assessment and by interview. We sought to develop clearly written administration procedures that aligned with the intent of SCIM-III items and that were feasible to implement in practice environments and clinical trial settings. We also sought to develop procedures to aid in accurate scoring of SCIM-III items when administered as a performance-based measure and interview.

METHODS

The Modified Delphi Technique is an iterative process that uses a systematic progression of repeated rounds of surveys to allow for feedback and revisions as a way to build consensus among experts in areas and for purposes where knowledge is complex, incomplete, or poorly understood [18, 19]. The defining feature of the Modified Delphi is that the collated responses from previous questionnaires are integrated into each new questionnaire, and the experts being questioned are able to reconsider their previous judgments, revising and responding as appropriate [20]. In health care, the Modified Delphi Method has been used for a variety of purposes including the development of clinical practice guidelines. In this study, we used the Modified Delphi to engage SCI clinicians and researchers for the purposes of developing standardized administration and scoring procedures for the SCIM-III.

As shown in Fig. 1, a multi-stage iterative process [19, 21, 22] (Fig. 1) was used to create initial administration and scoring procedures for each SCIM-III item and to garner feedback on clarity of wording related to administration, scoring guidelines, and feasibility of administering SCIM-III performance-based items as specified. This iterative process was used until consensus was reached at 80% agreement among experts that the administration and scoring were clearly written, that the procedures aided in administration and scoring, and it was feasible to execute. Two independent Modified Delphi processes were implemented. The first Modified Delphi focused on developing guidelines for the SCIM-III for administration and scoring as a performance-based measure (March 2020–May 2020), and the second Modified Delphi process was undertaken to develop guidelines for the SCIM-III for administration by interview (February 2021–March 2021). The two-staged approach was used to mitigate confusion about mode of administration.

As shown in Fig. 1, an initial set of administration and scoring guidelines were established based on the SCIM-III literature and SCIM-III experience and were used as the basis for the first survey.

Survey Development

As a starting point, initial administration and scoring procedures for each SCIM-III mobility and self-care item were developed for performance-based standardization. These procedures were slightly revised for SCIM-III mobility and self-care items for interview, and initial procedures for respiration and sphincter were created (we did not include respiration or sphincter in the performance-based guidelines) (Fig. 1).

Specifically, for each SCIM-III mobility and self-care item, an explicit administration and scoring procedure was provided followed by four "yes/no" questions: 1. Are the (administration) instructions clear? 2. Could you replicate this in your clinic? 3. Are the scoring instructions clear? 4. Would you be able to determine a SCIM-III score? This resulted in a total of 60

questions for the first survey for the Modified Delphi Process for performance-based standardization (15 SCIM-III mobility\self-care items X 4 "yes/no questions per item). Only two "yes/no" questions were asked for each SCIM-III item for interview standardization: 1. Are the interview prompts clear? 2. Could you replicate this in your clinic? Thus, for the first survey for the Modified Delphi process for interview standardization, there were 38 questions (19 SCIM-III items X 2 "yes/no" questions per item). Survey respondents were asked to answer each yes/no question, provide a rationale for a response of "no", and suggest alternatives to the content or wording to improve clarity, consistency, and standardization.

Consistent with Modified Delphi Methodology, subsequent surveys contained revised procedure guidelines based on expert input and the same "yes/no" questions. Surveys were completed anonymously and remained open for 2 weeks for each iterative round, with an e-mail reminder after one week to boost response rate. A statement of consent was provided in the email invitation.

Survey respondents

Purposeful [23] and snowball [24] sampling were used to recruit survey participants. Purposeful sampling was done by e-mailing an invitation to content experts known to the principal investigator with a description of the study purpose and a link to the survey. Content experts were asked to forward the e-mail invitation to their colleagues who work with the SCI population (snowball sampling).

Data analysis

Summary and descriptive statistics were used to assess percent agreement of survey responses. For yes-no questions that had at least 80% response rate of "yes", procedures were considered clearly written and feasible to implement in usual practice and clinical trial environments. For yes-no questions that did not reach 80% response rate of "yes", procedures were revised based on respondents' feedback and exposed to subsequent survey rounds until 80% agreement was achieved. Once 80% agreement was obtained for each question for each SCIM-III item, administration and scoring guidelines for performance-based administration and for administration by interview were formatted and prepared for widespread dissemination.

This study was approved by the Institutional Review Board.

RESULTS

As shown in Table 1, the majority of respondents were physical and occupational therapists, had over 10 years of SCI experience, and used the SCIM-III, primarily in rehabilitation and research settings. The majority of respondents reported administering the SCIM-III using a combination of observing performance and interview.

Three iterative rounds of the performance-based survey were required to reach 80% agreement for each of the four questions for each SCIM-III self-care and mobility items (Table 2). Two iterative rounds of the interview survey were required to reach 80% agreement for all items (Table 3).

Modified Delphi process for standardization of procedures for administration and scoring by performance-based measure

Table 2 shows the response rate of "Yes" to each of the four questions asked for the mobility and self-care items in the Modified Delphi process for standardization for the performance-based measure. As shown in the first survey, the response of "yes" for the administration procedures for 6 items and for the scoring procedures for 8 items was obtained in at least 80% of respondents. For the remaining questions ($n = 36$) that did not reach 80% response of "yes", feedback from respondents (Table 4) was used to revise the administration and scoring procedures, and these items were exposed to expert opinion and feedback in survey. As shown in Table 2, with the exception of the two questions asked for scoring procedures for one item (Mobility in Bed and Action to Prevent Pressure Injuries), survey 2 obtained at least an 80% response rate of "yes" for every question for each item. Respondent feedback was used to revise the scoring procedures for the Mobility in Bed and Action to Prevent Pressure Injuries (Table 4), and the item was again exposed to expert opinion and feedback in a third survey.

Modified Delphi process for standardization of procedures for administration and scoring by interview

Table 3 shows the response rate of "Yes" for each of the two questions asked for each of the SCIM IIII items in the Modified Delphi process for standardization for interview. As shown in the first survey, the response rate of "yes" for the questions, "are the interview prompts clear?" and "could you replicate this in your clinic?", were at least 80% for 13 items and for 16 items, respectively. For the items that did not reach 80% response rate for either question, feedback from survey 1 (Table 4) was used to revise the administration and scoring procedures, which were exposed to expert review in the second survey. Of note, although the 80% threshold was reached for two items (use of toilet; transfer from wheelchair to toilet/tub), administration and scoring were revised based on survey 1 feedback (Table 4) and included in the second survey. As shown in Table 3, an 80% or greater response rate of yes was obtained for each SCIM-III item.

DISCUSSION

In this study, a Modified Delphi Technique was used to iteratively obtain feedback and establish consensus among experts on the standardization of administration and scoring procedures for the SCIM-III when administered as a performance-based measure and as an interview. Owing to the widespread use of the SCIM-III in practice and research and the concerns about the lack of available standardized procedures for administration and scoring, we established procedural guidelines that are clear, mitigate variability, accessible, and that are feasible to implement in practice and research settings.

The inherent lack of guidance of the existing SCIM-III leaves it open to interpretation and variability. Given the potential variability in performance-based outcome measures, it is important to establish fidelity in administration and scoring to assess whether or not an individual implements an evaluation as intended [25, 26]. In a clinical trial setting, when administering outcome measures with different administration and scoring procedures, fidelity is essential to prevent "rater drift and contamination" [27]. This is significant considering the reliability and validity of a measure are "dependent on the consistency with which it can be administered, scored, and interpreted" [28]. Moreover, because users of clinical outcome measures frequently do not read clinical outcome measures users' manuals [29], we created scoring flow diagrams that are intuitive and allow for low burden and accurate scoring.

While we did not collect information on geographical location of the respondents, purposeful sampling included distribution of the email invitation to international experts. One possible limitation for not reaching the 80% threshold in the first surveys is due to regional variations in English lexicon. Certain terms such as "electric aide", "durable medical equipment" and "tick" seemed to be relevant for some regions and not for others. Medical jargon and use of abbreviations also likely contributed to responses of "no" for some items. Specifically, the definitions for durable medical equipment, adaptive devices, specific setting, total partial assistance, and partial assistance were created, as respondents identified these terms as being too complex or "technical" for use in SCIM-III interview. As such, the procedure manuals include an abbreviation list and glossary of terms with definitions to decrease areas of potential confusion and to make the measure easier to use for those with limited experience with SCIM-III. Gender-specific pronouns were removed to be more inclusive of all gender identities.

Table 1. Sample of survey respondents for performance-based and interview surveys.

	Performance-based Modified Delphi			Interview Modified Delphi	
	Round 1	**Round 2**	**Round 3**	**Round 1**	**Round 2**
	***N* = 51**	***N* = 48**	***N* = 64**	***N* = 27**	***N* = 12**
	Respondents N (%)			***Respondents N (%)***	
Discipline[a]					
Physical Therapist	23 (45.1)	22 (45.8)	33 (51.6)	11 (45.83)	6 (50)
Occupational Therapist	15 (29.4)	14 (29.2)	17 (26.6)	6 (25)	3 (25)
Physician (e.g., MD, DO)	6 (11.8)	7 (14.6)	5 (7.8)	5 (16.67)	1 (8.3)
Nurse	2 (3.9)	1 (2.1)	3 (4.7)	1 (4.2)	1 (8.3)
Physician Assistant	0 (0)	0 (0)	0 (0)	1 (4.2)	0 (0)
Psychologist	1 (2)	0 (0)	1 (1.6)	0 (0)	0 (0)
Other	4 (7.8)	4 (8.3)	5 (7.8)	1 (4.2)	1 (8.3)
Practice Setting (all that apply)					
Inpatient rehabilitation	42 (44.2)	38 (46.4)	44 (42.7)	20 (48.8)	8 (36.4)
Outpatient rehabilitation	24 (25.3)	16 (19.5)	29 (28.2)	5 (11.2)	4 (18.2)
Research environment	18 (19)	16 (19.5)	19 (18.5)	8 (19.5)	6 (27.3)
Acute care	9 (9.5)	11 (13.5)	6 (5.8)	6 (14.6)	3 (13.7)
Other – school, fitness/exercise gym	2 (2.2)	1 (1.2)	5 (4.8)	2 (4.8)	1 (4.6)
Patient Age Group (years) (all that apply)					
Pediatrics (birth to 18 years)	25 (24.3)	18 (20)	25 (20.3)	10 (23.8)	6 (31.6)
Adults (19 years or older)	78 (75.7)	72 (80)	98 (79.7)	32 (76.2)	13 (68.4)
Experience in SCI field (years)[b]					
<1 year	2 (4)	3 (6.5)	1 (1.6)	0 (0)	2 (16.7)
1–3 years	5 (10)	7 (15.2)	7 (11.3)	4 (17.4)	2 (16.7)
4–10 years	8 (16)	9 (19.6)	16 (25.8)	3 (13)	2 (16.7)
11–20 years	18 (36)	12 (26.1)	20 (32.3)	5 (21.7)	1 (8.3)
21 years or more	17 (34)	15 (32.6)	18 (29)	11 (47.8)	5 (41.7)
Do you have experience using the SCIM-III?[c]					
Yes	36 (70.6)	31 (64.6)	48 (75)	14 (58.3)	9 (75)
No	15 (29.4)	17 (35.4)	16 (25)	10 (41.7)	3 (25)
[If yes] How often do you use the SCIM-III in practice?					
Very often					
Often	11 (31.4)	4 (13.3)	12 (25.5)	3 (21.4)	0 (0)
Somewhat often	3 (8.6)	9 (30)	10 (21.3)	3 (21.4)	3 (33.3)
Not very often	5 (14.3)	4 (13.3)	10 (21.3)	2 (14.3)	3 (33.3)
Never	14 (40)	11 (36.7)	15 (31.9)	3 (21.4)	3 (33.3)
	2 (5.7)	2 (6.7)	0 (0)	3 (21.4)	0 (0)
If you have experience with the SCIM-III, how do you administer it as a performance based measure (observe patient do activity), self-report (patient self-reports function through an interview or by completing on own), or a combination of both?[a]					
Combination of both	19 (55.9)	20 (26.7)	28 (60.9)	4 (33.3)	6 (66.7)
Self-reports/Interview	10 (29.4)	8 (26.7)	15 (32.6)	3 (25)	1 (11.1)
Performance-based measure	5 (14.7)	2 (6.7)	3 (6.5)	5 (41.7)	2 (22.2)
When administering the SCIM-III as a performance-based measure, do you think the patient must always attempt the task, even if they are incapable of performing it?[d]					
No, score zero	N/A	27 (77.1)	26 (86.7)		
Yes, always attempt	N/A	7 (20)	3 (10)		
Other	N/A	1 (2.86)	1 (3.3)		

[a]Interview survey: For round 1, 2 participants did not respond (Discipline, $N = 25$; Administering SCIM-III, $N = 12$).
[b]Performance-based survey: for round 3, one participant did not respond to this question ($N = 63$);
Interview survey: For round 1, 4 participants did not respond to this question ($N = 23$).
[c]Interview survey: For round 1, 3 participants did not respond to this question ($N = 24$).
[d]Performance-based survey: added round 2.

Table 2. Response rate of "yes" to each of the four questions for mobility and self-care items for standardization when administered and scored as a performance-based measure.

	Administration procedures				**Scoring procedures**					
	Total respondents (Percent response "yes")									
Question	*Are the instructions clear?*		*Could you replicate this in your clinic?*		*Are the scoring instructions clear?*			*Would you be able to conclude your score?*		
Survey	1	2	1	2	1	2	3	1	2	3
	N (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)
Feeding	**45 (71.1)**	41 (92.7)	45 (95.6)	41 (92.7)	41 (82.9)	39 (92.7)	–	41 (85.4)	39 (92.7)	–
Bathing (UB/LB)	**38 (73.7)**	38 (94.7)	**38 (79)**	38 (86.8)	37 (94.6)	38 (92.1)	–	38 (92.1)	38 (92.1)	–
					36 (80.6)	38 (84.2)	–	36 (86.1)	38 (92.1)	–
Dressing (UB/LB)	**37 (73)**	38 (89.5)	37(81.1)	38 (97.4)	36 (81.6)	38 (81.6)	–	36 (94.4)	38 (92.1)	–
					36 (80.6)	38 (81.6)	–	36 (86.1)	38 (89.5)	–
Grooming	36 (80.6)	–	36 (94.4)	–	**36 (66.7)**	38 (86.8)	**–**	**36 (75)**	38 (89.5)	–
Mobility in Bed & Action to Prevent Pressure Injuries	**36 (77.8)**	38 (89.5)	36 (94.4)	37 (97.3)	35 (88.6)	**38 (63.2)**	58 (87.9)	35 (91.4)	**38 (71.1)**	58 (96.6)
Transfers: Bed to Wheelchair	35 (80)	–	35 (94.3)	–	**34 (79.4)**	38 (92.1)	**–**	**34 (76.5)**	38 (97.4)	–
Transfers: Wheelchair-Toilet-Tub	35 (82.9)	38 (86.8)	**34 (76.5)**	38 (94.7)	**34 (79.4)**	37 (81.1)	**–**	**35 (74.3)**	36 (88.9)	–
Mobility: Indoors, Outdoors, & on Uneven Surfaces	35 (88.6)	–	35 (91.4)	–	**35 (57.1)**	37 (86.5)	**–**	**35 (74.3)**	37 (89.2)	–
Stair Management	35 (82.9)	–	34 (94.1)	–	35 (100)	37 (97.3)	–	35 (100)	37 (94.6)	–
Transfers: Wheelchair-Car	35 (88.6)	–	34 (91.2)	–	**35 (75.3)**	37 (91.9)	**–**	35 (82.9)	37 (94.6)	–
Transfers: Wheelchair-Ground	34 (88.2)	–	33 (93.9)	–	35 (91.4)	37 (97.3)	–	35 (88.6)	37 (97.3)	–

Values in bold indicate response rate of "Yes" below the 80% threshold, indicating need for revision and inclusion in subsequent survey; a "–" (dash) indicates that item was not reviewed this round due to obtaining at least 80% response rate of "Yes" in previous survey. Note: the number of respondents (*N*) for questions vary, as not all respondents completed all items.
N number of respondents, *UB* upper body, *LB* lower body.

Table 3. Response rate of "Yes" to each of the two questions for all items for standardization when administered as an interview measure.

Administration procedures				
	Total respondents (Percent response "Yes")			
Question	*Are the interview*		*Could you replicate this in your clinic?*	
	prompts clear?			
Survey	1	2	1	2
	N (%)	*N* (%)	*N* (%)	*N* (%)
Feeding	21 (95.5)	–	20 (90.9)	–
UB Bathing	**15 (68.2)**	12 (100)	20 (90.9)	11 (100)
LB Bathing	**13 (61.9)**	12 (100)	18 (85.7)	10 (100)
UB Dressing	**15 (71.4)**	10 (83.3)	18 (85.7)	12 (100)
LB Dressing	16 (80)	10 (83.3)	18 (85.7)	12 (100)
Grooming	18 (85.7)	–	–	–
Respiration	**16 (76.2)**	11 (91.7)	18 (85.7)	11 (91.7)
Sphincter Management: Bladder	**15 (75)**	11 (91.7)	**15 (75)**	11 (91.7)
Sphincter Management: Bowel	**15 (75)**	10 (83.3)	**15 (75)**	11 (91.7)
Use of Toilet	17 (85)	11 (91.7)	17 (85)	10 (83.3)
Mobility in Bed & Action to Prevent Pressure Injuries	17 (85)	–	17 (85)	–
Transfers: Bed-Wheelchair	18 (90)	–	18 (90)	–
Transfers: Wheelchair-Toilet-Tub	18 (90)	12 (100)	18 (90)	11 (91.7)
Mobility Indoors	17 (85)	–	17 (85)	–
Mobility, Moderate Distances	17 (85)	–	17 (85)	–
Mobility Outdoors	17 (85)	–	17 (85)	–
Stair Management	18 (94.7)	–	18 (94.7)	–
Transfers: Wheelchair-Car	18 (90)	–	16 (84.2)	–
Transfers: Wheelchair-Ground	18 (90)	–	18 (90)	–

Values in bold indicate response rate of "Yes" below the 80% threshold, indicating need for revision and inclusion in subsequent survey; a "–" (dash) indicates that item was not reviewed this round due to obtaining at least 80% response rate of "Yes" in previous survey. Note: the number of respondents (*N*) for questions vary, as not all respondents completed all items.
N number of respondents, *UB* upper body, *LB* lower body.

Survey responses and feedback led to revision of administration and scoring procedures for the SCIM-III items, resulting in standardized and clearly written administration and scoring procedures for the SCIM-III. The methodology did not alter any established SCIM-III item or scoring criteria [1, 3, 5, 6]. Our intent was to maintain the SCIM-III as it was developed and make no changes to items or scoring. Several respondents reported concerns with the inconsistent scoring scale; meaning sometimes scoring is 0-2 while for other items it is 0-8. It was important to state that the scoring is directly from the original SCIM-III, and our team did not feel it within the scope of our work to revise the scoring structure. Our work focused entirely on providing clear and specific instructions for administering and scoring SCIM-III items via performance-based measure and individual interview, with the goal to provide the field with a resource to assess endpoints of usual care and clinical trials.

Performance-based version

An introduction and general instructions for the survey and SCIM-III performance-based procedures were necessary to provide background information and decrease administrator burden when completing the survey and administering this version of the SCIM-III. Respondents of the performance-based survey were instructed to review both the written instructions and scoring flowsheets prior to answering survey questions since these were developed to be used in conjunction with each other for each SCIM-III item. In both versions, the starting prompt of each flowsheet was moved towards the top-left corner and the final scores, which were made rectangular and blue to differentiate from the rest of the flowsheet, were organized in numerical order and placed on either the right or bottom of the flowsheet.

Administration and scoring procedures were not developed for the respiration and sphincter management subscale of the SCIM-III. This portion was excluded in the performance-based procedures because of the inability to adequately demonstrate and assess respiratory elements with a performance measure. Additionally, having participants demonstrate their bowel and bladder program would be an undue burden. In contrast, the self-care and mobility portions were included because clinical trials are more focused on mobility and upper extremity function, which are more directly related to those sections. Moreover, in the rehabilitation setting, it is frequently the occupational and physical therapists who use the SCIM-III; therefore, if transformed into a performance measure, the self-care and mobility portions would be assessed by an occupational or physical therapist.

Initial administration procedures gave examples of items that could be used or tasks to perform. However, survey respondents reported that the initial examples were not specific enough and the expectation was still unclear. For example, the administration procedures for the dressing item included "raise arms". In round 2, it was revised to "touch the back of your head". This created clear expectations for both the participant and administrator. We further needed to develop a list of recommended materials to be used for each item, adding to the standardization of our manual. This list included items such as an 8 oz sandwich container, 12 oz plastic cup, and a reusable metal knife and fork.

Table 4. Problems identified by respondents in performance-based and interview surveys of Modified Delphi process.

Performance-Based Survey: Revisions from round 1 to round 2		
Problems identified in round 1	**Sample of verbatim feedback**	**Revisions made for round 2**
Abbreviations: There was difficulty understanding abbreviations and terms. Examples of abbreviations and terms that caused confusion are: assistance, partial assist (PA), durable medical equipment (DME), edge of mat (EOM), and adaptive devices (AD). This pertained to all items.	*Recommend writing out PA, AD near flow chart* *Do not use the term 'DME' – is this a commode? Shower chair? Can only guess* *DME abbreviation should not be used – other countries, it may be called differently. Same comment for AD*	To increase clarity and understanding, abbreviations were written out in the administration and scoring procedures. A glossary of terms was also created with examples of durable medical equipment, adaptive devices, specific settings, and assistance.
Required Materials: Required materials were not listed for each item. Respondents also requested more specifications for each item. This increased administrator burden because respondents were not sure what materials were needed or could be used for each item.	*It is not clear as to what some of the tasks require to be sure they are standardized* *More clarity about materials needed* *What constitutes a sandwich container? Plastic bag, Tupperware, etc.*	A comprehensive required materials list was included for round 2 at the beginning on the survey. Further, each SCIM-III item was updated to include the required materials for that item. For example, the required materials for "Feeding" are: sandwich container, utensils (knife and fork), water bottle, cup, and foam block.
Lack of Materials[a]: Not all clinics/labs will have all materials/equipment. There was a reported need for options with alternate materials and/or simulated environments.	*Equipment not present in research lab* *Would need tub bench and tub* *DME and shower is not available*	Alternative equipment and specifications for simulated environments were provided if the clinic/lab lacks equipment. For example, for "Transfers: bed to wheelchair", if a bed is not available, then a raised mat can be used. Administrator must document the surface used, and use the same surface each time the participant is tested.
"Mobility in Bed and Action to Prevent Pressure Injuries": The original SCIM uses the term "pressure sores"; however, that term is outdated.	*…pressure injury (not sore)…*	The term "pressure sore" was replaced with "pressure injury" to reflect present terminology.
"Feeding": Respondents did not like the use of plastic utensils. They preferred reusable utensils because they were sustainable and sturdier (safer).	*Why a plastic fork/knife? It seems that most people would eat with standard utensils…* *It is far more environmentally conscious to utilize standard, metal utensils* *Cutting…with plastic knife sometimes causes the knife/fork to break…this might not be safe*	The utensils were changed from plastic utensils (round 1) to metal utensils (rounds 2 onward).
"Feeding": Respondents did not like the use of putty. There were sanitary concerns with the inability to properly clean and disinfect putty.	*Is the intent to re-use the theraputty? Once touched, theraputty cannot be sanitized. That seems like a large overhead cost (or waste)* *What kind of putty? Cutting the putty…may not be safe*	During round 2, respondents were asked what could replace putty as an alternative to a sandwich. Respondents suggested using foam, a sponge, or playdoh. The team conferred and decided to use foam.
"Dressing": Administration procedures previously stated to 'raise arms overhead', but that was considered too vague.	*What constitutes "reaching overhead"?* *A specific landmark that must be surpassed or reached* *UB dressing instructions from wheelchair should be more specific*	The instruction to "raise arms" was changed to "touch back of head" to increase specificity and clarity.
"Transfers: Wheelchair-Toilet-Tub": Settings/equipment were set to be level with each other (e.g., tub bench is set level to height of wheelchair); however, that is not always feasible.	*"Tub bench is set level to height of wheelchair" is not always achievable* *Not sure if always possible to adjust height of tub bench* *No height-adjustable tub bench available*	The clause, "or as close to the height of the wheelchair as possible", was added to the administration procedures.
"Mobility: Indoors, Outdoors, and on Uneven Surfaces": There was a lot confusion regarding the wording and organization for the 3 scoring procedures related to this item ("Mobility Indoors 10 feet", "Mobility Indoors 100 feet", "Mobility Outdoors").	*The flow from "can you walk 10' on smooth surfaces" does not make sense*	The scoring procedure was revised to address all areas of confusion. It was updated to be shorter and easier to follow.
	The first diamond to the right of the green diamond is unclear. The question…asks about walking, then YES answer continues about needing supervision in a wheelchair	One of the big areas of confusion was a connection that contradicted itself by asking about walking then, if someone answered yes", asking about supervision in a wheelchair. A box was added to the flowchart, "require a wheelchair", to address the error in the flowchart.
	The graph and the scoring system is confusing *Too long…too hard to follow….*	
"Transfers: Wheelchair-Car": The score of 0 for needs total assistance was missing from the scoring flowsheet, making the scoring procedures difficult to follow.	*I do not see an option for 0 score – if need total help for all* *There isn't a way to get a score of "0" on the algorithm*	A score of 0 (needs total assistance) was added to the scoring flowchart.
Scoring: For some of the mobility items ("Transfers: Bed to Wheelchair", "Transfers: Wheelchair-Toilet-Tub", & "Transfers: Wheelchair-Car"), the scoring is dependent on a number of activities; however, that was not made clear.	*Is the scoring based on average response to all questions or is there a score for each diamond* *Is the score of 1 for any of the activities?* *Need to clarify how to score if independent with some aspects of transfer, but not others*	The scoring procedures were revised to increase clarity for respondent. A disclaimer was added that scoring is altogether, not separate per task. For example, if one task requires partial assistance, then the participant receives a score of 1. A score of 2 means the participant is independent for all activities.
Setting: The setting for a few items ("Bathing", "Dressing", "Transfers: Bed to Wheelchair", "Transfers: Wheelchair-Toilet-Tub", & "Transfers: Wheelchair-Car"), were unclear. Survey respondents reported confusion and requested increased specificity (e.g., real vs. simulated activities, dry vs. wet environments).	*Unclear if this is a simulation or not…wet shower conditions have extra amount of difficulties* *Can we have a mock setup using a bench?* *Unclear about the set-up and use of equipment for transfers*	The instructions were revised to specific that a dry environment (vs. wet) for "Bathing" items and real settings should be used. The instructions further indicated that a simulated setting is acceptable if necessary, as well as specifications for if a simulated setting is used. This was included in the general instructions created for round 2 and within the administration procedures for each relevant item.
Use of Adaptive Devices: Many respondents inquired if adaptive devices were permitted. This needed to be clarified for future iterations.	*It is unclear if adapted equipment can be used for these tasks.* *Can the participant utilize adaptive equipment?* *It would be helpful to clarify whether adaptive equipment can be used…*	General instructions specified that "adaptive devices (AD), specific settings (SS), and durable medical equipment (DME) are allowed". Additionally, the required materials list was updated to include "adaptive devices typically used in feeding, bathing, grooming, and dressing tasks".

Table 4. continued

Performance-Based Survey: Revisions from round 1 to round 2		
Gender-specific Pronouns (he/him, she/her): Gender-specific pronouns can be exclusionary, and do not capture all persons.	*The terms his/her are not inclusive of all genders.* *Avoid he/she usage*	The administration and scoring procedures were revised with gender neutral terminology. Gendered pronouns like he/she were eliminated.
Red/Green Lines: The scoring procedures connected each prompt with blue lines for round 1. There was no clear visual marking or differentiation for "yes" or "no" branches.	There were no specific comments requesting this change. This change was made to increase clarity and decrease administrator burden when navigating the scoring procedures.	The scoring procedures were updated to include red/green lines leading to "yes" and "no".
Performance-Based Survey: Revisions from round 2 to round 3		
Problems identified in round 2	**Sample of verbatim feedback**	**Revisions made for round 3**
Formatting: The starting, green diamond leads into a single blue diamond, which caused confusion regarding the flow of the scoring procedures. Moreover, the flowchart was vertical, which suggested each step led into the next when each step is weighed equally. The overall formatting was not intuitive or easy to follow.	*It is unclear how to follow the first prompt and not intuitive how to tabulate* *There must be a better way to indicate all 4 assessment points must be made independently* *The green arrows connecting the 4 elements (tasks) are not how you have used the green arrows elsewhere*	To make the flowchart more intuitive, it was flipped horizontally. Rather than having green arrows connect each task, the green diamond had blue arrows branching into each of the four tasks. The tasks were placed next to one another, which clarified that they must be made independently. These changes allowed respondents to understand that the score is calculated by the sum of all "yes" responses.
Repeat: There were two tasks, or blue diamonds, that said "able to turn lower body on the mat".	*Two "turn lower body"* *Should one be upper body and other be lower body?*	The first "able to turn lower body on the mat" blue diamond was revised to "able to turn upper boy on the mat".
Time: There was an inconsistency between the administration and scoring procedures regarding the time to press up to lockout. The administration procedures indicated 30 s.	*The instructions on the first section said to hold the press up x 30 s, the diamond on the flow chart says 10 s* *Instructions say 30 s for press up, and flowchart says 10. Which is accurate?*	The administration procedure was updated to match the scoring flowchart. The 30 s was revised to 10 s.
Interview Survey: Revisions from round 1 to round 2		
Problems identified in round 1	**Sample of verbatim feedback**	**Revisions made for round 2**
Scoring difficulty: Among respondents, there was confusion around what certain scores implied or why there could be multiple paths to the same score.	*There are 2 scores of 2.*	General survey instructions were added specifying that "the original SCIM-III does not use a consistent score scale for each item. Therefore, the highest score for one item might be a 2, but for another item, the highest score might be an 8. There may be multiple routes through the flowsheets that will result in the same score."
	There are 2 scores of 1.	Given the scope of our work, it was not appropriate to make revisions to the scoring structure.
Relevance to practice area (including pediatric settings): Some respondents expressed certain items were not relevant to their discipline, domain of practice, or practice setting.	*We don't use the SCIM in the clinic…only for research purposes where I think this would be of valuable use.* *Tasks in prompts not applicable in pediatric hospital (e.g. cut with a knife)* *Not clear enough to describe child's function or measure progress* *Younger children will not be able to respond to these prompts* *Language not appropriate for younger children*	Although feedback within this theme is not tied to any explicit changes, certain language in the flowsheets was tailored so the interview could be more accessible to SCI patient populations across the lifespan. Certain medical terms were changed to reflect more colloquial language. For example, "gastrostomy" was changed to "feeding tube" as it was considered to be more familiar to patients.
Abbreviations: There was difficulty understanding abbreviations and terms. There were also concerns that patients may not understand certain concepts. Examples of abbreviations and terms that caused confusion are: adaptive devices (AD), specific settings (SS), total vs. partial assistance, intermittent assisted ventilation (IAV), tracheal/tracheostomy tube (TT)	*What is "AD" shouldn't you spell it out?* *Is assist physical? What is partial?* *Special setting is unclear* *Would skip the abbreviations; never heard of IAV or TT - just say trach or tracheostomy* *Need to distinguish total from partial assistance.* *What is an electric aid?* *Long and ring sitting are too technical*	To increase clarity and understanding, abbreviations were written out in flowsheets or a key was provided. A glossary of terms was also created. General instructions were added specifying that, "to stay as true to the original SCIM-III as possible, scoring, measurements (i.e., RUV values), and core tasks were not altered…Some measurements (especially for the Respiratory and Sphincter Management items) were preset by the original SCIM-III (i.e., RUV values/units). In order to adhere to the original SCIM-III, these are consistent with the original form."
Formatting: The positioning of the initial prompt was slightly different on the flowsheets considering varying complexity and number of prompts on each. The initial prompt was yellow and subsequent prompts were green, which was confusing to some respondents. All prompts were diamond shaped. The colors and shapes for prompts did not convey clear sequencing. Scoring boxes were not organized in one designated area of the flowsheet. The overall formatting proved confusing to participants and was not intuitive.	*I like the other flow diagrams because the first question is at top*	A sample flowsheet was provided at beginning of survey with instructions on how to follow prompts.
	Asking yes or no to 3 different questions is a bit daunting. I can't even answer one question no less 2 or 3!!	Initial prompt was moved to be positioned at the top of each flowsheet.
	I don't understand flow of questions *Start with green and if yes go to the yellow* *Maybe better if the scoring blue rectangles are arranged from lowest 0 at top to highest 4 at bottom*	Color and shape of initial and subsequent prompt boxes were changed so that new sequence was green circle (initial prompt), which was more intuitive for most people, to yellow diamond (subsequent prompts) then blue rectangle (scoring).
"UB Bathing": The initial prompt read "Are you able to stand in the shower," which focused on positioning of the participant during bathing rather than the adaptive devices (AD) and/or specific setting (SS) that is used to facilitate bathing. Use of AD and/or SS is the basis for scoring on the original SCIM-III.	*For the left green prompt - maybe add independently*	To create a clearer path through the flowsheet, the initial prompt was changed to read: "Do you use adaptive devices and/or a specific setting to shower?"
	"Unable to stand" prompt: I may say no because I need assistance	Also, scoring criteria were explicitly written to state if bathing required AD, SS, both, or none.
	I don't fully understand the last No: Total Assistance - I find the end choices confusing.	

Table 4. continued

Interview Survey: Revisions from round 1 to round 2		
Problems identified in round 1	**Sample of verbatim feedback**	**Revisions made for round 2**
"LB Bathing": The initial prompt, like in the UB Bathing flowsheet, focused on positioning of the participant and not necessarily on assistance level, which is the basis for scoring the item on the original SCIM-III.	*Again, I find the end splitting of choices confusing and think it should be split differently.*	Altered flowsheet in similar way to UB Bathing to decrease number of prompts asked and clarify interpretation of scoring.
	Not sure what the right two lowest boxes mean	
"UB Dressing": The initial prompt did not include orthoses as part of UB Dressing and subsequent prompt was too complex to differentiate between use of AD or SS for dressing.	*First green prompt: how do you know if special setting needed?*	Simplified prompts to distinguish whether AD and/or SS is used in any part of dressing. Added "by yourself" to initial prompt. Expanded scoring details to include type of supports required during dressing, whether it is physical assistance, AD, and/or SS.
	I think independently should be added	
"LB Dressing": One prompt referred to a "preferred setting" rather than the "specific setting" that the original SCIM-III uses.	*Second Green prompt: In preferred setting…, I don't see how you get to scores 0-2 without additional questions*	Edits made similar to UB Dressing, "preferred setting" was no longer used and replaced with "specific setting", which is consistent with original SCIM-III language.
	In the green middle prompt, there is the wording "by yourself." I think this should be added throughout.	Specifically asked about use of AD, and/or SS with zippers, buttons, hooks, laces since round 1 flowsheet only asked about assistance.
"Sphincter Management: Bladder": Round 1 flowsheet was arranged in a vertical orientation where the initial prompt was in line with one of the subsequent prompts. The arrangement of the flowsheet did not make progression through the prompts intuitive. Specific measurements from the SCIM-III were included with no explanation of their origin.	*What is the basis for RUV < or > 100? And is it cc or ml?*	Flowsheet was changed to a horizontal orientation which made flowsheet more intuitive.
	Original form is confusing. Can get up to 11 points with intermittent cath and residual > 100 cc	Noted that values such as RUV and the corresponding units were determined by the original SCIM-III and were not altered when developing flowsheets.
	Can't follow flowchart	
	Patients would not know their RUV	
"Sphincter Management: Bowel": Initial prompt read "Do you have regular bowel movements?" Definition of what is considered "regular timing" or "rare accidents" by the original SCIM-III was not mentioned until subsequent prompts.	*So don't you care if someone without a regular bowel program has accidents. I don't have a bowel program and don't have bowel accidents but for people with SCI I would like to know. What about use of diapers and incontinence briefs for both bowel and bladder?*	Changed initial prompt to read: "Is the timing of your bowel movements regular? (Irregular = less than once in 3 days)" and subsequent prompt defined a "rare accident" as "less than 2x/month" to ease administrator burden and clarify concepts for respondents.
	Twice a month not a reasonable timeframe for inpatient rehab; what is definition of "regular"?	Scoring criteria was expanded to define inclusion/exclusion of "regular timing" and/or "rare accidents".
"Use of Toilet": Reflected medical jargon of original SCIM-III with use of the word "perineal" which may not be a familiar term to respondents.	*Perineal is too advanced of a term*	Added "cleaning yourself" to prompt about "perineal hygiene"
"Transfers: wheelchair-toilet-tub": Initial prompt included "complete the transfer" as one of the steps to reflect the terms used in the original SCIM-III.	*What is meant by "complete the transfer" and why is that #3 (before all parts of transfer are completed)?*	In initial prompt, "complete the transfer" was replaced with "transfer (move to target surface)" to clarify action necessary for completion of transfer.

Respondents' questions and feedback (column 2) were incorporated into the administration and scoring procedures (column 3) and exposed to a second survey.

[a]*To better address this concern, a question was added to round 2 asking if respondents would be interested in a SCIM-III kit with all materials except durable medical equipment and adaptive devices (89.7% reported "yes").*

When standardizing administration procedures for dressing and bathing items, only a few specific tasks were selected to reduce time and participant and administrator burden. For example, for bathing, two body parts were selected to mimic upper body (non-dominant arm and low back) and lower body bathing (perineal area and dominant lower leg). The parts of the task that were selected were ones that incorporated most of the task or the more challenging aspects of the task. It was necessary to limit the task in this manner for the sake of time and clinical usability.

Lastly, modifications were made to all scoring procedures to increase flow and clarity. Although many of the scoring procedures obtained above 80% agreement in round 1, it was determined that it would be beneficial to apply the changes made to those with below 80% agreement to all flowsheets. The flowsheets were simplified and made as consistent as possible to ease administrator burden. They were then re-submitted in round 2 for review. After round 2, only 1 item required further revisions. This item was revised per the respondent's comments and submitted into a third and final round.

Interview version

An introduction was provided at the start of the round 2 survey to address some of the feedback provided in round 1. Participants provided comments that aligned with many of the limitations of the original SCIM-III such as the vagueness of core tasks, use of medical jargon and abbreviations, and the overall lack of accessibility of the measure to the patient population.

The original SCIM-III is not inclusive of all scenarios that individuals with SCI may participate in. To clarify the intent to retain fidelity to the original SCIM-III in terms of content, a general statement was made that: "Not every case/scenario may be captured in these flowsheets… Despite being an appropriate consideration for persons with SCI, some conditions and situations were not represented in the interview flowsheets in order to adhere to the original SCIM-III guidelines." One respondent emphasized that a person could perform a pressure relief in a wheelchair in more ways than just "doing push-ups in wheelchair," which is the only consideration in the original SCIM-III scoring criteria. These comments were reasoned to be issues and limitations with the original SCIM-III form and not necessarily with the interview flowsheets.

In developing the interview version, participants in round 1 provided feedback that the flowsheets were not intuitively organized and therefore the decision was made to alter the formatting and change the colors and shapes shown on the flowsheets. The shape of the initial prompt was changed to a circle and moved to the top of each flowsheet, making it apparent for administrators where to begin the sequence of prompts. The shape change also set the initial prompt apart from the subsequent diamond-shaped prompts. A sample flowsheet with

these changes was provided in the beginning of the round 2 survey with instructions on how to navigate the flowsheets.

Instead of focusing on the positioning of the individual during the bathing and dressing items, the prompts in the flowsheets were altered to be more consistent with the scoring guidelines which emphasize level of assistance provided, use of adaptive devices, and/or the use of specific settings to facilitate completion. In many cases, small modifications to prompts resulted in less prompts on the flowsheets overall, reducing the complexity of scoring and increasing ease of use of the interview format compared to the original SCIM-III. As indicated by 80% agreement on all the SCIM-III flowsheets in the round 2 survey, structural changes allowed for a conciseness of language in subsequent prompts and a clearer path through the flowsheets, decreasing the burden of administration and scoring. Standardizing the interview prompts while retaining fidelity to the original SCIM-III form makes the SCIM-III more accessible not only to those administering the assessment, but also to the individuals.

Limitations

A limitation in this study is our sampling methods. By utilizing purposeful and snowball sampling with health professionals and researchers known by the principal investigator, we may have limited the experience, practice settings, and familiarity with up-to-date practice recommendations of the participant pool. Moreover, the content experts selected are heavily involved in the SCI community, actively participate in SCI associations and their annual meetings, and engage in research and other educational opportunities. It is important to note that there are many health professionals and researchers who work with SCI populations but may not be included because they are not known to our PI. Moreover, because responses were not forced, not all survey participants responded to every question.

Dropout was evident in the interview survey between round 1 and round 2 where the response rate decreased by 15 total participants. Round 2 of the interview survey had fewer questions ($N = 26$) (9 SCIM-III items X 2 "yes/no" questions per item; 5 demographic questions; 3 SCIM-III-related questions) than round 1 ($N = 38$). There are a few factors that may have caused dropout. One possible reason is over-sampling. Despite the year gap between surveys, by using the same participant pool, individuals may have felt burdened by completing multiple surveys. Another reason may have been technical issues with accessing the survey link.

Future directions

In the performance-based survey, the questions regarding respondents' level of experience with the SCIM-III revealed that while majority had experience with the SCIM-III (65–75%), about 6.5–14.7% used the SCIM-III as a performance measure alone. Most respondents reported using the SCIM-III as a combination of both the performance and interview/self-report (56–67%). In the interview survey, 58.3–75.0% of participants were familiar with the SCIM-III. Of those individuals, 11.1–25.0% reported they use the SCIM-III as interview in their practice and 22.2–41.7% used the SCIM-III as performance measure alone. Approximately 33.3–66.7% reported to using the SCIM-III as a combination of a performance measure and interview/self-report. Based on these findings, it may be beneficial to compare the two distinct measures developed here to the original SCIM-III for accuracy and reliability.

CONCLUSION

Standardized administration and scoring procedures for both the performance-based and interview formats of the SCIM-III have been developed, and consensus on their clarity and feasibility established, and may be a useful resource for clinical trial programs and practice environments. They are available for download at no cost from the Jefferson College of Rehabilitation Sciences Center for Outcomes and Measurement website (https://www.jefferson.edu/university/rehabilitation-sciences/departments/outcomes-measurement/measures-assessments/spinal-cord-independence-measure-version-iii-administration-and-scoring-guidelines.html).

DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article [and its supplementary information files]. As stated in our conclusions, the completed guidelines resulting from this study are available on the Center for Outcomes and Measurement website, https://www.jefferson.edu/academics/colleges-schools-institutes/rehabilitation-sciences/departments/outcomes-measurement/measures-assessments/spinal-corde-independence-measure-v3.html.

REFERENCES

1. Catz A, Itzkovich M, Steinberg F, Philo O, Ring H, Ronen J, et al. The Catz-Itzkovich SCIM: a revised version of the spinal cord independence measure. Disabil Rehabil. 2001;23:263–8.
2. De Almeida C, Coelho JN, Riberto M. Applicability, validation and reproducibility of the spinal cord independence measure version III (SCIM III) in patients with non-traumatic spinal cord lesions. Disabil Rehabil. 2016;38:2229–34.
3. Catz A, Itzkovich M, Agranov E, Ring H, Tamir A. SCIM-spinal cord independence measure: a new disability scale for patients with spinal cord lesions. Spinal Cord. 1997;35:850–6.
4. Ackerman P, Morrison SA, McDowell S, Vazquez L. Using the spinal cord independence measure III to measure functional recovery in a post-acute spinal cord injury program. Spinal Cord. 2010;48:380–7.
5. Catz A, Itzkovich M, Tesio L, Biering-Sorenson F, Weeks C, Laramee MT, et al. A multi-center international study on the spinal cord independence measure, version III: Rasch psychometric validation. Spinal Cord. 2007;45:275–91.
6. Itzkovich M, Gelernter I, Biering-Sorenson F, Weeks C, Laramee MT, Craven BC, et al. The spinal cord independence measure (SCIM) version III: reliability and validity in a multi-center international study. DisabilRehabil. 2007;29:1926–33.
7. Bluvshtein V, Front L, Itzkovich M, Aidinoff E, Gelernter I, Hart J, et al. SCIM III is reliable and valid in a separate analysis for traumatic spinal cord lesions. Spinal Cord. 2011;49:292–6.
8. Anderson KD, Acuff ME, Arp BG, Backus D, Chun S, Fisher K, et al. United States (US) multi-center study to assess the validity and reliability of the spinal cord independence measure (SCIM III). Spinal Cord. 2011;49:88–885.
9. Scivolette G, Tamburella F, Laurenza L, Molinari M. The spinal cord independence measure: how much change is clinically significant for spinal cord injury subjects. DisabilRehabil. 2013;35:1808–13.
10. Aidinoff E, Front L, Itzkovich M, Bluvshtein V, Gelernter I, Hart J, et al. Expected spinal cord independence measure, third version, scores for various neurological levels after complete spinal cord lesions. Spinal Cord. 2011;49:893–6.
11. Fekete C, Eriks-Hoogland I, Baumberger M, Catz A, Itzkovich M, Luthi H, et al. Development and validation of a self-report version of the spinal independence measure (SCIM III). Spinal Cord. 2013;51:40–47.
12. Reed R, Mehra M, Kirshblum S, Maier D, Lammertse D, Blight A, et al. Spinal cord ability ruler: an interval scale to measure volitional performance after spinal cord injury. Spinal Cord. 2017;55:730–38.
13. Anderson K, Aito S, Atkins M, Biering-Sorenson F, Charlifue S, Curt A, et al. Functional recovery measures for spinal cord injury: an evidence based review for clinical practice and research. J Spinal Cord Med. 2008;31:133–44.
14. Macdermid JC, Law M, Michlovitz SL. Outcome measurement in evidence based rehabilitation. In: Law M, Macdermid J, editors. Evidence based rehabilitation: a guide to practice (3rd edition). Slack Incorporated; 2014. p. 65–104.
15. Saberi H, Vosoughi F, Derakhshanrad N, Yekaninejad M, Khan ZH, Kohan AH, et al. Development of Persian version of the spinal cord independence measure III assessed by interview: a psychometric study. Spinal Cord. 2018;56:980–6.
16. Itzkovich M, Shefler H, Front L, Gur-Pollack R, Elkayam K, Bluvshtein V, et al. SCIM III (Spinal Cord Independence Measure version III): reliability of assessment by interview and comparison with assessment by observation. Spinal Cord. 2017;56:46–51.
17. Itzkovich M, Tamir A, Philo O, Steinberg F, Ronen J, Sprasser R, et al. Reliability of the Catz-Itzkovich spinal cord independence measure assessment by interview and comparison with observation. Am J Phys Med Rehabil. 2003;82:267–72.
18. Hasson F, Keeney S, Mckenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32:1008–15.

19. Meshkat B, Cowman S, Gethin G, Ryan K, Wiley M, Brick A, et al. Using an e-Delphi technique in achieving consensus across disciplines for developing best practice in a day surgery in Ireland. J Hosp Adm. 2014;3:1–8.
20. Niederberger M, Spranger J. Delphi technique in health sciences: a map. Front Public Health. 2020; https://doi.org/10.3389/fpubh.2020.00457.
21. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess. 1998;2:1–88.
22. Green B, Jones M, Hughes D, Willimas A. Applying the Delphi technique in a study of GP's information requirements. Health Soc Care Community. 1999;7:198–205.
23. Emmel, N. Purposeful sampling. In: Sampling and choosing cases in qualitative research: a realist approach. SAGE Publications Ltd; 2013. p. 33–44.
24. Morgan, D. Snowball sampling. In: Given L, editor. The SAGE Encyclopedia of Qualitative Research Methods. Thousand Oaks: SAGE Publications Inc.; 2008. p. 816–7.
25. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. Am J Eval. 2003;24:315–40.
26. Walton H, Spector A, Williamson M, Tombar I, Michie S. Developing quality fidelity and engagement measures for complex health interventions. Br J Health Psychol. 2020;25:39–60.
27. Spell LA, Richardson JD, Basilakos A, Stark BC, Teklehaimanot A, Hillis AE, et al. Developing, implementing, and improving assessment and treatment fidelity in clinical aphasia research. Am J Speech-Lang Pathol. 2020;29:286–98.
28. Reed DK, Sturges KM. An examination of assessment fidelity in the administration and interpretation of reading tests. Remedial Spec Educ. 2012;34:259–68.
29. Novick DG, Ward K. Why don't people read the manual. Departmental Papers. 2006;15.

ACKNOWLEDGEMENTS

We acknowledge Isa A. McClure, MAPT, Kessler Institute for Rehabilitation and Rebecca Martin, OTD, Kennedy Krieger Institute, Johns Hopkins University School of Medicine for their contributions in the development of the standardized administration and scoring procedures. We further acknowledge our study respondents who participated in the surveys unnamed. Their expertise greatly informed the development of the SCIM-III standardized administration and scoring procedures.

AUTHOR CONTRIBUTIONS

RYK was responsible for leading data collection and analysis of Delphi for SCIM performance-based assessment, developing Delphi surveys for interview assessment, and leading the writing of the manuscript. CCT contributed by assisting in developing Delphi surveys for performance-based and interview assessment, collecting and analyzing data, and writing the manuscript. GH led the data collection and analysis of Delphi for SCIM interview assessment and assisted with writing the interview assessment portions of the manuscript. MJM was responsible for the conceptual design of the study, leading the development of the Delphi surveys, assisting in data analysis, and an equal in writing of the manuscript.

FUNDING

This study was funded by departmental funds from the Department of Occupational Therapist in the Jefferson College of Rehabilitation Sciences in Thomas Jefferson University.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICAL APPROVAL

This study was reviewed and approved by the Institutional Review Board (IRB), "Spinal Cord Independence Measure, Version III (SCIM-III) Procedure Development: A Modified Delphi Survey" (Departmental) Control #20E.150. We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers/animals were followed during the course of this research.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41393-023-00891-5.

Correspondence and requests for materials should be addressed to Rachel Y. Kim.

Reprints and permission information is available at http://www.nature.com/reprints

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