

UCLA

UCLA Previously Published Works

Title

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

Permalink

<https://escholarship.org/uc/item/4mz668gq>

Journal

Arthritis Care & Research, 66(11)

ISSN

2151-464X

Authors

Nagaraja, Vivek

Hays, Ron D

Khanna, Puja P

et al.

Publication Date

2014-11-01

DOI

10.1002/acr.22337

Peer reviewed

Construct validity of the Patient Reported Outcomes Measurement Information System (PROMIS®) Gastrointestinal Symptom Scales in Systemic Sclerosis

Vivek Nagaraja, MD¹; Ron D. Hays, PhD^{2, 3}; Puja P. Khanna¹; MD, MPH¹; Brennan M.R. Spiegel, MD, MSHS^{4, 5, 6}; Lin Chang, MD^{5, 7}; Gil Y. Melmed, MD, MS⁸; Roger Bolus, PhD^{5, 6}; Dinesh Khanna, MD, MSc¹

(1) Division of Rheumatology, University of Michigan, Ann Arbor, MI; (2) Division of General Internal Medicine & Health Services Research, David Geffen School of Medicine at UCLA, and Department of Health Services, UCLA School of Public Health; (3) RAND, Santa Monica, CA; (4) Department of Gastroenterology, VA Greater Los Angeles Healthcare System; (5) Division of Digestive Diseases, David Geffen School of Medicine at UCLA; (6) UCLA/VA Center for Outcomes Research and Education; (7) Gail and Gerald Oppenheimer Family Center for Neurobiology of Stress, David Geffen School of Medicine at UCLA; and (8) Division of Gastroenterology, Department of Medicine, Cedars-Sinai Medical Center

Corresponding author:

Dinesh Khanna, MD, MSc

Associate Professor of Medicine

Director, University of Michigan Scleroderma Program

Division of Rheumatology/Dept. of Internal Medicine

300 North Ingalls Street, Suite 7C27

Ann Arbor, MI 48109

Email: khannad@med.umich.edu

Phone: 734.647.8173

Fax: 734.763.5761

Grant support: The research reported in this publication was supported by NIH/NIAMS U01 AR057936A, the National Institutes of Health through the NIH Roadmap for Medical Research Grant (AR052177) The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dinesh Khanna was also supported by the NIH/NIAMS K24AR063120-02. Ron D. Hays was supported by NIH/NIA Grants P30-AG028748 and P30-AG021684, and NCMHD Grant 2P20MD000182.

PROMIS® was funded with cooperative agreements from the National Institutes of Health (NIH) Common Fund Initiative (Northwestern University, PI: David Cella, PhD, U54AR057951, U01AR052177; Northwestern University, PI: Richard C. Gershon, PhD, U54AR057943; American Institutes for Research, PI: Susan (San) D. Keller, PhD, U54AR057926; State University of New York, Stony Brook, PIs: Joan E. Broderick, PhD and Arthur A. Stone, PhD, U01AR057948,

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/acr.22337

© 2014 American College of Rheumatology

Received: Oct 11, 2013; Revised: Mar 18, 2014; Accepted: Mar 25, 2014

U01AR052170; University of Washington, Seattle, PIs: Heidi M. Crane, MD, MPH, Paul K. Crane, MD, MPH, and Donald L. Patrick, PhD, U01AR057954; University of Washington, Seattle, PI: Dagmar Amtmann, PhD, U01AR052171; University of North Carolina, Chapel Hill, PI: Harry A. Guess, MD, PhD (deceased), Darren A. DeWalt, MD, MPH, U01AR052181; Children's Hospital of Philadelphia, PI: Christopher B. Forrest, MD, PhD, U01AR057956; Stanford University, PI: James F. Fries, MD, U01AR052158; Boston University, PIs: Alan Jette, PT, PhD, Stephen M. Haley, PhD (deceased), and David Scott Tulsky, PhD (University of Michigan, Ann Arbor), U01AR057929; University of California, Los Angeles, PIs: Dinesh Khanna, MD (University of Michigan, Ann Arbor) and Brennan Spiegel, MD, MSHS, U01AR057936; University of Pittsburgh, PI: Paul A. Pilkonis, PhD, U01AR052155; Georgetown University, PIs: Carol M. Moinpour, PhD (Fred Hutchinson Cancer Research Center, Seattle) and Arnold L. Potosky, PhD, U01AR057971; Children's Hospital Medical Center, Cincinnati, PI: Esi M. Morgan DeWitt, MD, MSCE, U01AR057940; University of Maryland, Baltimore, PI: Lisa M. Shulman, MD, U01AR057967; and Duke University, PI: Kevin P. Weinfurt, PhD, U01AR052186). 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-Sztejn, MD, William Phillip Tonkins, DrPH, Ellen Werner, PhD, Tisha Wiley, PhD, and James Witter, MD, PhD. The contents of this article uses data developed under PROMIS. These contents do not necessarily represent an endorsement by the US Federal Government or PROMIS. See <http://www.nihpromis.org> for additional information on the PROMIS® initiative.

Objective: Gastrointestinal (GI) involvement is common in patients with systemic sclerosis (SSc). 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: 167 patients with SSc were administered the PROMIS GI bank and the UCLA 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 US 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.

Significance and innovations

1. Patient reported outcomes (PRO) play an important role in clinical practice. They help to assess the disease burden and guide treatment.
2. PROMIS instruments are more precise than existing legacy measures.

3. Newly developed PROMIS GI Symptom item bank captures 8 GI-specific symptom scales for luminal GI disorders.
4. In patients with SSc, PROMIS GI scales showed construct (convergent and discriminant) validity relative to a legacy instrument (UCLA Scleroderma Clinical Trials Consortium Gastrointestinal scale [GIT 2.0]).
5. Compared to GIT 2.0, PROMIS GI bank has additional scales that are applicable to patients with SSc.

INTRODUCTION

Patient-reported outcomes (PROs) are widely used in research and they 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 symptoms 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 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) (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 from hereon) 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®) Roadmap to develop, evaluate, and standardize item banks to measure patient-reported outcomes 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 SF-36 physical functioning-10 and health assessment questionnaire-disability index (18).

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

MATERIALS AND METHODS

Patient-reported outcome 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 4 centers in United States: University of Michigan Hospital, University of California Los Angeles Medical Center, Cedars Sinai

Medical Center and VA West Los Angeles Medical Center and 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 item response theory 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 two-parameter item response theory (IRT) graded response model and scored on a T-score metric with a mean of 50 and SD of 10 in the U.S. 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 gastrointestinal 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 (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 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

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 PPK or DK. This subset of patients were administered the PROMIS GI Symptom item bank and GIT 2.0 instrument.

Statistical analysis

Descriptive statistics are presented as mean (standard deviations) for continuous variables and percentages for categorical variables. In addition to mean scores and standard deviations (SD), 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 two components of construct validity.

Convergent validity is supported when different methods of assessing the same construct (e.g. two measures of reflux) should be highly correlated. 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%), Caucasian (54%), and highly educated (98% with some college degree); the mean age of the sample was 53 years (SD = 13; Table 1).

Table 2 provides descriptive statistics and reliability estimates for the PROMIS GI Symptom and GIT 2.0 scales. Patients in the sample had PROMIS GI scale scores 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, the GIT 2.0 reflux scale has a single item assessing solid food dysphagia, and 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 out of 40 hetero-method (convergent correlations compared with correlations among different constructs measured by different methods) and 39 out 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 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 GERD, 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 to 0.54). During our qualitative phase for development of 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 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. Next, 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 same metric that allows comparison of prevalence/ severity of symptoms in patients with SSc with 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 bank. 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 females (91%), mainly Caucasians 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 vs. 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.

References

1. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. *Journal of neurogastroenterology and motility*. 2013;19(2):137-48.
2. Khanna D, Nagaraja V, Gladue H, Chey W, Pimentel M, Frech T. Measuring response in the gastrointestinal tract in systemic sclerosis. *Current opinion in rheumatology*. 2013;25(6):700-6.
3. Khanna D, Tsevat J. Health-related quality of life--an introduction. *The American journal of managed care*. 2007;13 Suppl 9:S218-23.
4. Lock G, Holstege A, Lang B, Scholmerich J. Gastrointestinal manifestations of progressive systemic sclerosis. *The American journal of gastroenterology*. 1997;92(5):763-71.
5. Sjogren RW. Gastrointestinal motility disorders in scleroderma. *Arthritis and rheumatism*. 1994;37(9):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*. 2011;50(2):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. *Alimentary pharmacology & therapeutics*. 2003;18(8):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. *Clinical and experimental rheumatology*. 2013;31(2 Suppl 76):57-63.

9. Baron M, Hudson M, Steele R, Lo E, Canadian Scleroderma Research G. Validation of the UCLA Scleroderma Clinical Trial Gastrointestinal Tract Instrument version 2.0 for systemic sclerosis. *The Journal of rheumatology*. 2011;38(9):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. *The Journal of rheumatology*. 2011;38(9):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 and rheumatism*. 2009;61(9):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. *Medical care*. 2007;45(5 Suppl 1):S3-S11.
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 & research*. 2011;63 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. *The American journal of gastroenterology*. 2011;106(3):380-5.
15. Spiegel BM, Hays RD, Bolus R, Melmed GY, Chang L, Whitman C, Khanna PP, Paz SH, Hays T, Khanna D. Development of the NIH Patient Reported Outcomes Measurement Information System (PROMIS®) Gastrointestinal Symptom Scales. *The American journal of gastroenterology*. 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. *Medical care*. 2007;45(5 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). *Medical care*. 2007;45(5 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. *The Journal of rheumatology*. 2009;36(9):2061-6.
19. DeWalt DA, Rothrock N, Yount S, Stone AA, Group PC. Evaluation of item candidates: the PROMIS qualitative item review. *Medical care*. 2007;45(5 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. *Clinical and experimental rheumatology*. 2011;29(2 Suppl 65):S15-21.

21. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis and rheumatism*. 1980;23(5):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. *Psychological bulletin*. 1959;56(2):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 in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2012;15(1):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. *British journal of rheumatology*. 1989;28(4):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*. 2011;50(7):1280-7.
27. Ntounmazios SK, Voulgari PV, Potsis K, Koutis E, Tsifetaki N, Assimakopoulos DA. Esophageal involvement in scleroderma: gastroesophageal reflux, the common problem. *Seminars in arthritis and rheumatism*. 2006;36(3):173-81.
28. Sallam H, McNearney TA, Chen JD. Systematic review: pathophysiology and management of gastrointestinal dysmotility in systemic sclerosis (scleroderma). *Alimentary pharmacology & therapeutics*. 2006;23(6):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. *The American journal of gastroenterology*. 1993;88(6):870-6.
30. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. *The Journal of rheumatology*. 2011;38(8):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). *Journal of clinical epidemiology*. 2008;61(1):17-33.

Table 1 describes the baseline characteristics of the study participants

Variables	Total sample (n=167)
Age in years, Mean (range), SD	53 (22-80), SD = 13
Women (%)	91%
Race / Ethnicity	
White	67%
African American	11%
Hispanic	10%
Asian	7%
Other	5%
Education	
Less than or equal to college education	2%
Some college	16%
College graduate	20%
Graduate degree	63%
Marital status	
Married	65%
Never married	10%
Widowed / separated / divorced	25%
Employment	
Employed full-time/part-time	38%
Retired	24%
Unemployed	4%
On disability	22%

Homemakers
Full-time student

10%
2%

Table 2: Descriptive statistics of PROMIS GI item bank and 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	N/A
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

Accepted Article

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

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/ bloat/ 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

*GIT 2.0 Reflux scale asks about reflux, dysphagia to solid foods, and nausea/ vomiting. Bold represents hypothesized correlation coefficients.

Table 4: Multitrait-Multimethod Matrix (MTMM) table of correlations

Method	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

N = 167; DFS = 164

Average convergent validity correlation is 0.766

Average off-diagonal correlation is 0.225

Table 1: Baseline characteristics of the study participants

Variables	Total sample (n=167)
Age in years, Mean (range), SD	53 (22-80), SD = 13
Women (%)	91%
<i>Race / Ethnicity</i>	
White	67%
African American	11%
Hispanic	10%
Asian	7%
Other	5%
<i>Education</i>	
Less than or equal to college education	2%
Some college	16%
College graduate	20%
Graduate degree	63%
<i>Marital status</i>	
Married	65%
Never married	10%
Widowed / separated / divorced	25%
<i>Employment</i>	
Employed full-time/part-time	38%
Retired	24%
Unemployed	4%
On disability	22%
Homemakers	10%
Full-time student	2%

Table 2: Descriptive statistics of PROMIS GI item bank and 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	N/A
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

Table 3: Product-moment correlations between PROMIS GI Symptoms item bank and UCLA SCTC GIT 2.0 symptoms 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/ bloat/ 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

*GIT 2.0 Reflux scale asks about reflux, dysphagia to solid foods, and nausea/ vomiting. Bold represents hypothesized correlation coefficients.

Table 4: Multitrait-Multimethod Matrix (MTMM) table of correlations

<i>Method</i>	<i>Trait</i>	GIT 2.0					PROMIS-GI			
		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

N = 167; *DFS* = 164

Average convergent validity correlation is 0.766

Average off-diagonal correlation is 0.225