Arthritis Care & Research
Vol. 66, No. 11, November 2014, pp 1725–1730
DOI 10.1002/acr.22337

ORIGINAL ARTICLE

Construct Validity of the Patient-Reported Outcomes Measurement Information System Gastrointestinal Symptom Scales in Systemic Sclerosis

VIVEK NAGARAJA,[1] RON D. HAYS,[2] PUJA P. KHANNA,[1] BRENNAN M. R. SPIEGEL,[3] LIN CHANG,[4] GIL Y. MELMED,[5] ROGER BOLUS,[6] AND DINESH KHANNA[1]

Objective. **Gastrointestinal (GI) involvement is common in patients with systemic sclerosis (SSc; scleroderma). The Patient-Reported Outcomes Measurement Information System (PROMIS) GI symptom item bank captures upper and lower GI symptoms (reflux, disrupted swallowing, nausea/vomiting, belly pain, gas/bloating/flatulence, diarrhea, constipation, and fecal incontinence). The objective of this study was to evaluate the construct validity of the PROMIS GI bank in SSc.**

Methods. **A total of 167 patients with SSc were administered the PROMIS GI bank and the University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale (GIT 2.0) instrument. GIT 2.0 is a multi-item instrument that measures SSc-associated GI symptoms. Product-moment correlations and a multitrait–multimethod analysis of the PROMIS GI scales with the GIT 2.0 symptom scales were used to evaluate convergent and discriminant validity.**

Results. **Patients with SSc GI involvement had PROMIS GI scale scores 0.2–0.7 SD worse than the US general population. Correlations among scales measuring the same domains for the PROMIS GI and GIT 2.0 measures were large, ranging from 0.61 to 0.87 (average r = 0.77). The average correlation between different symptom scales was 0.22, supporting discriminant validity.**

Conclusion. **This study provides support for the construct validity of the PROMIS GI scales in SSc. Future research is needed to assess the responsiveness to change of these scales in patients with SSc.**

INTRODUCTION

Patient-reported outcomes (PROs) are widely used in research and are playing an increasingly important role in clinical practice (1). In a clinical practice, PROs can be administered to identify presence/absence of symptoms or assess symptom severity, which can assist in clinical decision making (2). Unlike the traditional measures of disease burden (direct and indirect expenditures of a disease), PRO instruments document the burden of disease in terms

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, nor does it represent an endorsement by the US Federal Government or the Patient-Reported Outcomes Measurement Information System (PROMIS).

PROMIS was funded with cooperative agreements from the NIH Common Fund Initiative (Northwestern University, grants U54-AR-057951, U01-AR-052177, and U54-AR-057943; American Institutes for Research, grant U54-AR-057926; State University of New York, Stony Brook, grants U01-AR-057948 and U01-AR-052170; University of Washington, Seattle, grants U01-AR-057954 and U01-AR-052171; University of North Carolina, Chapel Hill, grant U01-AR-052181; Children's Hospital of Philadelphia, grant U01-AR-057956; Stanford University, grant U01-AR-052158; Boston University, grant U01-AR-057929; University of California, Los Angeles, grant U01-AR-057936; University of Pittsburgh, grant U01-AR-052155; Georgetown University, grant U01-AR-057971; Children's Hospital Medical Center, Cincinnati, grant U01-AR-057940; University of Maryland, Baltimore, grant U01-AR-057967; and Duke University, grant U01-AR-052186). See Appendix A for principal investigators and NIH science officers.

Dr. Hays's work was supported by the NIH/National Institute on Aging (grants P30-AG-028748 and P30-AG-021684) and the National Center on Minority Health and Health Disparities (grant 2P20-MD-000182). Dr. Dinesh Khanna's work was supported by the NIH/National Institute for Arthritis and Musculoskeletal and Skin Disease (grant K24-AR-063120-02).

[1]Vivek Nagaraja, MD, Puja P. Khanna, MD, MPH, Dinesh Khanna, MD, MSc: University of Michigan, Ann Arbor; [2]Ron D. Hays, PhD: RAND Corporation, Santa Monica, California, and David Geffen School of Medicine at University of California, Los Angeles; [3]Brennan M. R. Spiegel, MD,

Significance & Innovations

- Patient-reported outcomes play an important role in clinical practice. They help to assess the disease burden and guide treatment.
- The Patient-Reported Outcomes Measurement Information System (PROMIS) instruments are more precise than existing legacy measures.
- The newly developed PROMIS gastrointestinal (GI) symptom item bank captures 8 GI-specific symptom scales for luminal GI disorders.
- In patients with systemic sclerosis (SSc; scleroderma), PROMIS GI scales showed construct (convergent and discriminant) validity relative to a legacy instrument (University of California Los Angeles Scleroderma Clinical Trials Consortium gastrointestinal scale [GIT 2.0]).
- Compared to GIT 2.0, the PROMIS GI bank has additional scales that are applicable to patients with SSc.

of impact on daily functioning and well-being, or health-related quality of life (HRQOL) (3).

Gastrointestinal tract (GI) involvement occurs in approximately 90% of patients with systemic sclerosis (SSc; scleroderma) (4,5) and is associated with decline in HRQOL (6,7). The University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale (UCLA SCTC GIT 2.0 [GIT 2.0]) is a multi-item instrument that measures GI symptoms and their impact on HRQOL. Support for the reliability and validity (including responsiveness to change) of the GIT 2.0 was found in different observational cohorts (8–11). It is considered the "legacy" instrument to assess GI involvement in patients with SSc (2).

The National Institutes of Health (NIH) funded the Patient-Reported Outcomes Measurement Information System (PROMIS; trademarked by the NIH) Roadmap initiative (available at www.nihpromis.org) to develop, evaluate, and standardize item banks to measure PROs across patients with different medical conditions and in the US general population (12,13). PROMIS GI symptom item banks that assess 8 GI domains were recently developed (14,15). The goal of PROMIS is to develop reliable and valid item banks using item response theory (IRT) that can be administered in a variety of formats, including short forms and computerized adaptive tests (12,16,17). PROMIS has several advantages over the traditional instruments. First, a consistent qualitative process is employed with detailed systematic review, focus groups, cognitive interviews, and translatability for each item bank. Second, PROMIS static items produce more reliable information than existing measures, such as the Short Form 36 health survey physical functioning component-10 and the Health Assessment Questionnaire disability index (18).

The goal of this study was to evaluate the construct validity of the PROMIS GI symptom scales.

MSHS: VA Greater Los Angeles Healthcare System, David Geffen School of Medicine at University of California, Los Angeles, and VA Center for Outcomes Research and Education, Los Angeles, California; [4]Lin Chang, MD: Gail and Gerald Oppenheimer Family Center for Neurobiology of Stress, David Geffen School of Medicine at University of California Los Angeles; [5]Gil Y. Melmed, MD, MS: Cedars-Sinai Medical Center, Los Angeles, California; [6]Roger Bolus, PhD: David Geffen School of Medicine at University of California, Los Angeles, and the VA Center for Outcomes Research and Education, Los Angeles, California.

Address correspondence to Dinesh Khanna, MD, MSc, Associate Professor of Medicine, Frederick G. Huetwell Professor, Director, University of Michigan Scleroderma Program, Division of Rheumatology/Department of Internal Medicine, 300 North Ingalls Street, Suite 7C27, SPC 5422 Ann Arbor, MI 48109. E-mail: khannad@med.umich.edu.

Submitted for publication October 11, 2013; accepted in revised form March 25, 2014.

MATERIALS AND METHODS

PRO measures. *NIH PROMIS GI symptom item bank.* The GI symptom item banks were developed using the standard PROMIS qualitative and quantitative methodology (19). Briefly, the qualitative aim was achieved through completion of a systematic review of the literature to identify extant GI PRO items followed by a comprehensive review and evaluation of these items (15). The individual items from existing instruments were grouped based on different symptoms. This was complemented by focus group discussions of patients with GI conditions to evaluate their symptoms. New items were developed based on extant items and input from the focus group participants, followed by fine tuning of item wording based on cognitive interviews with GI patients. The items were administered to 865 patients with different GI disorders (including SSc) at the following 4 centers in the US: University of Michigan Hospital, UCLA Medical Center, Cedars Sinai Medical Center, and VA West Los Angeles Medical Center, and to 1,177 individuals from the US general population. The US general population was included to develop norms for clinical care and research. Items were finalized based on psychometric analyses, including categorical confirmatory factor analyses and IRT modeling (15).

The final PROMIS GI symptom instrument (15) has 60 items and assesses 8 domains: gastroesophageal reflux (13 items), disrupted swallowing (7 items), diarrhea (5 items), bowel incontinence/soilage (4 items), nausea and vomiting (4 items), constipation (9 items), belly pain (6 items), and gas/bloating/flatulence (12 items). All scales were calibrated using the 2-parameter IRT graded response model and scored on a T score metric with a mean ± SD of 50 ± 10 in the US general population. A higher score denotes more GI symptoms. The recall period for PROMIS GI symptom items is 1 week.

GIT 2.0. The GIT 2.0 was developed to assess presence/absence and severity of GI involvement and the consequent impairment in social and emotional well-being in patients with SSc (11). Previous work has provided support for the reliability and validity of the GIT 2.0 scales

Table 1. Baseline characteristics of the study participants*

Variables	Total sample (n = 167)
Age, mean ± SD (range) years	53 ± 13 (22–80)
Women	91
Race/ethnicity	
White	67
African American	11
Hispanic	10
Asian	7
Other	5
Education	
≤ College education	2
Some college	16
College graduate	20
Graduate degree	63
Marital status	
Married	65
Never married	10
Widowed/separated/divorced	25
Employment	
Full time/part time	38
Retired	24
Unemployed	4
On disability	22
Homemaker	10
Full-time student	2

* Values are the percentage unless indicated otherwise.

(11,20). The GIT 2.0 scales were found to be sensitive to the presence of abnormalities on structural/motility testing and can be routinely used as an initial screening test in clinical practice (8). It is the "legacy" PRO measure to assess the severity of GI involvement and its impact on HRQOL in patients with SSc. The GIT 2.0 has 34 items; the 7 multi-item scales include reflux (8 items), distention/bloating (4 items), diarrhea (2 items), fecal soilage (1 item), constipation (4 items), emotional well-being (9 items), and social functioning (6 items) (2). The reflux scale has 1 item each for solid food dysphagia, nausea, and vomiting. The items are scored from 0.0–3.0, except diarrhea (0.0–2.0) and constipation (0.0–2.5). Higher values indicate worse HRQOL. The total GIT 2.0 score averages 6 of 7 scales (excluding constipation) and is scored from 0.0 (no GI symptoms) to 2.8 (severe GI symptoms). The recall period for the GIT 2.0 items is 1 week.

Participants. A total of 167 patients with SSc are a subset of 865 patients who were recruited predominantly at the University of Michigan Scleroderma Clinic. The diagnosis of SSc was made based on the 1980 American College of Rheumatology criteria (21) and/or clinical diagnosis by 2 physicians (PPK and DK). This subset of patients was administered the PROMIS GI symptom item bank and GIT 2.0 instrument.

Statistical analysis. Descriptive statistics are presented as means and SDs for continuous variables and percentages for categorical variables. In addition to mean scores and SDs, ranges and percentages of respondents scoring the minimum and maximum possible scores were calculated to evaluate scale score distributions for the PROMIS GI symptom and GIT 2.0 scales. Internal consistency reliability for all scales was estimated using Cronbach's alpha, and reliability for the GIT 2.0 total score was estimated using Mosier's formula. Reliability ≥0.70 was considered satisfactory for group comparisons (22).

Convergent and discriminant validity are 2 components of construct validity. Convergent validity is supported when different methods of assessing the same construct (e.g., 2 measures of reflux) should be highly correlated.

Table 2. Descriptive statistics of PROMIS GI item band and UCLA SCTC GIT 2.0*

Scale	Mean ± SD	Minimum score	Maximum score	% with minimum	% with maximum	Cronbach's α
PROMIS GI symptom item bank						
Reflux	54 ± 8	33	75	1	1	0.83
Disrupted swallowing	56 ± 10	41	83	14	1	0.91
Nausea/ vomiting	54 ± 11	41	84	20	1	0.73
Belly pain	55 ± 10	37	79	2	1	0.88
Gas/bloat/flatulence	57 ± 10	38	79	1	1	0.94
Diarrhea	55 ± 11	40	82	1	1	0.89
Constipation	52 ± 9	37	75	7	1	0.88
Fecal incontinence	54 ± 13	44	91	47	1	0.90
UCLA SCTC GIT 2.0						
Reflux	0.85 ± 0.63	0.0	3.0	8	1	0.80
Distention/bloating	1.38 ± 0.87	0.0	3.0	4	6	0.78
Diarrhea	0.65 ± 0.70	0.0	2.0	45	9	0.72
Fecal soilage	0.47 ± 0.81	0.0	3.0	69	4	NA
Constipation	0.61 ± 0.59	0.0	2.5	26	1	0.74
Emotional well-being	0.66 ± 0.74	0.0	3.0	27	1	0.90
Social functioning	0.52 ± 0.58	0.0	2.33	34	1	0.78
Total GI score	0.75 ± 0.51	0.0	2.34	1	0	0.92

* PROMIS = Patient-Reported Outcomes Measurement Information System; GI = gastrointestinal; UCLA SCTC GIT 2.0 = University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale; NA = not applicable.

Table 3. Product-moment correlations between PROMIS GI symptom item bank and UCLA SCTC GIT 2.0 symptom scales*

	UCLA SCTC GIT 2.0				
	Reflux†	Distention/ bloating	Diarrhea	Constipation	Fecal incontinence
PROMIS GI					
Reflux	0.77‡	0.44	0.13	0.25	−0.03
Disrupted swallowing	0.61‡	0.39	0.16	0.21	0.13
Nausea and vomiting	0.66‡	0.44	0.20	0.22	0.18
Belly pain	0.45	0.49	0.23	0.34	0.04
Gas/bloating/flatulence	0.46	0.73‡	0.30	0.29	0.10
Diarrhea	0.25	0.25	0.65‡	0.02	0.54
Constipation	0.37	0.32	0.05	0.76‡	−0.01
Fecal incontinence	0.12	0.11	0.43	−0.18	0.87‡

* PROMIS = Patient-Reported Outcomes Measurement Information System; GI = gastrointestinal; UCLA SCTC GIT 2.0 = University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale.
† GIT 2.0 reflux scale asks about reflux, dysphagia to solid foods, and nausea/vomiting.
‡ Hypothesized correlation coefficients.

Pearson's product-moment correlations of the PROMIS GI symptom scales with corresponding GIT 2.0 scales were used to assess convergent validity. Discriminant validity is supported when measures of different constructs (e.g., diarrhea and constipation) do not correlate highly with each other. We conducted multitrait, multimethod matrix (MTMM) analyses to evaluate convergent and discriminant validity (23). We hypothesized that correlations among scales measuring the same construct would be significantly larger than other correlations (24). A coefficient of ≥0.50 was considered large for current analysis.

RESULTS

The majority of participants were female (91%), white (54%), and highly educated (98% with some college degree); the mean ± SD age of the sample was 53 ± 13 years (Table 1).

Table 2 provides descriptive statistics and reliability estimates for the PROMIS GI symptom and GIT 2.0 scales. The PROMIS GI scale scores of patients in the sample were 0.2–0.7 SD worse than the US population. The percentage of patients with minimum scores on the PROMIS scales ranged from 1% (for reflux, bloating, and diarrhea scales) to 47% (for fecal incontinence scale), while the percentage with maximum scores was 1% for all scales. Cronbach's coefficient alpha was >0.70 for all scales.

For GIT 2.0, the mean ± SD scores ranged from 0.47 ± 0.81 for fecal soilage, to 1.38 ± 0.87 for distension/bloating scale. The percentage of patients with minimum scores ranged from 4% (for distension/bloating scale) to 69% (for fecal soilage scale), while the percentage with maximum scores ranged from 1% (for reflux, constipation, emotional well-being, and social functioning scales) to 9% (for diarrhea scale). Cronbach's alpha was >0.70 for all scales.

Pearson's product-moment correlations between corresponding scale scores were large (ranging from 0.61 to 0.87) (Table 3). GIT 2.0 does not have separate scales for disrupted swallowing and nausea/vomiting. However,

Table 4. Multitrait, multimethod matrix table of correlations (n = 167)*

	GIT 2.0					PROMIS-GI			
Trait	Reflux	Gas	Diarrhea	Constipation	Incontinence	Reflux	Gas	Diarrhea	Constipation
GIT 2.0									
Reflux	1.00								
Gas	0.57	1.00							
Diarrhea	0.13	0.23	1.00						
Constipation	0.34	0.22	−0.14	1.00					
Incontinence	0.10	0.14	0.38	−0.12	1.00				
PROMIS GI									
Reflux	0.77†	0.44	0.13	0.25	−0.03	1.00			
Gas	0.46	0.73†	0.30	0.29	0.10	0.39	1.00		
Diarrhea	0.25	0.25	0.65†	0.02	0.54	0.23	0.31	1.00	
Constipation	0.37	0.32	0.05	0.76†	−0.01	0.35	0.36	0.20	1.00
Incontinence	0.12	0.11	0.43	−0.18	0.87†	0.06	0.11	0.61	−0.06

* Average convergent validity correlation is 0.766. Average off-diagonal correlation is 0.225 (164 df). GIT 2.0 = University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale; PROMIS = Patient-Reported Outcomes Measurement Information System; GI = gastrointestinal.
† Convergent correlation between the different constructs when the 2 instruments are compared.

the GIT 2.0 reflux scale has a single item assessing solid food dysphagia, as well as nausea and vomiting items. This accounts for the relatively high correlations of the reflux, disrupted swallowing, and nausea/vomiting scales on the PROMIS GI symptom scales with the GIT 2.0 reflux scale.

The average convergent validity correlation in the MTMM was 0.77 and the average off-diagonal correlation was 0.22 (Table 4). *T*-tests of the significance of differences between relevant corresponding correlations for evaluating discriminant validity showed that 39 of 40 hetero-method (convergent correlations compared with correlations among different constructs measured by different methods) and 39 of 40 mono-method (convergent correlations compared with correlations among the different constructs measured by the same method) comparisons were statistically significant in the hypothesized direction, providing strong support for construct validity.

DISCUSSION

GI involvement affects approximately 90% of SSc patients (25) and the majority of patients have symptoms. Although the preferred approach for evaluation of GI pathology is tests such as endoscopy and manometry, it is impractical to perform these tests in every patient, particularly as symptoms evolve over time (2). PRO measures complement objective tests (8), and GI symptoms in SSc are independently associated with poor HRQOL (26).

The PROMIS GI scales assess symptoms that can be used to assess the general population and patients with different GI disorders. In the current study, PROMIS GI symptom scales were compared to the widely used GIT 2.0 to explore its construct validity in patients with SSc. Correlations between PROMIS GI scales and corresponding GIT 2.0 scales were large, and correlations of scales measuring different constructs were small, providing support for construct validity.

The largest correlation between PROMIS GI scales and corresponding GIT 2.0 scales was for fecal incontinence ($r = 0.87$), and this may be attributed to a single item that is worded very similarly to GIT 2.0. A high correlation was also noted between the PROMIS reflux, disrupted swallowing, and nausea/vomiting scales and the GIT 2.0 reflux scale ($r = 0.77$, 0.61, and 0.66, respectively). This is likely, as the GIT 2.0 does not have separate scales for disrupted swallowing and nausea and vomiting; the reflux scale of GIT 2.0 includes an item each for solid food dysphagia, nausea, and vomiting. Solid and liquid dysphagia and nausea/vomiting are common in patients in SSc due to gastroesophageal reflux disease, esophageal dysmotility, and gastroparesis (25,27–29). Hence, separate scales for disrupted swallowing and nausea/vomiting are more meaningful in patients with SSc. The correlations between diarrhea and fecal incontinence scales for the 2 instruments were also noteworthy (range 0.43–0.54). During our qualitative phase for development of the PROMIS GI scale, some patients stated that loose/frequent bowel movements and fecal incontinence were in a continuum rather than separate constructs. This is also supported by negative correlations between fecal incontinence and constipation scales (range −0.01 to −0.18).

PROMIS GI scales and GIT 2.0 demonstrated satisfactory reliability (>0.70 for all scales). The percentage of patients with minimum and maximum scores can limit responsiveness to change in a longitudinal study. In our study, PROMIS scales had lower percentage of patients in different scales who achieved minimum and maximum scores compared to GIT 2.0, suggesting that measurement precision may be better for the PROMIS bank over a wide range compared to GIT 2.0. This will likely increase the ability to detect true change and to fulfill power and sample size requirements (30,31).

Our study has several strengths. First, it provides support for the PROMIS GI symptom scales in patients with SSc. Second, it adds to the limited repertoire of psychometrically sound instruments to assess the GI burden of SSc. In clinical practice and trials, incorporation of either GIT 2.0 or PROMIS GI scales is appropriate. Third, PROMIS GI has advantage of separate scales for disrupted swallowing and nausea/vomiting (applicable in SSc) and will have data available on the same metric that allow comparison of prevalence/severity of symptoms in patients with SSc with the general population and other GI disorders, such as inflammatory bowel disease and irritable bowel disorder. On the other hand, PROMIS GI scales require a computer to calculate the scores, whereas GIT 2.0 can be scored in the office setting.

The study is not without limitations. First, the GIT 2.0 was one of the "legacy" instruments used during the development of GI symptom item banks. Although there are other measures to assess GI involvement in SSc, none have been comprehensively evaluated as GIT 2.0 (2). Second, the study population was quite homogenous, predominantly involving women (91%), mainly whites, and patients who were highly educated (63% with graduate degree). Larger studies in SSc and other GI disorders will need to be conducted to assess responsiveness to change of PROMIS GI item bank versus other "legacy" instruments.

In conclusion, this study provides support for the construct validity of the PROMIS GI symptom item scales in patients with SSc. These items are ready for use in clinical practice to assess the presence and severity of GI symptoms.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. D. Khanna had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Nagaraja, Hays, P. Khanna, Spiegel, Chang, Melmed, Bolus, D. Khanna.

Acquisition of data. P. Khanna, Spiegel, Chang, Melmed, D. Khanna.

Analysis and interpretation of data. Hays, Spiegel, Melmed, Bolus, D. Khanna.

REFERENCES

1. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. J Neurogastroenterol Motil 2013;19:137–48.
2. Khanna D, Nagaraja V, Gladue H, Chey W, Pimentel M, Frech T. Measuring response in the gastrointestinal tract in systemic sclerosis. Curr Opin Rheumatol 2013;25:700–6.

3. Khanna D, Tsevat J. Health-related quality of life: an introduction. Am J Manag Care 2007;Suppl 9:S218–23.
4. Lock G, Holstege A, Lang B, Scholmerich J. Gastrointestinal manifestations of progressive systemic sclerosis. Am J Gastroenterol 1997;92:763–71.
5. Sjogren RW. Gastrointestinal motility disorders in scleroderma. Arthritis Rheum 1994;37:1265–82.
6. Bodukam V, Hays RD, Maranian P, Furst DE, Seibold JR, Impens A, et al. Association of gastrointestinal involvement and depressive symptoms in patients with systemic sclerosis. Rheumatology (Oxford) 2011;50:330–4.
7. Kulig M, Leodolter A, Vieth M, Schulte E, Jaspersen D, Labenz J, et al. Quality of life in relation to symptoms in patients with gastro-oesophageal reflux disease: an analysis based on the ProGERD initiative. Aliment Pharmacol Ther 2003;18:767–76.
8. Bae S, Allanore Y, Furst DE, Bodukam V, Coustet B, Morgaceva O, et al. Associations between a scleroderma-specific gastrointestinal instrument and objective tests of upper gastrointestinal involvements in systemic sclerosis. Clin Exp Rheumatol 2013;Suppl 76:57–63.
9. Baron M, Hudson M, Steele R, Lo E, and the Canadian Scleroderma Research Group. Validation of the UCLA Scleroderma Clinical Trial Gastrointestinal Tract Instrument version 2.0 for systemic sclerosis. J Rheumatol 2011;38:1925–30.
10. Khanna D, Furst DE, Maranian P, Seibold JR, Impens A, Mayes MD, et al. Minimally important differences of the UCLA Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. J Rheumatol 2011;38:1920–4.
11. Khanna D, Hays RD, Maranian P, Seibold JR, Impens A, Mayes MD, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009;61:1257–63.
12. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007;Suppl 1:S3–11.
13. Khanna D, Krishnan E, Dewitt EM, Khanna PP, Spiegel B, Hays RD. The future of measuring patient-reported outcomes in rheumatology: Patient-Reported Outcomes Measurement Information System (PROMIS). Arthritis Care Res (Hoboken) 2011;Suppl 11:S486–90.
14. Spiegel BM, Khanna D, Bolus R, Agarwal N, Khanna P, Chang L. Understanding gastrointestinal distress: a framework for clinical practice. Am J Gastroenterol 2011;106:380–5.
15. Spiegel BM, Hays RD, Bolus R, Melmed GY, Chang L, Whitman C, et al. Development of the NIH Patient Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am J Gastroenterol 2014. In press.
16. Hays RD, Liu H, Spritzer K, Cella D. Item response theory analyses of physical functioning items in the medical outcomes study. Med Care 2007;Suppl 1:S32–8.
17. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care 2007;Suppl 1:S22–31.
18. Fries JF, Cella D, Rose M, Krishnan E, Bruce B. Progress in assessing physical function in arthritis: PROMIS short forms and computerized adaptive testing. J Rheumatol 2009;36:2061–6.
19. DeWalt DA, Rothrock N, Yount S, Stone AA, Group PC. Evaluation of item candidates: the PROMIS qualitative item review. Med Care 2007;Suppl 1:S12–21.
20. Bae S, Allanore Y, Coustet B, Maranian P, Khanna D. Development and validation of French version of the UCLA Scleroderma Clinical Trial Consortium gastrointestinal tract instrument. Clin Exp Rheumatol 2011;Suppl 65:S15–21.
21. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 1980;23:581–90.
22. Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.
23. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56:81–105.
24. Khanna D, Maranian P, Rothrock N, Cella D, Gershon R, Khanna PP, et al. Feasibility and construct validity of PROMIS and "legacy" instruments in an academic scleroderma clinic. Value Health 2012;15:128–34.
25. Akesson A, Wollheim FA. Organ manifestations in 100 patients with progressive systemic sclerosis: a comparison between the CREST syndrome and diffuse scleroderma. Brit J Rheumatol 1989;28:281–6.
26. Frech T, Hays RD, Maranian P, Clements PJ, Furst DE, Khanna D. Prevalence and correlates of sleep disturbance in systemic sclerosis: results from the UCLA scleroderma quality of life study. Rheumatology (Oxford) 2011;50:1280–7.
27. Ntoumazios SK, Voulgari PV, Potsis K, Koutis E, Tsifetaki N, Assimakopoulos DA. Esophageal involvement in scleroderma: gastroesophageal reflux, the common problem. Semin Arthritis Rheum 2006;36:173–81.
28. Sallam H, McNearney TA, Chen JD. Systematic review: pathophysiology and management of gastrointestinal dysmotility in systemic sclerosis (scleroderma). Aliment Pharmacol Ther 2006;23:691–712.
29. Yarze JC, Varga J, Stampfl D, Castell DO, Jimenez SA. Esophageal function in systemic sclerosis: a prospective evaluation of motility and acid reflux in 36 patients. Am J Gastroenterol 1993;88:870–6.
30. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and internet administration. J Rheumatol 2011;38:1759–64.
31. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). J Clin Epidemiol 2008;61:17–33.

APPENDIX A: PRINCIPAL INVESTIGATORS AND NIH SCIENCE OFFICERS FOR THE PROMIS INITIATIVE

Principal investigators for PROMIS are as follows: David Cella, PhD, Richard C. Gershon, PhD (Northwestern University); Susan (San) D. Keller, PhD (American Institutes for Research); Joan E. Broderick, PhD, Arthur A. Stone, PhD (State University of New York, Stony Brook); Heidi M. Crane, MD, MPH, Paul K. Crane, MD, MPH, Donald L. Patrick, PhD, Dagmar Amtmann, PhD (University of Washington, Seattle); Harry A. Guess, MD, PhD (deceased), Darren A. DeWalt, MD, MPH (University of North Carolina, Chapel Hill); Christopher B. Forrest, MD, PhD (Children's Hospital of Philadelphia); James F. Fries, MD (Stanford University); Alan Jette, PT, PhD, Stephen M. Haley, PhD (deceased), David Scott Tulsky, PhD (current address: University of Michigan, Ann Arbor) (Boston University); Dinesh Khanna, MD (current address: University of Michigan, Ann Arbor), Brennan Spiegel, MD, MSHS (University of California, Los Angeles); Paul A. Pilkonis, PhD (University of Pittsburgh); Carol. M. Moinpour, PhD (current address: Fred Hutchinson Cancer Research Center, Seattle), Arnold L. Potosky, PhD (Georgetown University); Esi M. Morgan DeWitt, MD, MSCE (Children's Hospital Medical Center, Cincinnati); Lisa M. Shulman, MD (University of Maryland, Baltimore); Kevin P. Weinfurt, PhD (Duke University).

NIH Science Officers on this project have included Deborah Ader, PhD, Vanessa Ameen, MD (deceased), Susan Czajkowski, PhD, Basil Eldadah, MD, PhD, Lawrence Fine, MD, DrPH, Lawrence Fox, MD, PhD, Lynne Haverkos, MD, MPH, Thomas Hilton, PhD, Laura Lee Johnson, PhD, Michael Kozak, PhD, Peter Lyster, PhD, Donald Mattison, MD, Claudia Moy, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Peter Scheidt, MD, Ashley Wilder Smith, PhD, MPH, Susana Serrate-Sztein, MD, William Phillip Tonkins, DrPH, Ellen Werner, PhD, Tisha Wiley, PhD, and James Witter, MD, PhD.