



Published in final edited form as:

J Sex Med. 2018 December ; 15(12): 1792–1810. doi:10.1016/j.jsxm.2018.09.015.

Psychometric Evaluation of PROMIS® Sexual Function and Satisfaction Measures in a Longitudinal Population-based Cohort of Men with Localized Prostate Cancer

Bryce B. Reeve^{1,2}, Mian Wang³, Kevin Weinfurt^{1,2,4}, Kathryn E. Flynn⁵, Deborah S. Usinger^{3,6}, and Ronald C. Chen^{3,6,7}

¹Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, USA

²Duke Cancer Institute, Duke University, Durham, NC, USA

³Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁴Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA

⁵Department of Medicine, Medical College of Wisconsin, Milwaukee, WI, USA

⁶Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁷Department of Radiation Oncology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Abstract

Background—There are multiple treatment options for men with localized prostate cancer that provide similar curative efficacy but differ in their impact on sexual functioning.

Aim—This study evaluated the psychometric properties of the Patient-Reported Outcomes Measurement Information System® (PROMIS®) Sexual Function and Satisfaction (SexFS) measures, including items from version 1 and 2 of the short forms.

Methods—A population-based cohort of men across North Carolina completed surveys via phone interviews at baseline (prior to treatment) and at 3-, 12-, and 24-months post cancer

Corresponding Author: Bryce B. Reeve, PhD, Professor, Population Health Sciences, Duke University School of Medicine, DUMC 104023, 2200 West Main St, Room #771, Durham, NC 27705, Bryce.Reeve@Duke.edu, Phone: 919-613-7812.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

treatment initiation. Surveys included the PROMIS SexFS domains of Interest in Sexual Activity, Erectile Function, Orgasm, and Satisfaction and the Prostate Cancer Symptom Indices (PCSI). Analyses included descriptive statistics, assessment of factorial validity using confirmatory factor analysis (CFA) and item response theory (IRT), tests for differential item functioning, assessment of convergent validity using correlations, and evaluation of responsiveness of the PROMIS SexFS measures over time. We hypothesized that men undergoing surgery (prostatectomy) would report the poorest sexual function at 3-month survey.

Results—Sample size varied by assessment point and ranged from 332 to 939 men, consisting of 30% non-white men and 30% of sample with a high school degree or less. The items within the PROMIS orgasm domain did not associate together to form a unidimensional scale. PROMIS measures of Interest in Sexual Activity, Erectile Function, and Satisfaction were unidimensional and highly correlated with related PCSI measures (e.g., erectile function, $r=0.84-0.95$). Erectile Function in the Surgery group declined more at 3-months compared to the No-Surgery group (2 points); this difference narrowed at 12- and 24-months as the Surgery group recovered over time. Results were similar for PROMIS Interest in Sexual Activity and PROMIS Satisfaction scales.

Clinical Implications—The PROMIS SexFS measures may be used to identify effective interventions to treat sexual dysfunction and monitor sexual functioning in men with prostate cancer over time.

Strengths & Limitations—This study was limited to men living in North Carolina who could self-report their HRQOL in English. However, this study was able to include more men from vulnerable populations by allowing the men to self-report over the phone.

Conclusion—This study provided strong support for use of the PROMIS SexFS (version 2) measures in men with localized prostate cancer to assess sexual interest, erectile function and satisfaction over time.

Keywords

PROMIS; erectile function; satisfaction with sex; prostate cancer

Introduction

Prostate cancer is the most prevalent cancer among men with an estimated 164,590 new diagnoses in 2018¹. There are multiple treatment options for men with localized prostate cancer which provide similar curative efficacy but differ in their impact on health-related quality of life (HRQOL)²; therefore, the latter is often an important consideration in patients' decision-making process. Many patients with prostate cancer report that choosing a treatment which allows preservation of their sexual function is highly important³. For example, prostatectomy has been shown to reduce erectile functioning⁴⁻⁷. The importance of assessing sexual function and satisfaction in this patient population is well-recognized, and all existing, validated prostate cancer-specific PRO instruments include sexual items and/or domains^{8,9}.

In 2004, the National Institutes of Health (NIH) initiated the Patient-Reported Outcomes Measurement Information System[®] (PROMIS[®]) to provide standardized high quality PRO

measures¹⁰. The PROMIS investigators used both qualitative and quantitative methods to design and validate measures of common PROs (e.g., fatigue, pain, depression, anxiety, physical function) in the general population and patients with a variety of chronic diseases¹¹. Sexual functioning was not included in the original list of PRO domains for development; however, the National Cancer Institute (NCI) identified these concepts as a priority given how they can be impacted by cancer and its treatments. Through a supplement from the NCI, Drs. Weinfurt and Flynn led a multidisciplinary group to design and publish version 1 of the sexual function and satisfaction (SexFS) measures¹². Relevant for men with prostate cancer, the PROMIS SexFS measures assess interest in sexual activity, erectile function, orgasm, and global satisfaction with sex life. With subsequent NIH funding, Drs. Weinfurt, Flynn, and colleagues refined the SexFS measures, including adding new domains and items, as well as evaluating the measure in the general population. A result of this work is the SexFS V2¹³.

The goal of this study is to evaluate the level of evidence for the validity and reliability of the PROMIS SexFS measures in a sample of men receiving treatment for localized prostate cancer. Men were participants of a larger comparative effectiveness research (CER) study examining short term and long term outcomes of treatment modalities¹⁴. At the time of the launch of the CER study, PROMIS SexFS V2 was not available; however the CER study included items from PROMIS SexFS V1 and new candidate items being considered for PROMIS SexFS V2. Further, all men in this study completed the PRO measures via phone interview, which is a departure from previous studies of the PROMIS measures that collected data mostly by web-based surveys.

Methods

Study Design and Participants

As a longitudinal, population-based, observational, CER study, the North Carolina Prostate Cancer Comparative Effectiveness and Survivorship Study (NC ProCESS) examines the impact of localized prostate cancer and treatments on the lives of men¹⁴. The overarching goal of the study was to prospectively evaluate cancer-specific and patient-reported outcomes of these men with newly diagnosed prostate cancer. Participants completed surveys via phone interviews at baseline (prior to treatment, if any) and at 3-, 12-, and 24-months post treatment initiation.

From January 2011 to June 2013, out of the 2473 eligible English-speaking men, 1419 of them with newly diagnosed prostate cancer were recruited from all 100 counties in North Carolina (NC) using the rapid case ascertainment (RCA) mechanism of the North Carolina Central Cancer Registry. This current PROMIS psychometric study operated within the parent NC ProCESS project, and therefore the PROMIS measures (including SexFS items) were administered to subsamples of NC ProCESS participants at the four assessment points¹⁵. However, since the PROMIS psychometric study started after the launch of the parent study, sample sizes are smaller at baseline and increase over assessment points.

The study (#10–1483) was approved by the University of North Carolina Institutional Review Board.

Measures

PROMIS short forms included measures of Fatigue, Pain, Depression, Anxiety, Physical Functioning and four PROMIS SexFS domains of Interest in Sexual Activity, Erectile Function, Orgasm, and Global Satisfaction with Sex Life. There were a few PROMIS SexFS domains not included in this study because they were deemed not relevant for this prostate cancer study, including oral discomfort, anal discomfort, and female-specific domains such as vaginal discomfort. The PROMIS SexFS items have undergone rigorous evaluation, including validation in cancer patients^{12 13, 16–19}. Tables

Table 1 presents the SexFS items along with their question stems and response options. Table 1 indicates which items were part of the original PROMIS SexFS-V1 and items that were included for PROMIS SexFS-V2. Table 1 also includes items that were potential additions for SexFS-V2 but were not selected because of poor performance on subsequent testing. At the time of our study, we did not know which candidate items for SexFS-V2 would be selected. Please note, however, this CER study is missing one item included in the SexFS-V2 of Sexual Satisfaction (SFSAT103: *How often have you thought that your sex life is wonderful?*).

In addition, the parent NC ProCESS study used the Prostate Cancer Symptom Indexes (PCSI) for the assessment of prostate cancer patients' quality of life⁸, which includes two subscales on Erectile Function and Sexual Problems. The two subscales are provided in Table 1 and included in our study to facilitate the evaluation of convergent validity of PROMIS SexFS measures.

Analysis

Descriptive statistics and missing data patterns across the four assessment points are summarized and reported first. Confirmatory factor analysis (CFA) and item response theory (IRT) models were used to evaluate the unidimensionality of the scale. These approaches were important for determining the structural validity of each PROMIS SexFS scale. We performed differential item functioning (DIF) between surgery (prostatectomy) and non-surgery treatment to make sure the PROMIS SexFS items perform similarly across groups. Multidimensional and longitudinal two-group (surgery versus no-surgery) IRT models were used to assess relationships among PROMIS SexFS domains at each assessment point and longitudinally. These approaches were important for assessing the responsiveness of the PROMIS SexFS measures over time. Single-factor ordinal CFA models are fit by the WLSMV estimator²⁰, unidimensional graded-response IRT models (including within-wave two-group models) by the conventional quadrature based expectation-maximization algorithm²¹, and multidimensional IRT models by the more computationally-efficient Metropolis–Hastings Robbins–Monro algorithm²². Details of the analysis plan are provided below in each sub-section.

Descriptive statistics—Total sample sizes of the parent study are reported, as well as sample sizes of those who responded to the PROMIS SexFS items. Frequencies based on key demographic and clinical background information are also reported, and the corresponding percentages are calculated out of the PROMIS SexFS subsamples.

Missing data.—True missing values (e.g., skipped, unknown, or refused responses) are distinguished from missing values representing the “inapplicable” responses provided by sexually-inactive participants. Missing data percentages are calculated either with or without the sexually-inactive participants within each time point. In preparation for the unidimensional CFA models, each domain within each wave is also tested for the missing completely at random (MCAR) assumption²³ with their p values adjusted²⁴.

Evaluation of modeling assumptions—To ensure that unidimensional CFA/IRT models are appropriate for each domain, three important modeling assumptions are evaluated: 1) unidimensionality (i.e., items belonging to the same SexFS domain should measure only one common underlying factor); 2) item local independence (i.e., items should have little/no associations with other items other than being related through one common SexFS factor); and 3) monotonicity (i.e., men with higher SexFS domain scores should be more likely to endorse item response categories that reflect higher levels of the measured domain).

During the evaluation process, single-factor ordinal CFA models are carried out, in conjunction with unidimensional parametric IRT graded response²⁵ and nonparametric IRT Mokken scale analysis models²⁶. Results obtained from these models are then compared against commonly-accepted criteria to determine whether the three assumptions hold for a given domain.

For unidimensionality, a model must show high communalities ($h^2 \geq 0.6$, which is equivalent to having a factor loading $\lambda \geq 0.775$ under a unidimensional model;^{27,28} large comparative fit index (CFI ≥ 0.95 ;^{29,30} minimal residual mean square error of approximation (RMSEA < 0.06 ;³⁰ and small weighted root mean square residuals (WRMR < 1 ;³¹. For item local independence, each item pair must show no significant residual covariation (i.e., LD- χ^2 statistic should be smaller than the 95% cutoff of a χ^2 distribution with $df = [K-1][K'-1]$, where K and K' are the number of categories for the two items³². For monotonicity, each item must show no significant monotonicity violations when conditioning on rest scores^{33,34}.

For every PROMIS SexFS domain at each time point, the above procedures are carried out using all available items in the NC ProCESS study to identify a subset of items that best meet the three assumptions, resulting in a SexFS-Preferred version for that domain. The best fitting version is selected as the basis for all subsequent multidimensional models.

In addition, the two PCSI sexual functioning subscales are also evaluated using the same procedures and criteria before entering the multidimensional models, because PCSI were originally developed under classical test theory³⁵.

Assessment of convergent and discriminant validity—Within each assessment point, multidimensional IRT models (all domain factors are fixed to be standard normal, and they can covary) are implemented to capture the between-factor correlations, as an assessment of convergent and discriminant validity.

For convergent validity, we expect strong correlations between measures of the same attribute (e.g., the two erectile function measures), and medium to strong correlations between all sex-related domains. For discriminant validity, we expect relatively weaker correlations between non-sex-related domains (i.e., PROMIS Fatigue and Physical Function) and sex-related domains.

Detection of differential item functioning—To ensure that PROMIS SexFS items (of the best-fitting version) have invariant (or unbiased) factor loadings and intercepts, they are also tested for both between-group (BG-DIF) and longitudinal differential item functioning (L-DIF). BG-DIF was evaluated with respect to surgery vs no-surgery, because differences in sexual functioning between prostatectomy and other non-surgery procedures are often used to inform treatment choices for men with prostate cancer. The BG-DIF analysis repeats through each assessment time point for every SexFS domain, and it is based on a traditional two-group approach that simultaneously fits two unidimensional models. The L-DIF test only repeats through every domain, and it is based on a longitudinal (bifactor) approach that fits four correlated primary factors (to capture cross-time factor autocorrelations) and S number of uncorrelated specific factors (to capture cross-time item autocorrelations; S = test length; ³⁶.

Regardless of the approach, we first adopt the likelihood ratio test all-others-as-anchors procedure (LRT-AOAA ³⁷ and the MaxA1 criterion (MinG²1 criterion is applied if all items are significant during AOAA ³⁸)) to select a group-/time-invariant anchor item, which is then used to link groups or time points onto a common metric. Once anchors are identified, non-anchor items are tested (one at a time) for DIF using LRT ^{37, 39} while parameters of the designated anchor are constrained equal between groups or across waves. Items that show nonsignificant between-group/cross-time parameter differences are later used to link the two groups and/or the four time points in the final models (see next section). Note that, for a given domain, there will be four group-invariant anchors (one for each time point, though they can overlap) and at most two time-invariant anchors, because BG-DIF tests always repeat through the four time points whereas L-DIF tests need to repeat through the two groups only if BG-DIF is found. On a related note, if no significant BG-DIF effect is found across waves, L-DIF tests for that domain will utilize all available sample (i.e., ignoring group membership).

Investigation of responsiveness of PROMIS SexFS over time—Finally, a longitudinal two-group IRT model is fit for every PROMIS SexFS measure that survives both the assumption checks and the DIF tests (see Figure 1 for an example).

Within this final model, a common metric is first set by fixing the baseline no-surgery group factor to a standard normal latent distribution, leaving the other factor means and variances to be estimated. Then, the two groups are linked by equating items that show no BG-DIF, and the four time points associated with each group are linked by equating items that show no L-DIF (linking items can be different for each group depending on DIF test results), so that the remaining factor means and variances are estimable.

These longitudinal analyses examine the responsiveness of the PROMIS measures to changes over time in sexual functioning and satisfaction that are consistent with what is theorized. It is expected that prostatectomy will have a strong negative impact (i.e., lower factor means) on erectile function especially within the first year, because prostatectomy is closely tied with erectile dysfunction^{29, 40–43}. We also expect sexual satisfaction to be negatively affected by prostatectomy due to erectile dysfunction. We also expect a decrease in sexual interest for men undergoing prostatectomy, though the change is expected to be smaller than decreases in erectile function or sexual satisfaction⁴².

Software—Analyses for the current study are conducted in R version 3.3.2⁴⁴. Besides the built-in base package, the following R packages are applied: *BaylorEdPsych* version 0.5⁴⁵ for MCAR tests, *mirt* version 1.23 for parametric IRT modeling⁴⁶, *lavaan* version 0.5–22 for CFA modeling⁴⁷, *mokken* version 2.8.5 for nonparametric IRT modeling³⁴, *ggplot2* version 2.2.1 for plotting⁴⁸.

Results

Demographics

Table 2 shows the demographic and clinical characteristics of participants across the four assessment time points. Over time, the number of men who completed the PROMIS SexFS measures steadily increased, whereas the total sample sizes of NC ProCESS show an average attrition rate of 12.5% over time.

Across the four assessment points, the PROMIS SexFS subsample consisted of mostly non-Hispanic Whites (> 70%) and Blacks (approximately 25%), and approximately 30% of men had a high school diploma or less. More than one third of men opted to undergo prostatectomy to treat their cancer.

Missing Data

Across the four time points, when sexually-inactive participants are included, the overall missing percentages range from 18% to 26% for PROMIS SexFS measures and from 20% to 32% for PCSI subscales. When excluding sexually-inactive participants, missing percentages all drop to values below 1%. A large number of missing data are present among sexually-inactive men due to inapplicability of the sex-related questions; and, thus, skipped out of answering the questions in the survey. To avoid potential bias of including large missing values and to ensure interpretability of the final results, we will restrict our subsequent analyses to sexually-active men. In addition, we will perform listwise deletion given that missing values are completely at random among the sexually-active men²³, and removal of incomplete cases results in less than 1% loss of data.

Modeling Assumptions

As shown in Table 3, none of the NC ProCESS versions of the PROMIS SexFS measures (a mixture of SexFS-V1, SexFS-V2, and candidate items that did not make it in either version) fully met the three modeling assumptions – model fit indices in general are far from satisfactory, internal consistency of PROMIS Orgasm items is particularly low, several

domains have one or two items with extremely low or even negative factor loadings, and all domains show signs of locally dependent pairs and/or non-monotonically increasing items. The findings regarding some of the poorly-fitting items are consistent with past research¹³ which also excluded those items from SexFS-V2 due to their unsatisfactory psychometric properties.

For PROMIS SexFS domains of Interest in Sexual Activity, Erectile Function, and Satisfaction with Sex Life, we developed SexFS-Preferred versions consisting of well performing items. Also consistent with¹³, no set of well performing items for the Orgasm domain could be identified; thus, the Orgasm domain was excluded from subsequent analyses. Model selection procedures are implemented using a top-down approach. The unidimensional model of a given domain is pruned by removing one offending item at a time until all three assumptions are satisfied. As shown in Table 3, the resulting new versions (SexFS-Preferred) are highly internally consistent, free of local dependence, and fully monotonic, even though their model fit indices are unavailable/meaningless because they are just-identified models (3 items per domain) with zero degrees of freedom.

In addition, the two PCSI sexual functioning subscales met the three assumptions (except one MOS Sexual Problem item which has factor loadings slightly below 0.775 at 3- and 12-month), and therefore they are used in the upcoming multidimensional models without any modifications. In addition, both PCSI subscales showed adequate reliability (all Cronbach alpha estimates reported in Table 3 were above 0.80).

Convergent and Discriminant Validities

Table 4 shows the between-factor correlations estimated at each time point using single-group multidimensional IRT models. High convergent validity is reflected in the large correlations between the PROMIS and PCSI Erectile Function measures (correlations ranged from 0.84 to 0.95 over four assessment time points), as well as the moderate to large correlations between all sex-related domains except for PROMIS Sexual Interest. As expected, the non-sex-related domains (Fatigue and Physical Functioning) correlated weakly with the sex-related domains, which is an indication of high discriminant validity.

DIF

Based on the previous analyses, only the three PROMIS SexFS-Preferred measures are tested for DIF. In terms of BG-DIF, significant DIF effects ($p < .05$) are detected between Surgery and No-Surgery groups for SEXFCN3 at 3-Month, SEXFCN1 and SEXFCN3 at 24-Month, and GLOBSAT2 at 24-Month. Signed, unsigned, and standardized effect size measures for the differentially functioning items are presented in Table 5. All items show nonuniform DIF effects except SEXFCN1 at 24-Month which shows uniform DIF (the surgery group is more likely to endorse higher categories). Based on the expected score standardized difference (ESSD) measure⁴⁹, all the BG-DIF effects are quite small. Nonetheless, for our purpose of mean comparisons, it is safer to link the metric using only DIF-free items. The underlying reasons for observing these DIF effects are beyond the scope of the current study and therefore are not investigated. In addition, all items are invariant across time (no L-DIF). Hence, for the upcoming longitudinal two-group model below, a

measure will be linked longitudinally and cross-sectionally by equating only items that show no BG-DIF.

Responsiveness of Measures over Time

Factor means are presented in Figure 2, and the results are as expected. Factor mean for Erectile Function and Sexual Satisfaction of the Surgery group is always lower than that of the No-Surgery group at 3-Month assessment.

For PROMIS Sexual Interest, at baseline, the Surgery group reported statistically significantly lower sexual interest than the No-Surgery group (95% C.I. for surgery group = $[-0.41, -0.13]$; see the top plot in Figure 2). The two groups are similar in terms of factor mean change over time, since prostatectomy is not expected to have a strong impact on sexual interest. Both groups show a noticeable decline of factor mean at 3-Month, and then they stabilize at about half a standard deviation below the reference value (i.e., zero).

For PROMIS Erectile Function, even though factor mean of the Surgery group falls more drastically at 3-Month in contrast to the No-Surgery group likely due to the impact of surgery, the difference between the two means gradually narrows at 12- and 24-Month as the Surgery group recovers over time.

For PROMIS Sexual Satisfaction, the results are overall similar to those of PROMIS Erectile Function. Nonetheless, the impact of prostatectomy on sexual satisfaction is not as salient as its impact on erectile function at 3-Month. In addition, at baseline, the Surgery group showed statistically significantly lower sexual satisfaction than the No-Surgery group (95% C.I. for surgery group = $[-0.26, -0.04]$).

Discussion

The timing of this study provided an early spotlight on how well the items within select domains of the PROMIS SexFS profile of measures perform for assessing sexual functioning and satisfaction within a cohort of men with localized prostate cancer. This study included a mixed set of items from an early version of PROMIS SexFS (version 1) and items being considered for the next version of PROMIS SexFS measure. It was not expected that all items would perform well because the included items were still under investigation by PROMIS investigators.

Applying the criteria that all items within a SexFS domain must fully meet the three IRT model assumptions of unidimensionality, local independence and monotonicity, we were able to find three well-performing items in each of the domains of Sexual Interest, Erectile Function, and Sex Satisfaction. A few items were found to have DIF between Surgery and non-Surgery groups, but the items did not have longitudinal DIF. Subsequently, these study-specific PROMIS SexFS-Preferred measures performed well in the evaluation of the reliability and validity of measures including evaluation of convergent and discriminant validity, and the responsiveness of the measures over time.

All items included in the PROMIS SexFS-Preferred measures were originally on the PROMIS SexFS Version 1 measure, but not all items on the PROMIS SexFS Version 1 were

selected for the PROMIS SexFS-Preferred measure. PROMIS SexFS Version 1 items that did not perform well in our study were also found to be problematic by Weinfurt et al when creating version 2¹³. All 3 items in each of the PROMIS SexFS-Preferred Sex Satisfaction and SexFS-Preferred Erectile Function domains were included on version 2 of the PROMIS SexFS measures¹³. Two of the three items in the PROMIS SexFS Sex Interest domain was included in the respective PROMIS SexFS version 2 domain (note that Version 2 only includes these 2 items). These study results in men with localized prostate cancer provide further psychometric evidence for the reliability and validity of the PROMIS SexFS Version 2 measures.

The items for the PROMIS SexFS Orgasm domain did not perform well and could not be used in subsequent analysis of the validity of the measure. This finding is consistent with psychometric work conducted by Weinfurt et al.¹³, who recommended that additional work needs to be performed on optimal approaches to assess the orgasm domain.

A unique aspect of this study is that all men completed the surveys over the phone with an interviewer, in contrast to most previous evaluations of PROMIS measures that ask participants to complete the questionnaires privately by computer. A couple of previous studies⁵⁰ supported the measurement invariance of the PROMIS measures across computers, personal digital assistants, paper-pencil, and interactive voice response assessment modes, but not with a live interviewer. Completing questionnaires with an interviewer over the phone is a different experience than independently completing a questionnaire on a computer. Previous research on the effect of mode of administration between computer and phone interviewer within the same cohort in this study found men were more likely to report better erectile functioning to the phone interviewer than on the PC⁵¹. The same study did not find invariance across phone interviewer and PC for the Sex Interest and Sex Satisfaction domains. The NC ProCESS selected the phone interviewer format to be inclusive of those men who may be too illiterate to read the survey.

This study was limited to men living in North Carolina who could self-report their HRQOL in English. Thus, there is concern about the generalizability of the study findings; however, this study was able to include more men from vulnerable populations by allowing the men to self-report over the phone. We were unable to fully evaluate the psychometric properties of all items in version 2 of the PROMIS SexFS as we did not have access at the time to all the items that were included after we launched our study.

This study provided psychometric evidence for use of the PROMIS SexFS (version 2) measures in men with localized prostate cancer to assess sexual interest, erectile function and sex satisfaction over time. Starting with a larger set of SexFS items from Version 1 of the measure and newly developed items, our evaluation selected a subset of items that performed well psychometrically. A better understanding of how prostate cancer and its treatments impacts the lives of men will inform the identification of interventions to treat dysfunction and inform the development of decision aids for men to make better decisions on their treatment choices.

Acknowledgments

Funding: This research was supported by grants from the Agency for Healthcare Research and Quality (HHS29020050040ITO6) and the National Cancer Institute (R01CA174453).

References

1. American Cancer Society. Cancer Facts & Figures 2018 Atlanta: American Cancer Society; 2018.
2. Hamdy FC, Donovan JL, Lane JA, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *N Engl J Med* 2016;375(15): 1415–24. [PubMed: 27626136]
3. Broughman JR, Basak R, Nielsen ME, et al. Prostate Cancer Patient Characteristics Associated With a Strong Preference to Preserve Sexual Function and Receipt of Active Surveillance. *J Natl Cancer Inst* 2018;110(4): 420–25. [PubMed: 29045679]
4. Stanford JL, Feng Z, Hamilton AS, et al. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study. *JAMA* 2000;283(3): 354–60. [PubMed: 10647798]
5. Briganti A, Capitanio U, Chun FK, et al. Prediction of sexual function after radical prostatectomy. *Cancer* 2009;115(13 Suppl): 3150–9. [PubMed: 19544544]
6. Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. *N Engl J Med* 2013;368(5): 436–45. [PubMed: 23363497]
7. Chen RC, Basak R, Meyer AM, et al. Association Between Choice of Radical Prostatectomy, External Beam Radiotherapy, Brachytherapy, or Active Surveillance and Patient-Reported Quality of Life Among Men With Localized Prostate Cancer. *JAMA* 2017;317(11): 1141–50. [PubMed: 28324092]
8. Clark J, Talcott J. Prostate Cancer Symptom Indexes and Symptom Distress Scales (PCSISDS) Galway, Ireland: Measurement Instrument Database for the Social Sciences (MIDSS); 2005.
9. Wei JT, Dunn RL, Litwin MS, et al. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. *Urology* 2000;56(6): 899–905. [PubMed: 11113727]
10. Cella D, Yount S, Rothrock N, 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;45(5 Suppl 1): S3–s11.
11. Reeve BB, Hays RD, Bjorner JB, 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;45(5 Suppl 1): S22–31. [PubMed: 17443115]
12. Flynn KE, Lin L, Cyranowski JM, et al. Development of the NIH PROMIS (R) Sexual Function and Satisfaction measures in patients with cancer. *J Sex Med* 2013;10 Suppl 1: 43–52. [PubMed: 23387911]
13. Weinfurt KP, Lin L, Bruner DW, et al. Development and Initial Validation of the PROMIS((R)) Sexual Function and Satisfaction Measures Version 2.0. *J Sex Med* 2015;12(9): 1961–74. [PubMed: 26346418]
14. Chen RC, Carpenter WR, Kim M, et al. Design of the North Carolina Prostate Cancer Comparative Effectiveness and Survivorship Study (NC ProCESS). *Journal of comparative effectiveness research* 2015;4(1): 3–9. [PubMed: 25565065]
15. Quach CW, Langer MM, Chen RC, et al. Reliability and validity of PROMIS measures administered by telephone interview in a longitudinal localized prostate cancer study. *Qual Life Res* 2016;25(11): 2811–23. [PubMed: 27240448]
16. Fortune-Greeley AK, Flynn KE, Jeffery DD, et al. Using cognitive interviews to evaluate items for measuring sexual functioning across cancer populations: improvements and remaining challenges. *Qual Life Res* 2009;18(8): 1085–93. [PubMed: 19672697]
17. Flynn KE, Jeffery DD, Keefe FJ, et al. Sexual functioning along the cancer continuum: focus group results from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)). *Psychooncology* 2011;20(4): 378–86. [PubMed: 20878833]

18. Flynn KE, Reeve BB, Lin L, et al. Construct validity of the PROMIS(R) sexual function and satisfaction measures in patients with cancer. *Health and quality of life outcomes* 2013;11: 40. [PubMed: 23497200]
19. Weinfurt KP, Lin L, Dombek CB, et al. Accuracy of 30-day recall for components of sexual function and the moderating effects of gender and mood. *J Sex Med* 2014;11(3): 678–96. [PubMed: 23802907]
20. Muthén B, du Toit SHC, Spisic D. Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. (Unpublished technical report) Unpublished technical report; 1997.
21. Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika* 1981;46(4): 443–59.
22. Cai L High-dimensional exploratory item factor analysis by a Metropolis–Hastings Robbins–Monro Algorithm. *Psychometrika* 2010;75(1): 33–57.
23. Little RJA. A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association* 1988;83(404): 1198–202.
24. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001;29(4): 1165–88.
25. Samejima F Estimation of latent ability using a response pattern of graded scores. *ETS Research Bulletin Series* 1968;1968(1): i–169.
26. Sijtsma K, van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *Br J Math Stat Psychol* 2017;70(1): 137–58. [PubMed: 27958642]
27. MacCallum RC, Widaman KF, Zhang SB, et al. Sample size in factor analysis. *Psychol Methods* 1999;4(1): 84–99.
28. MacCallum RC, Widaman KF, Preacher KJ, et al. Sample size in factor analysis: The role of model error. *Multivariate behavioral research* 2001;36(4): 611–37. [PubMed: 26822184]
29. Chen RC, Clark JA, Talcott JA. Individualizing quality-of-life outcomes reporting: how localized prostate cancer treatments affect patients with different levels of baseline urinary, bowel, and sexual function. *J Clin Oncol* 2009;27(24): 3916–22. [PubMed: 19620493]
30. Hu LT, Bentler PM. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Struct Equ Modeling* 1999;6(1): 1–55.
31. Yu C- Y. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes Los Angeles: University of California, Los Angeles; 2002:168.
32. Mislevy JL, Rupp AA, Harring JR. Detecting Local Item Dependence in Polytomous Adaptive Data. *Journal of Educational Measurement* 2012;49(2): 127–47.
33. Sijtsma K, Molenaar IW. Mokken Models. In: Van der Linden WJ, ed. *Handbook of Item Response Theory Models* Boca Raton, FL: CRC Press, Taylor & Francis Group; 2016:303–21.
34. Van Der Ark LA. Mokken scale analysis in R. *Journal of Statistical Software* 2007;20(11): 1–19.
35. Clark JA, Talcott JA. Symptom indexes to assess outcomes of treatment for early prostate cancer. *Med Care* 2001;39(10): 1118–30. [PubMed: 11567174]
36. Hill CD. *Two models for longitudinal item response data.* Psychology and Neuroscience Chapel Hill, NC: The University of North Carolina at Chapel Hill; 2006:91.
37. Thissen D, Steinberg L, Wainer H. Detection of different item functioning using the parameters of item response models. In: Holland PW, Wainer H, eds. *Differential item functioning* Hillsdale, N.J.: Lawrence Erlbaum; 1993:67–113.
38. Meade AW, Wright NA. Solving the measurement invariance anchor item problem in item response theory. *J Appl Psychol* 2012;97(5): 1016–31. [PubMed: 22468848]
39. Thissen D, Steinberg L, Wainer H. Use of item response theory in the study of group differences in trace lines. In: Wainer H, Braun HI, eds. *Test validity* New York; London: Routledge/Taylor & Francis Group; 1988:147–69.
40. Litwin MS, Gore JL, Kwan L, et al. Quality of life after surgery, external beam irradiation, or brachytherapy for early-stage prostate cancer. *Cancer* 2007;109(11): 2239–47. [PubMed: 17455209]

41. Pardo Y, Guedea F, Aguiló F, et al. Quality-of-life impact of primary treatments for localized prostate cancer in patients without hormonal treatment. *J Clin Oncol* 2010;28(31): 4687–96. [PubMed: 20921463]
42. Perez MA, Meyerowitz BE, Lieskovsky G, et al. Quality of life and sexuality following radical prostatectomy in patients with prostate cancer who use or do not use erectile aids. *Urology* 1997;50(5): 740–6. [PubMed: 9372885]
43. Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate- cancer survivors. *N Engl J Med* 2008;358(12): 1250–61. [PubMed: 18354103]
44. R. Core Team. R: Language and Environment for Statistical Computing Vienna, Austria: R Foundation; 2017.
45. Beaujean AA. BaylorEdPsych: R Package for Baylor University Educational Psychology Quantitative Courses Waco, TX: Baylor University; 2012.
46. Chalmers RP. mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software* 2012;48(6): 1–29.
47. Rosseel Y lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 2012;48(2): 1–36.
48. Wickham H Ggplot2 : elegant graphics for data analysis. Dordrecht; New York: Springer; 2009.
49. Meade AW. A taxonomy of effect size measures for the differential functioning of items and scales. *J Appl Psychol* 2010;95(4): 728–43
50. Bjorner JB, Rose M, Gandek B, et al. Difference in method of administration did not significantly impact item response: an IRT-based analysis from the Patient-Reported Outcomes Measurement Information System (PROMIS) initiative. *Qual Life Res* 2014;23(1): 217–27. [PubMed: 23877585]
51. Wang M, Chen RC, Usinger DS, et al. Evaluating measurement invariance across assessment modes of phone interview and computer self-administered survey for the PROMIS measures in a population-based cohort of localized prostate cancer survivors. *Qual Life Res* 2017;26(11): 2973–85. [PubMed: 28681335]

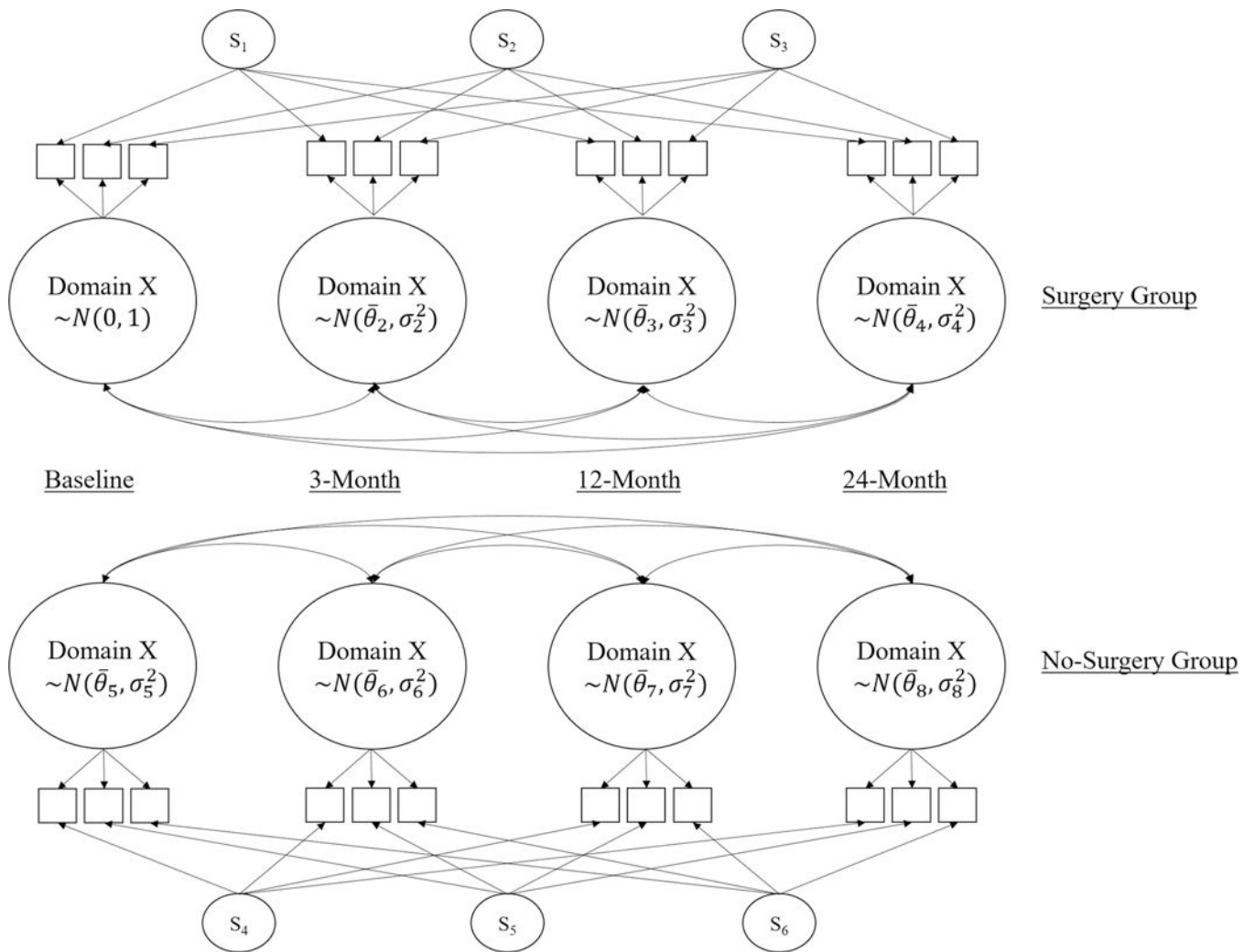


Figure 1.
 Example Path Diagram for a Longitudinal Two-Group IRT Model.
 $\bar{\theta}_*$ = latent mean estimate
 σ_*^2 = latent variance estimate
 S_* = specific factors (fixed to standard normal) that capture item autocorrelations.

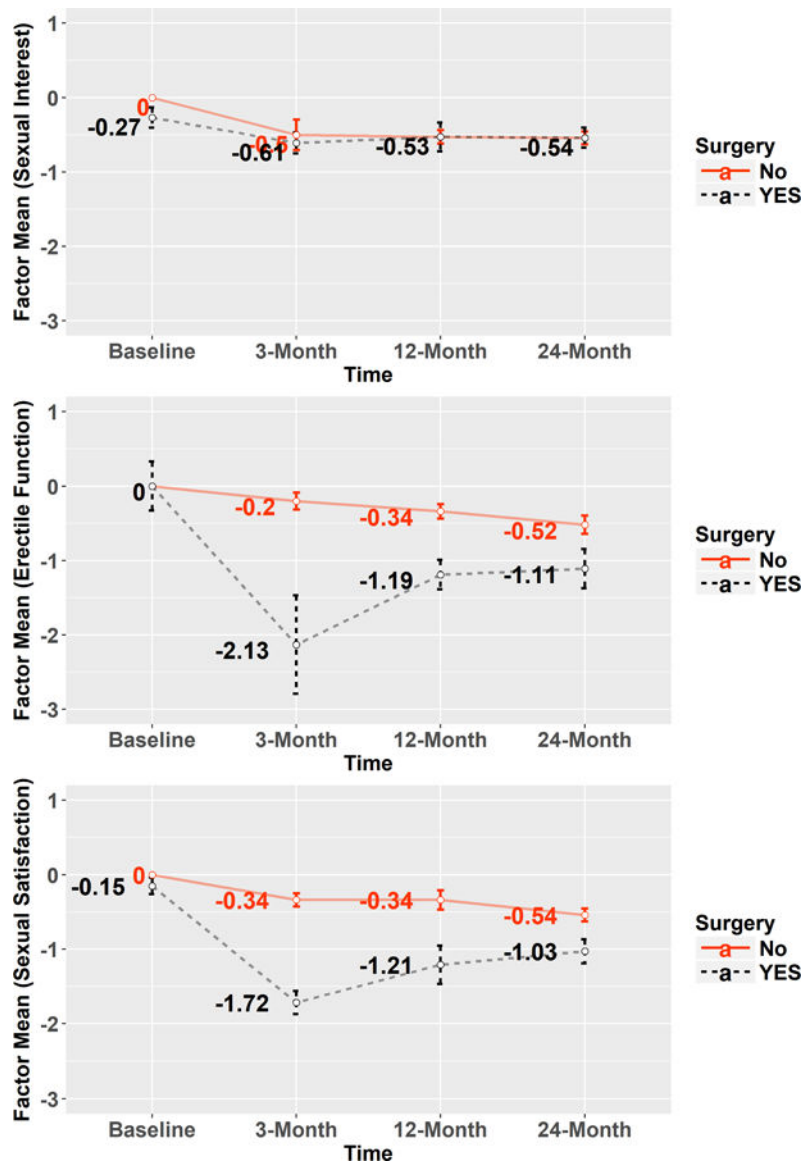


Figure 2. Changes of Factor Means (by Surgery Groups) Over Time for Three PROMIS SexFS Domains. Within each domain, the metric is defined by fixing the no-surgery group at baseline as standard normal. Vertical bars represent 95% confidence intervals for the corresponding mean estimates.

Table 1.

Question Stems and Response Options for Evaluated Items.

Domain	Item ID (as appeared on NC ProCESS surveys)	Item Stem (as appeared on NC ProCESS surveys)	Item Categories
PROMIS Sexual Interest	SEXFCN1 + ≠ \$	How interested have you been in sexual activity?	1) Not at all 2) A little bit 3) Somewhat 4) Quite a bit 5) Very
	SEXFCN2 + \$	How often were you interested enough to start a sexual activity if you could have?	1) Never 2) Rarely
	SEXFCN3 + ≠ \$	How often have you felt like you wanted to have sexual activity?	3) Sometimes 4) Often
	SEXFCN4 +	How often have you had sexual thoughts or fantasies while you were awake?	5) Always
	SEXFCN5	How often has your level of Sexual Interest bothered you?	1) Not at all 2) A little bit 3) Somewhat 4) Quite a bit 5) Very much
PROMIS Erectile Function	ERECFCN1 + ≠ \$	Please rate your ability to have an erection or get hard.	1) Very Good 2) Good 3) Fair 4) Poor 5) Very Poor
	ERECFCN2 ≠	How hard have your erections been?	6) You have not tried to get an erection or get hard in the past 30 days 1) Penis was completely hard 2) Penis was hard enough for penetration but not completely hard 3) Penis was hard but not hard enough for Penetration 4) Penis was larger but not hard 5) You have not had an erection in the past 30 days
	ERECFCN3 + ≠ \$	How difficult has it been for you to get an erection or get hard when you wanted to?	1) Not at all 2) A little bit
	ERECFCN4 + ≠ \$	How difficult has it been to keep an erection or stay hard when you wanted to?	3) Somewhat 4) Quite a bit 5) Very

Domain	Item ID (as appeared on NC ProCESS surveys)	Item Stem (as appeared on NC ProCESS surveys)	Item Categories
			6) Have not tried to get an erection in the past 30 days
	ERECFCN5 +	How often have your erections been physically uncomfortable or painful?	1) Never 2) Rarely
	ERECFCN6	How often have you been able to get an erection or get hard when you wanted to?	3) Sometimes 4) Often
	ERECFCN7 ≠	How often have you been able to keep an erection or stay hard as long as you wanted to?	5) Always 6) Have not tried to get an erection in the past 30 days
PROMIS Orgasm	ORG1 +	How would you rate your ability to have a satisfying orgasm/climax?	1) Excellent 2) Very Good 3) Good 4) Fair 5) Poor 6) You have not tried to have an orgasm or climax in the past 30 days
	ORG2 +	How often have you had an orgasm or climax more quickly than you would like?	1) Never 2) Rarely
	ORG3 +	How often have you had pain and/or burning during or after ejaculation?	3) Sometimes 4) Often
	ORG4 ≠	How often have you been able to have an orgasm/climax when you wanted to?	5) Always 6) You have not tried to ejaculate in the past 30 days
PROMIS Sexual Satisfaction	GLOBSATI + ≠ \$	How satisfied have you been with your sex life?	1) Not at all 2) A little bit 3) Somewhat 4) Quite a bit 5) Very
	GLOBSAT2 + ≠ \$	How much pleasure has your sex life given you?	1) None 2) A little bit 3) Somewhat 4) Quite a bit 5) A lot
	GLOBSAT3 + ≠ \$	How satisfied have you been with your sexual relationship or relationships?	1) Not at all 2) A little bit 3) Somewhat 4) Quite a bit 5) Very 6) You have not had sexual relationships

Domain	Item ID (as appeared on NC ProCESS surveys)	Item Stem (as appeared on NC ProCESS surveys)	Item Categories
	GLOBSAT4 [†] ‡	When you have had sexual activity, how much have you enjoyed it?	<ol style="list-style-type: none"> 1) Not at all 2) A little bit 3) Somewhat 4) Quite a bit 5) Very much 6) Have not had sexual activity in the past 30 days
	GLOBSAT5 [†]	When you have had sexual activity, how satisfying has it been?	
PCSI Erectile Function (Clark, 2005)	E10	What is the most erect or hard your penis has become at any time?	<ol style="list-style-type: none"> 1) A full erection 2) A nearly full erection sufficient for penetration without manual assistance 3) A partial erection capable of penetration with manual assistance 4) A partial erection not capable of penetration even with manual assistance 5) You've had no erection at all in the past 4 weeks
	E30	How much difficulty have you had getting an erection during sexual activity?	<ol style="list-style-type: none"> 1) No difficulty 2) A little difficulty 3) Some difficulty 4) A lot of difficulty 5) You have not had sexual activity in the past 4 weeks
	E40	How much difficulty have you had keeping an erection during sexual activity?	
PCSI MOS Sexual Problems (Clark, 2005)	E140	How big a problem, if any, has each of the following been for you? (Your level of sexual desire?)	<ol style="list-style-type: none"> 1) No problem 2) Very small problem 3) Small problem
	E150	How big a problem, if any, has each of the following been for you? (Your ability to have or keep an erection?)	<ol style="list-style-type: none"> 1) No problem 2) Very small problem 3) Small problem 4) Moderate problem 5) Big problem
	E160	How big a problem, if any, has each of the following been for you? (Your ability to reach orgasm or climax?)	

Notes: PROMIS items have a 30-day recall period (i.e., each question starts with "In the past 30 days...").

PCSI items have a 4-week recall period (i.e., each question starts with "In the past 4 weeks...").

PCSI Erectile Function is derived from the Sexual Dysfunction subscale (Clark, 2005).

PCSI MOS Sexual Problems is missing two items from its original subscale (Clark, 2005).

[†] Items from PROMIS SexFS-V1 (Flynn et al., 2013). The five underlined items are worded slightly differently than in PROMIS SexFS-V1.

~~‡~~ Items from PROMIS SexFS-V2 (Weinfurt et al., 2015).

Items in PROMIS SexFS-Preferred (best-fitting items in the current study; see Table 3 for model and item fit).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Sample Characteristics at Each Wave of Assessment.

Sample Characteristics at Baseline (unless noted otherwise)	Baseline	3-Month	12-Month	24-Month
<i>N (%) out of total</i>				
Total (within each wave)	1449 (100%)	1164 (100%)	1083 (100%)	964 (100%)
Responded to PROMIS SexFS domains¹ (within each wave)	332 (22.9%)	408 (35.1%)	768 (70.9%)	939 (97.4%)
<i>N (%) out of those who responded to PROMIS SexFS domains</i>				
Seniority (time-specific)				
< Age 65	168 (50.6%)	191 (46.8%)	319 (41.5%)	346 (36.8%)
Age 65	164 (49.4%)	217 (53.2%)	449 (58.5%)	593 (63.2%)
Race				
White	235 (70.8%)	290 (71.1%)	550 (71.6%)	678 (72.2%)
Black	89 (26.8%)	109 (26.7%)	200 (26.0%)	229 (24.4%)
Asian or Pacific Islander	0	1 (0.2%)	1 (0.1%)	2 (0.2%)
American Indian or Alaskan Native	8 (2.4%)	8 (2.0%)	15 (2.0%)	14 (1.5%)
Other & Unknown	0	0	2 (0.3%)	16 (1.7%)
Ethnicity				
Non-Hispanic	326 (98.2%)	402 (98.5%)	757 (98.6%)	914 (97.3%)
Hispanic	4 (1.2%)	3 (0.7%)	9 (1.2%)	10 (1.1%)
Unknown	2 (0.6%)	3 (0.7%)	2 (0.3%)	15 (1.6%)
Highest level of education				
Eighth grade or less	18 (5.4%)	17 (4.2%)	23 (3.0%)	22 (2.3%)
Some high school	18 (5.4%)	29 (7.1%)	59 (7.7%)	63 (6.7%)
High school graduate	64 (19.3%)	87 (21.3%)	164 (21.4%)	200 (21.3%)
Some college	100 (30.1%)	114 (27.9%)	222 (28.9%)	264 (28.1%)
College graduate	132 (39.8%)	161 (39.5%)	300 (39.1%)	379 (40.4%)
Unknown	0	0	0	11 (1.2%)
Marital status				
Married	265 (79.8%)	324 (79.4%)	625 (81.4%)	753 (80.2%)
Divorced	35 (10.5%)	38 (9.3%)	70 (9.1%)	74 (7.9%)
Widowed	11(3.3%)	18(4.4%)	31(4.0%)	47 (5.0%)
Never married	16(4.8%)	19(4.7%)	28(3.6%)	35 (3.7%)
Separated	5 (1.5%)	9 (2.2%)	14(1.8%)	18 (1.9%)
Other & Unknown	0	0	0	12 (1.3%)
Employment status				
Employed full time	120 (36.1%)	143 (35.0%)	259 (33.7%)	319 (34.0%)
Employed part time	32(9.6%)	34(8.3%)	68(8.9%)	77 (8.2%)
Unemployed	14(4.2%)	16(3.9%)	26(3.4%)	32 (3.4%)

Sample Characteristics at Baseline (unless noted otherwise)	Baseline	3-Month	12-Month	24-Month
Retired	135 (40.7%)	178 (43.6%)	356 (46.4%)	440 (46.9%)
Disabled and not working	31(9.3%)	37(9.1%)	59(7.7%)	60 (6.4%)
Unknown	0	0	0	11 (1.2%)
Income				
Less than \$10,000	15(4.5%)	20(4.9%)	40(5.2%)	39 (4.2%)
\$10,000 to \$20,000	36(10.8%)	43(10.5%)	68(8.9%)	80 (8.5%)
\$20,001 to \$40,000	75(22.6%)	85(20.8%)	162 (21.1%)	198 (21.1%)
\$40,001 to \$70,000	97(29.2%)	115 (28.2%)	219 (28.5%)	262 (27.9%)
\$70,001 to \$90,000	35(10.5%)	50(12.3%)	99(12.9%)	118 (12.6%)
more than \$90,000	60(18.1%)	79(19.4%)	153 (19.9%)	198 (21.1%)
Unknown	14(4.2%)	16(3.9%)	27(3.5%)	44 (4.7%)
Gleason score – indicator of the aggressiveness of prostate cancer				
< 7 (slowly growing cells)	177 (53.3%)	230 (56.4%)	447 (58.2%)	537 (57.2%)
= 7 (intermediate risk)	127 (38.3%)	141 (34.6%)	521 (67.8%)	321 (34.2%)
> 7 (high grade)	28(8.4%)	37(9.1%)	70(9.1%)	81 (8.6%)
Treatments² (post-baseline)				
Radiation		119 (29.2%)	234 (30.5%)	285 (30.4%)
Hormone with Radiation		28(6.9%)	65(8.5%)	78 (8.3%)
Prostatectomy		152 (37.3%)	278 (36.2%)	380 (40.5%)
Other treatments		14(3.4%)	27(3.5%)	29 (3.1%)
No treatment (e.g., active surveillance, watchful waiting, and/or supplements)		119 (29.2%)	228 (29.7%)	252 (26.8%)
Unknown		6 (1.5%)	8 (1.0%)	7 (0.7%)

Note: The columns are not mutually exclusive men. The Table shows the sample size available at each assessment point.

¹Including cases with at least one response on Sexual Interest, Satisfaction with Sex Life, Orgasm, and/or Erectile Function domains of PROMIS. Therapeutic Aids is excluded because no psychometric analysis is conducted using this domain.

²A small proportion of patients received multiple treatments.

Table 3.

Model and Item Fit of PROMIS and PCSISDS Sex-Related Measures.

Domain – Version	Time	Descriptive Statistics	Model Fit	Item ID	λ	LD Grouping Labels (Time-Specific)	# of Significant Monotonicity Violations
PROMIS Sexual Interest - NC ProCESS	Baseline	N = 329 Alpha = 0.812	$\chi^2_5 = 36.46, p < 0.001$ RMSEA = 0.138 CFI = 0.996 WRMR = 0.591	SEXFCN1 $\neq \neq \neq$	0.923	i, ii	None
				SEXFCN2 $\neq \neq$	0.979	iii	None
				SEXFCN3 $\neq \neq \neq$	0.920	iv	None
				SEXFCN4 \neq	0.735	i, iv	None
				SEXFCN5	0.121	ii, in	1
	3-Month	N = 397 Alpha = 0.738	$\chi^2_5 = 56.87, p < 0.001$ RMSEA = 0.162 CFI = 0.992 WRMR = 0.881	SEXFCN1 $\neq \neq \neq$	0.942	i	None
				SEXFCN2 $\neq \neq$	0.927	ii	None
				SEXFCN3 $\neq \neq \neq$	0.901	lii	None
				SEXFCN4 \neq	0.768	lii	None
				SEXFCN5	-0.124	i, ii	6
	12-Month	N = 759 Alpha = 0.757	$\chi^2_5 = 100.73, p < 0.001$ RMSEA = 0.159 CFI = 0.995 WRMR = 1.054	SEXFCN1 $\neq \neq \neq$	0.930	i	None
				SEXFCN2 $\neq \neq$	0.957	i	None
				SEXFCN3 $\neq \neq \neq$	0.938	ii, in	None
				SEXFCN4 \neq	0.775	ii	None
				SEXFCN5	-0.114	i, in	30
	24-Month	N = 928 Alpha = 0.772	$\chi^2_5 = 62.42, p < 0.001$ RMSEA = 0.111 CFI = 0.996 WRMR = 0.881	SEXFCN1 $\neq \neq \neq$	0.914	i	None
				SEXFCN2 $\neq \neq$	0.959	ii	None
				SEXFCN3 $\neq \neq \neq$	0.917	lii	1
				SEXFCN4 \neq	0.756	lii	None
				SEXFCN5	-0.054	i, ii	18
PROMIS Sexual Interest — SexFS-Preferretl	Baseline	N = 330 Alpha = 0.918	Just-identified	SEXFCN1 $\neq \neq \neq$	0.907	None	None
				SEXFCN2 $\neq \neq$	0.998	None	None
				SEXFCN3 $\neq \neq \neq$	0.908	None	None
	3-Month	N = 400 Alpha = 0.897	Just-identified	SEXFCN1 $\neq \neq \neq$	0.945	None	None
				SEXFCN2 $\neq \neq$	0.933	None	None
				SEXFCN3 $\neq \neq \neq$	0.885	None	None
	12-Month	N = 760 Alpha = 0.923	Just-identified	SEXFCN1 $\neq \neq \neq$	0.928	None	None
				SEXFCN2 $\neq \neq$	0.966	None	None

Domain – Version	Time	Descriptive Statistics	Model Fit	Item ID	λ	LD Grouping Labels (Time-Specific)	# of Significant Monotonicity Violations
PROMIS Erectile Function - NC ProCESS	24-Month	N = 932 Alpha = 0.909	Just-identified	SEXFCN3 \neq	0.926	None	None
				SEXFCN1 \neq	0.912	None	None
				SEXFCN2 \neq	0.964	None	None
				SEXFCN3 \neq	0.906	None	None
	Baseline	N = 245 Alpha = 0.590	$\chi^2_{14} = 5.73, p < 0.001$ RMSEA = 0.173 CFI = 0.981 WRMR = 0.980	ERECFCN1 \neq	0.917	None	None
				ERECFCN2 \neq	0.889	None	None
				ERECFCN3 \neq	0.915	i. il	None
				ERECFCN4 \neq	0.899	i. in	None
				ERECFCN5 \neq	0.157	ii	None
				ERECFCN6	0.840	iv	None
				ERECFCN7 \neq	0.820	iii. iv	None
	3-Month	N = 243 Alpha = 0.900	$\chi^2_{14} = 162.26, p < 0.001$ RMSEA = 0.209 CFI = 0.979 WRMR = 1.573	ERECFCN1 \neq	0.879	None	None
ERECFCN2 \neq				0.890	None	None	
ERECFCN3 \neq				0.957	i	None	
ERECFCN4 \neq				0.925	i	None	
ERECFCN5 \neq				0.310	i	None	
ERECFCN6				0.819	ii	None	
ERECFCN7 \neq				0.839	ii	None	
12-Month	N = 495 Alpha = 0.900	$\chi^2_{14} = 320.32, p < 0.001$ RMSEA = 0.210 CFI = 0.978 WRMR = 1.711	ERECFCN1 \neq	0.900	i	None	
			ERECFCN2 \neq	0.886	None	None	
			ERECFCN5 \neq	0.922	ii. in	None	
			ERECFCN4 \neq	0.935	ii	None	
			ERECFCN5 \neq	0.151	i. iii	None	
			ERECFCN6	0.808	iv	None	
			ERECFCN7 \neq	0.882	iv	None	
24-Month	N = 591 Alpha = 0.891	$\chi^2_{14} = 402.42, p < 0.001$ RMSEA = 0.217 CFI = 0.977 WRMR = 2.231	ERECFCN1 \neq	0.902	i	None	
			ERECFCN2 \neq	0.862	i	None	
			ERECFCN3 \neq	0.929	ii, in	None	
			ERECFCN4 \neq	0.937	ii	None	
			ERECFCN5 \neq	0.091	iii	2	

Domain – Version	Time	Descriptive Statistics	Model Fit	Item ID	λ	LD Grouping Labels (Time-Specific)	# of Significant Monotonicity Violations			
PROMIS Erectile Function - SexFS-Preferretl	Baseline	N = 247 Alpha = 0.595	Just-identified	ERECFCN6	0.787	iv	None			
				ERECFCN7 \neq	0.842	iv	None			
	3-Month	N = 246 Alpha = 0.903	Just-identified	ERECFCN1 \neq	0.892	None	None			
				ERECFCN3 \neq	0.946	None	None			
				ERECFCN4 \neq	0.911	None	None			
				ERECFCN1 \neq	0.821	None	None			
				ERECFCN3 \neq	0.999	None	None			
				ERECFCN4 \neq	0.936	None	None			
				12-Month	N = 499 Alpha = 0.910	Just-identified	ERECFCN1 \neq	0.862	None	None
							ERECFCN \neq	0.955	None	None
							ERECFCN4 \neq	0.954	None	None
				24-Month	N = 595 Alpha = 0.914	Just-identified	ERECFCN1 \neq	0.852	None	None
							ERECFCN5 \neq	0.967	None	None
							ERECFCN4 \neq	0.957	None	None
PROMIS Orgasm - NC ProCESS	Baseline	N = 233 Alpha = 0.371	$\chi^2_2 = 4.02, p = 0.134$ RM SEA = 0.066 CFI = 0.993 WRMR = 0.370	ORG1 \neq	0.984	None	None			
				ORG2 \neq	-0.019	i	None			
				ORG3 \neq	0.051	i	None			
				ORG4 \neq	0.690	None	None			
	3-Month	N = 227 Alpha = 0.267	$\chi^2_2 = 3.06, p = 0.217$ RM SEA = 0.048 CFI = 0.997 WRMR = 0.304	ORG1 \neq	0.865	i	None			
				ORG2 \neq	-0.37	i, li	1			
				ORG3 \neq	-0.064	None	None			
				ORG4 \neq	0.861	ii	None			
	12-Month	N = 475 Alpha = 0.325	$\chi^2_2 = 1609, p < 0.001$ RMSEA = 0.122 CFI = 0.991 WRMR = 0.736	The IRT model did not converge after 5000 iterations.						
	24-Month	N = 583 Alpha = 0.326	$\chi^2_2 = 17.28, p < 0.001$ RMSEA = 0.115 CFI = 0.987 WRMR = 0.752	ORG1 \neq	0.989	i	1			
				ORG2 \neq	-0.252	i, ii, lii	12			
				ORG3 \neq	-0.042	ii	2			
				ORG4 \neq	0.761	iii	3			

Domain – Version	Time	Descriptive Statistics	Model Fit	Item ID	λ	LD Grouping Labels (Time-Specific)	# of Significant Monotonicity Violations
PROMIS Sexual Satisfaction – NC ProCESS	Baseline	N = 225 Alpha = 0.929	$\chi^2_5 = 53.12, p < 0.001$ RMSEA = 0.207 CFI = 0.992 WRMR = 0.919	GLOBSAT1 \neq	0.893	i	None
				GLOBSAT2 \neq	0.931	i	None
				GLOBSAT4 \neq	0.911	li	None
				GLOBSAT5 \neq	0.903	ii	None
				GLOBSAT1 \neq	0.922	None	None
	3-Month	N = 204 Alpha = 0.939	$\chi^2_5 = 38.40, p < 0.001$ RMSEA = 0.181 CFI = 0.995 WRMR = 0.614	GLOBSAT1 \neq	0.966	None	None
				GLOBSAT3 \neq	0.931	None	None
				GLOBSAT4 \neq	0.859	i	None
				GLOBSAT5 \neq	0.92	i	None
				GLOBSAT1 \neq	0.894	i	None
	12-Month	N = 447 Alpha = 0.940	$\chi^2_5 = 203.38, p < 0.001$ RMSEA = 0.295 CFI = 0.988 WRMR = 1.618	GLOBSAT2 \neq	0.926	i	None
				GLOBSAT3 \neq	0.932	None	None
				GLOBSAT4 \neq	0.92	ii	None
				GLOBSAT5 \neq	0.946	ii	None
				GLOBSAT1 \neq	0.872	i	None
	24-Month	N = 535 Alpha = 0.921	$\chi^2_5 = 216.29, p < 0.001$ RMSEA = 0.281 CFI = 0.983 WRMR = 1.919	GLOBSAT2 \neq	0.898	i	None
				GLOBSAT3 \neq	0.895	i	None
				GLOBSAT4 \neq	0.906	ii	None
				GLOBSAT5 \neq	0.904	ii	None
				GLOBSAT1 \neq	0.918	None	None
PROMIS Sexual Satisfaction - SexFS-Preferred	Baseline	N = 225 Alpha = 0.908	Just-identified	GLOBSAT2 \neq	0.963	None	None
				GLOBSAT3 \neq	0.896	None	None
				GLOBSAT1 \neq	0.930	None	None
	3-Month	N = 206 Alpha = 0.929	Just-identified	GLOBSAT2 \neq	0.974	None	None
				GLOBSAT3 \neq	0.927	None	None
				GLOBSAT1 \neq	0.923	None	None
	12-Month	N = 447 Alpha = 0.914	Just-identified	GLOBSAT2 \neq	0.952	None	None
				GLOBSAT3 \neq	0.920	None	None
				GLOBSAT1 \neq	0.901	None	None
	24-Month	N = 535 Alpha = 0.893	Just-identified	GLOBSAT2 \neq	0.932	None	None

Domain – Version	Time	Descriptive Statistics	Model Fit	Item ID	λ	LD Grouping Labels (Time-Specific)	# of Significant Monotonicity Violations
				GLOBSAT3 <u>†</u> ‡ [§]	0.895	None	None
PC SI Erectile Function - Clark (2005)	Baseline	N = 882 Alpha = 0.860	Just-identified	E10	0.804	None	None
				E30	0.949	None	None
				E40	0.935	None	None
	3-Month	N = 464 Alpha = 0.900	Just-identified	E10	0.851	None	None
				E30	0.968	None	None
				E40	0.937	None	None
	12-Month	N = 474 Alpha = 0.891	Just-identified	E10	0.836	None	None
				E30	0.953	None	None
				E40	0.938	None	None
	24-Month	N = 421 Alpha = 0.881	Just-identified	E10	0.818	None	None
				E30	0.961	None	None
				E40	0.930	None	None
PCSI MOS Sexual Problems – Clark (2005), less two items	Baseline	N = 886 Alpha = 0.817	Just-identified	E140	0.807	None	None
				E150	0.913	None	None
				E160	0.868	None	None
	3-Month	N = 467 Alpha = 0.811	Just-identified	E140	0.773	None	None
				E150	0.906	None	None
				E160	0.872	None	None
	12-Month	N = 475 Alpha = 0.809	Just-identified	E140	0.754	None	None
				E150	0.909	None	None
				E160	0.858	None	None
	24-Month	N = 421 Alpha = 0.804	Just-identified	E140	0.801	None	None
				E150	0.872	None	None
				E160	0.818	None	None

λ = standardized factor loading

LD = local dependence

Alpha = Ordinal coefficient alpha, an internal consistency measure computed based on polychoric correlations (Gadermann et al., 2012)

χ^2_{df} = scaled chi-square statistic with *df* degrees of freedom

p = *p* value associated with the scaled chi-square statistic

RMSEA = scaled root mean square error of approximation

CFI = scaled comparative fit index

WRMR = weighted root mean square residuals

† Items from PROMIS SexFS-V1 (Flynn et al., 2013). The underlined items are worded slightly differently than in PROMIS SexFS-V1.

~~‡~~ Items from PROMIS SexFS-V2 (Weinfurt et al., 2015).

[§] Items in PROMIS SexFS-Preferred (best-fitting items in the current study)

Table 4.

Within-Time Factor Correlations for the Assessment of Convergent and Discriminant Validity.

Time	Non-Sex-Related Domains		Sex-Related Domains					
	F1 Fatigue	F2 Phys. Fn.	PROMIS			PCSISDS		
			F3 Sex. Int.	F4 Erect. Fn.	F5 Sex. Sat.	F6 Erect. Fn.	F7 Sex. Prob	
Baseline	F1	---	---	---	---	---	---	---
	F2	-0.67	---	---	---	---	---	---
	F3	0.03	0.04	---	---	---	---	---
	F4	-0.24	0.22	0.44	---	---	---	---
	F5	-0.25	0.19	0.64	0.7	---	---	---
	F6	-0.26	0.07	0.39	0.95	0.59	---	---
	F7	0.2	-0.14	-0.36	-0.9	-0.59	-0.87	---
3-Month	F1	---	---	---	---	---	---	---
	F2	-0.54	---	---	---	---	---	---
	F3	-0.21	0.33	---	---	---	---	---
	F4	-0.16	0.34	0.41	---	---	---	---
	F5	-0.27	0.24	0.58	0.72	---	---	---
	F6	-0.2	0.26	0.35	0.9	0.76	---	---
	F7	0.33	-0.3	-0.36	-0.84	-0.81	-0.94	---
12-Month	F1	---	---	---	---	---	---	---
	F2	-0.66	---	---	---	---	---	---
	F3	-0.15	0.19	---	---	---	---	---
	F4	-0.24	0.22	0.35	---	---	---	---
	F5	-0.31	0.28	0.46	0.74	---	---	---
	F6	-0.23	0.24	0.23	0.88	0.73	---	---
	F7	0.25	-0.2	-0.26	-0.86	-0.76	-0.88	---
24-Month	F1	---	---	---	---	---	---	---
	F2	-0.59	---	---	---	---	---	---
	F3	0.05	0.01	---	---	---	---	---
	F4	-0.18	0.2	0.35	---	---	---	---
	F5	-0.21	0.24	0.44	0.68	---	---	---
	F6	-0.19	0.18	0.26	0.84	0.66	---	---
	F7	0.23	-0.14	-0.33	-0.84	-0.74	-0.86	---

F1: PROMIS Item Bank V1 Fatigue short form 4a.

F2: PROMIS Item Bank V2 Physical Function short form 4a.

F3: PROMIS Sexual Interest - SexFS-Preferred.

F4: PROMIS Erectile Function - SexFS-Preferred.

F5: PROMIS Sexual Satisfaction - SexFS-Preferred.

F6: PCSISDS Erectile Function - Clark (2005).

F7: PCSISDS MOS Sexual Problems - Clark (2005), less two items.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5.

Effect sizes for items that function differentially between surgery and no-surgery groups

Domain – Version	Item Name	Time	SIDS	UIDS	SIDN	UIDN	D-Max	ESSD
PROMIS Sexual Interest – SexFS-Preferred	SEXFCN1	24-Month	0.13	0.13	0.14	0.14	0.28	0.11
	SEXFCN3	3-Month	0.08	0.18	0.07	0.18	0.57	0.09
		24-Month	-0.01	0.04	-0.01	0.04	-0.09	-0.02
PROMIS Sexual Satisfaction – SexFS-Preferred	GLOBSAT2	24-Month	-0.09	0.19	-0.08	0.19	0.38	-0.08

SIDS = signed item difference in sample; UIDS = unsigned item difference in sample; SIDN = signed item difference in normal distribution; UIDN = unsigned item difference in normal distribution; D-Max = maximum difference in sample; ESSD = expected score standardized difference (Meade, 2010).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript