Development of the NIH PROMIS® Sexual Function and Satisfaction Measures in Patients with Cancer

Kathryn E. Flynn, PhD[1,2], **Li Lin, MS**[1], **Jill M. Cyranowski, PhD**[3], **Bryce B. Reeve, PhD**[4], **Jennifer Barsky Reese, PhD**[5], **Diana D. Jeffery, PhD**[6], **Ashley Wilder Smith, MPH, PhD**[7], **Laura S. Porter, PhD**[2], **Carrie B. Dombeck, MA**[1], **Deborah Watkins Bruner, RN, PhD, FAAN**[8], **Francis J. Keefe, PhD**[2], and **Kevin P. Weinfurt, PhD**[1,2]

[1]Duke Clinical Research Institute, Duke University School of Medicine, Durham, NC, USA

[2]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

[3]Departments of Psychiatry and Psychology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[4]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[5]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[6]TRICARE Management Activity, Department of Defense - Health Affairs, Falls Church, VA, USA

[7]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

[8]Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA, USA

Abstract

Introduction—We describe the development and validation of the PROMIS Sexual Function and Satisfaction (PROMIS SexFS) measures version 1.0 for cancer populations.

Aim—To develop a customizable self-report measure of sexual function and satisfaction as part of the U.S. National Institutes of Health PROMIS® Network.

Methods—Our multidisciplinary working group followed a comprehensive protocol for developing psychometrically robust patient reported outcome (PRO) measures including qualitative (scale development) and quantitative (psychometric evaluation) development. We performed an extensive literature review, conducted 16 focus groups with cancer patients and multiple discussions with clinicians, and evaluated candidate items in cognitive testing with patients. We administered items to 819 cancer patients. Items were calibrated using item response theory and evaluated for reliability and validity.

Main Outcome Measures—The PROMIS Sexual Function and Satisfaction (PROMIS SexFS) measures version 1.0 include 79 items in 11 domains: interest in sexual activity, lubrication, vaginal discomfort, erectile function, global satisfaction with sex life, orgasm, anal discomfort, therapeutic aids, sexual activities, interfering factors, and screener questions.

Corresponding Author Name: Kathryn E Flynn, PhD, kathryn.flynn@duke.edu.
Conflict of Interest: None

Results—In addition to content validity (patients indicate that items cover important aspects of their experiences) and face validity (patients indicate that items measure sexual function and satisfaction), the measure shows evidence for discriminant validity (domains discriminate between groups expected to be different), convergent validity (strong correlations between scores on PROMIS and scores on conceptually-similar older measures of sexual function), as well as favorable test-retest reliability among people not expected to change (inter-class correlations from 2 administrations of the instrument, 1 month apart).

Conclusions—The PROMIS SexFS offers researchers a reliable and valid set of tools to measure self-reported sexual function and satisfaction among diverse men and women. The measures are customizable; researchers can select the relevant domains and items comprising those domains for their study.

Keywords

Introduction

Cancer and other chronic diseases, as well as their treatments, often have adverse effects on sexual function including sexual interest, arousal, orgasm, and comfort during intercourse.[1, 2] In addition, chronic diseases can have indirect effects on sexual function by altering relationships and self-image as a result of experiencing pain, fatigue, disfigurement, and dependency.[1, 3–5] Tools to assess sexual function and satisfaction in a comprehensive, valid and precise manner are required to develop targeted and effective treatments. A recent review highlighted the critical need for improved patient-reported outcome measures of sexual function.[6]

The National Institutes of Health's Patient-Reported Outcomes Measurement Information System® (PROMIS®) Network was created in 2004 to develop standardized, precise, and valid measures of patient-reported outcomes for use in research across chronic diseases.[7] Unique to the PROMIS initiative is the extensive, comprehensive development of its item banks, i.e., a large set of questions and response options (items) for each measured domain. PROMIS item banks undergo qualitative evaluation using focus groups, cognitive testing, and psychometric evaluation incorporating item response theory (IRT). Content represents substantial input from patients, clinicians, and survey methodologists, making the PROMIS measures both clinically relevant and valid for research. Finally, PROMIS item banks provide an unusual degree of flexibility to researchers, allowing them to select relevant items to accurately measure the domains of interest for their target population. In 2004, the National Cancer Institute identified sexual function as highly relevant for cancer survivors and provided support for creating a new measure of sexual function through the PROMIS network.

Aims

We describe the development and validation of the PROMIS Sexual Function and Satisfaction (PROMIS SexFS) measures version 1.0 for cancer populations. Our goal was to develop a customizable self-report set of measures that was a) comprehensive in scope, b) flexible to meet diverse research needs, c) broadly applicable with respect to age, gender, sexual orientation, partner status, and literacy level, and d) disease-neutral yet useful across a wide spectrum of diseases with symptoms that potentially interfere with functioning.

Methods

Our multidisciplinary working group followed the comprehensive PROMIS protocol for developing psychometrically robust patient reported outcome (PRO) measures.[8, 9] The qualitative steps (scale development) are reported in detail elsewhere[6, 10, 11] and described in brief here. The quantitative steps (psychometric evaluation) are described in detail here. All patient participants provided informed consent. The Institutional Review Board of the XXX Health System approved this study.

Scale Development

Review of Extant Items—Using a consensus-driven approach, we conducted a literature search for articles published from 1991–2007 that reported the administration of a self-reported measure of sexual function in a cancer population; see Jeffery 2009 for detail.[6] Based on this literature review, we developed a preliminary conceptual model to reflect domains to be included in the measures: interest in sexual activity, lubrication, vaginal discomfort, erectile function, orgasm, anal discomfort, frequency of sexual activity, and sexual satisfaction. We categorized more than 1100 items from existing measures into these domains and selected ~50 clinically relevant items for further testing.

Item banks are dynamic and can incorporate extant items on the same scoring metric. We incorporated into the PROMIS SexFS some items that are publically available or for which the copyright holders granted permission. Thus, some PROMIS SexFS instruments include modified items from other sexual function instruments (e.g., UCLA-Prostate Cancer Index,[12] Female Sexual Function Index [FSFI][13]).

Focus Groups—We conducted 16 focus groups with 109 patients to establish the content validity of the PROMIS SexFS measures.[10] Patient focus groups focused on physical and psychosocial impacts of cancer on sexual function and intimate relationships; see Flynn 2011 for detail. We also discussed the clinical relevance of the proposed conceptual model with oncology researchers and clinicians and sought their views on how cancer and its treatment affected patients' sexual health. We published our working conceptual measurement model[10] and provide an adapted version to show the domains covered and how they may relate to each other (Figure). We wrote ~40 new items based on the results of these groups/discussions and added domains describing symptoms and other factors that interfere with sexual satisfaction.

Cognitive Interviews—To evaluate face validity, we tested items in cognitive interviews. Each of 83 candidate items was evaluated for clarity, sensitivity, and relevance by 5 or more participants, at least 2 of whom had less than a 9th grade education or reading level as tested by the Wide Range Achievement Test 4;[14] see Fortune-Greeley 2009 for detail.[11] The 39 participants (16 low literacy) were diverse with respect to sex and race as well as cancer type and stage. Items were revised based on results of the interviews. Substantially revised items were retested in additional interviews.

Other Qualitative Review—The team responsible for translating all PROMIS measures reviewed all items and minor changes were made to reduce potential difficulty with translation to non-English languages.[15] We also convened 7 additional clinical and academic experts on sexual function and cancer to review the conceptual model and item development to date.

Psychometric Evaluation

Study Sample—Patients who participated in item testing were recruited from the Duke University tumor registry (mailed invitation, 56% of sample), Duke's private diagnostic oncology clinics (in person, 14% of sample), as well as the NexCura internet panel (emailed invitation, 30% of sample). For patients recruited through NexCura, we verified cancer diagnosis and treatment status with their treating physician. We also targeted recruitment of additional lesbian, gay, and bisexual cancer patients and survivors through online communities, though this strategy yielded few participants (n=14). Patients were eligible if they were 18 years of age or older, had been diagnosed with cancer, and were able to speak English. We aimed for a diverse sample regarding sex, race, tumor site, and whether the person was undergoing active treatment for cancer or in post-treatment follow-up (Table 1). The initial target sample comprised ~250 men and ~250 women, but roughly 40% of this sample had not been sexually active in the past 30 days. Thus, we recruited a supplemental sample of ~300 sexually active cancer patients and survivors from the same sources. Participants completed 2 screener items, 79 PROMIS SexFS items, and 49 items from other commonly used measures of sexual function (FSFI, UCLA Prostate Cancer Index, Medical Outcomes Study Sexual Problems Survey,[16] International Index of Erectile Function [IIEF][17]).

Analytic Approach—We performed a psychometric evaluation of the SexFS domains of: Interest in Sexual Activity, Lubrication [women], Vaginal Discomfort [women], Erectile Function [men], Orgasm, Anal Discomfort, and Global Satisfaction with Sex Life using the established PROMIS methodology based on classical test theory and IRT.[9] We first examined the distribution of item responses and examined item-total correlations within the same SexFS domain. Multiple statistical software programs were used, including Mplus (version 5.21, Muthen & Muthen) and MULTILOG (ver 7.03, Scientific Software International) as noted below, and SAS software 9.2 (SAS Institute Inc., Cary, NC, USA) for other analyses. Psychometric analyses were not conducted for the subdomains not thought to comprise latent variables (Sexual Activities, Interfering Factors, Therapeutic Aids, and Screener Items).

After adequate distribution of item responses and item-total correlations within the same SexFS domain being verified were established, confirmatory factor analyses (CFA) were conducted in Mplus to supply statistics and factor score estimates for evaluating assumptions applied in IRT modeling, namely unidimensionality, local independence, and monotonicity. *Unidimensionality* means that the items measure a common underlying aspect of functioning. This was tested by examining the fit of a single-factor CFA model for ordinal categorical item responses.[18] For this, multiple indices were computed, including the comparative fit index (criterion for good fit is >0.95), Tucker-Lewis index (>0.95 for good fit), and root mean square error of approximation (<0.06 for good fit). When CFA showed a poor fit, exploratory factor analysis (EFA) was used to explore alternative models. *Local independence* requires that there are no significant associations among item responses once the dominant factor influencing a person's response to an item is controlled. Item pairs with high residual correlations (>0.2) were flagged as possible cases of local dependence. *Monotonicity* assumes that the probability of endorsing an item response indicative of better health status should increase as the underlying level of health increases. For each item, we examined the relationship between the item score and the factor score graphically and via a polyserial correlation.

Item Calibration—IRT modeling was used to evaluate the psychometric properties of each item within a scale and to calibrate the metric for the items, i.e., to generate statistical parameters that allow translations of a person's item responses into an estimate of that

person's underlying functioning. The probability of choosing each item-response category is modeled as a function of the underlying latent trait, denoted by theta (θ). Samejima's Graded Response Models (GRM) was chosen as the primary IRT model, following PROMIS conventions.[19, 20] In the GRM, each item is assigned a set of parameter estimates: a slope parameter for discrimination and a number of between-category thresholds. Item characteristic curves, as well as item and test information curves were produced and examined to ensure adequate properties of items and measures. Scores for each respondent were calculated based on their responses to each of the items within the SexFS item bank and rescaled to a *T* distribution (mean = 50, standard deviation = 10). IRT analysis was performed by using MULTILOG.

Differential Item Functioning (DIF) and Measurement Invariance—DIF occurs when people from different groups with the same latent trait have a different probability of giving a certain response on a question.[21] Differences between groups (e.g., males and females) in individual item behavior were assessed via an analysis of DIF. We tested DIF for males vs females (on domains that both men and women answered, including Global Satisfaction with Sex Life and Interest in Sexual Activity) and web- versus phone-based mode of administration (on all domains). We used ordinal logistic regression to examine DIF at the item level. An item was classified as displaying DIF if (1) the Likelihood ratio chi-square test in logistic regression had a p-value 0.01, and (2) the Zumbo-Thomas effect size, which is increment of the R-square, was 0.130.[22]

We also used multiple-group analysis in MPlus to examine measurement invariance at the domain level. Because our sample size for some DIF analyses was smaller than typically desired, we also relied on graphical methods, producing item-predicted factor score mean curves to visually verify the magnitude of the DIF and whether non-uniform DIF exists. Intra-class correlation coefficients were calculated to assess the agreement between factor scores generated from group-specific and common IRT parameter estimates.

Concurrent and Construct Validity—To establish concurrent validity, we first examined Pearson correlation coefficients between subdomains of the PROMIS SexFS and other widely used measures of conceptually similar constructs, namely the 19-item FSFI[13] and the 15-item IIEF.[17] Second, we examined whether scores on selected subdomains of the PROMIS SexFS could discriminate between groups that should, in theory, differ in terms of their sexual experiences. During item testing, participants were asked whether they had ever asked an oncology professional about sexual problems. We hypothesized that asking for help with sexual problems may indicate a clinically meaningful decrement in function.

Reliability Analyses—We calculated Cronbach's alpha for men and women separately to demonstrate internal consistency. To examine the validity of the 30-day recall period and test-retest reliability, we conducted a separate 30-day diary study in 202 men and women for whom we did not expect changes over the month.[23] To express the consistency of scores over time, we report intraclass correlations coefficients (ICCs) and their corresponding 95% confidence intervals (CIs) for two administrations of the measure, one month apart.

Results

Psychometric Evaluation

A summary of fit statistics is shown in Table 2. We expect future data collection efforts will allow for calibration of the Orgasm and Anal Discomfort subdomains, but there were not enough responses (Anal Discomfort) or items (Orgasm) to support calibration at present. For the other domains, all calibrated items satisfied the basic assumptions for adequate

distribution (none displayed sparseness or monotonicity and higher item scores corresponded to higher scale sum-scores). All calibrated item banks had high comparative fit indices, supporting unidimensional models, though some lack of fit was suggested by the RMSEA. With rare exception, items had factor loading and polyserial correlation of 0.8 or greater. Item properties (from item calibration) are available online.[24]

DIF

For DIF analysis by survey mode (phone versus internet), there was mode equivalence (i.e., no DIF) between phone and internet surveys at both the domain- and item-level; however, the multiple group models failed to converge for the Vaginal Discomfort and Erectile Function domains due to inadequate sample size. For Satisfaction, many of the likelihood-ratio tests from ordinal logistic regression were significant for DIF ($p < 0.001$); however, the effect sizes were small and did not reach the pre-specified 0.130 cut-point, tempering the finding of differences in item response by gender. Multiple-group and graphical analyses confirmed the lack of noticeable DIF. The intraclass correlation between the factor scores predicted by the common and sex-specific IRT models was greater than 0.99, which reflected an almost perfect psychometric agreement between males and females. For Interest in Sexual Activity, none of the items displayed significant DIF by the ordinal logistic regression method, though the multiple-group analysis suggested possible DIF. Graphical analyses did not reveal any substantial DIF, though the ICC between the factor scores predicted by the common and gender-specific IRT models was 0.88. This ambiguity regarding gender-related DIF for Interest in Sexual Activity suggests the need for more research.

Scoring

Calibrated subdomain scores are expressed as T scores (mean = 50, standard deviation= 10). A T score of 50 corresponds to the mean response among the cancer survivors used for item testing. Higher scores corresponded to higher levels of the item or domain.

Administration

All items in the PROMIS SexFS item banks are not intended to be administered together, but rather that researchers should have the flexibility to select the items that are relevant to their specific sample. For the calibrated scales, if one or more items from within that instrument are administered, a respondent's score will be calculated using IRT parameters (either through the PROMIS Assessment Center or look-up tables provided in the user manual). For the 6 non-calibrated item banks (e.g., Interfering Factors), the items within those instruments are not combined in any way to create a score. Each item in these instruments measures a very specific construct corresponding only to that item (e.g., how much has fatigue affected satisfaction with sex life). For any given item in these uncalibrated instruments, the researcher can use the raw item responses directly for analyses.

In addition to the full banks, from which researchers can select items, we produced a brief "off the shelf" measure, the PROMIS Sexual Function and Satisfaction Brief Profile. It includes 1 to 3 of the best general-purpose items from each of the key domains of sexual function (8 total items for men, 10 for women).

Validity

In general, the correlations between subdomains of the PROMIS SexFS and the corresponding subdomains of the FSFI and IIEF provide strong evidence for the construct validity of the PROMIS SexFS and the brief profile measures (Table 3). Additionally, patients who had asked a provider for help with sexual problems were significantly different

from those who had not asked and in the directions we expected (Table 4). Askers had significantly greater interest in sexual activity, increased vaginal discomfort, and lower levels of erectile function, lubrication, orgasm, and overall satisfaction. These effect sizes were greater than or equal to the effects for the corresponding subscales of the FSFI and IIEF. In three cases, the PROMIS SexFS and brief profiles detected statistically significant (p<.05) differences between those who did and did not ask, whereas the FSFI or IIEF did not.

Reliability

Estimates of internal consistency (Cronbach's alpha) were calculated separately for men and women and were high, ranging from 0.87–0.95 (Table 5). Test-retest reliability was also favorable, based on intraclass correlation coefficients (ranging from 0.71–0.87).

Discussion

The PROMIS SexFS is a customizable self-reported measure of sexual function with demonstrated validity for use in cancer populations. Version 1.0 consists of 11 domains. It has 79 items in 5 calibrated scales (Interest in Sexual Activity, Lubrication [women], Vaginal Discomfort [women], Erectile Function [men], Global Satisfaction with Sex Life) and 6 collections of stand-alone items (Sexual Activities, Orgasm, Interfering Factors, Therapeutic Aids, Anal Discomfort, Screener Items). The PROMIS SexFS instruments are available for download on the Assessment Center™ website (http://assessmentcenter.net/).

The PROMIS SexFS offers researchers an outcome measure for use with men and women with common domains where possible (e.g., Satisfaction). It allows researchers to record frequencies of specific sexual activities but does not generally refer to specific activities when measuring function, making it neutral to the respondent's partner status or sexual orientation. It offers a group of items (Interfering Factors) to capture reasons for respondent dissatisfaction with sex life. This greater conceptual precision in understanding the problems experienced by individuals could lead to more focused interventions.

While it was developed primarily as an outcome measure, the PROMIS SexFS may offer benefit when used in clinical practice. For instance, screening patients for sexual problems or tracking the trajectory of patients' sexual activities, function, satisfaction, and/or use of therapeutic aids could help practitioners identify patients who may be in need of intervention or treatment. Administering this measure may also provide psycho-educational value by helping patients to rate and track their sexual satisfaction and to identify specific aspects of sexual function that they wish to improve in order to increase the quality of their sexual lives. Finally, clinicians may find that offering their patients the opportunity to report on their sexual function and satisfaction in a confidential manner will facilitate conversations with patients about this sensitive topic.

Most PROMIS SexFS items are not specific to cancer but have thus far only been validated in cancer samples. However, we designed the measure to accommodate various health states or conditions that could affect sexual function. Ongoing research is currently testing the PROMIS SexFS measure in other targeted groups (e.g., adults with diabetes, heart disease, anxiety, depression, age >65, lesbian, gay, or bisexual, as well as a nationally representative sample of U.S. adults). We are also testing the expansion of the existing domains (e.g., Orgasm), and the addition of new domains (e.g., subjective arousal, oral discomfort, and vulvar discomfort)..

Conclusions

The PROMIS SexFS measures version 1.0 offer researchers a reliable and valid tool to measure self-reported sexual function and satisfaction among diverse men and women with cancer. The measure is customizable in that researchers can select the relevant SexFS domains and items comprising those domains for their study. The measures are comprehensive in scope, covering both physical and psychological components. They are broadly applicable with respect to age, gender, sexual orientation, partner status, and literacy level. Finally, they are disease-neutral yet also able to capture relevant symptoms of cancer and its treatment that interfere with sexual satisfaction. These features should enhance our ability to describe and intervene on specific aspects of the sexual function and satisfaction of patients with cancer.

Acknowledgments

References

1. Basson R, Schultz WW. Sexual sequelae of general medical disorders. Lancet. 2007; 369:409–424. [PubMed: 17276781]

2. Namiki S, Ishidoya S, Nakagawa H, Ito A, Kaiho Y, Tochigi T, Takegami M, Arai Y. The relationships between preoperative sexual desire and quality of life following radical prostatectomy: a 5-year follow-up study. J Sex Med. 2012; 9:2448–2456. [PubMed: 22616799]

3. Clayton A, Ramamurthy S. The impact of physical illness on sexual dysfunction. Adv Psychosom Med. 2008; 29:70–88. [PubMed: 18391558]

4. Panjari M, Bell RJ, Burney S, Bell S, McMurrick PJ, Davis SR. Sexual function, incontinence, and wellbeing in women after rectal cancer - a review of the evidence. J Sex Med. 2012 Epub ahead of print.

5. Bober SL, Giobbie-Hurder A, Emmons KM, Winer E, Partridge A. Psychosexual functioning and body image following a diagnosis of ductal carcinoma in situ. J Sex Med. 2012 Epub ahead of print.

6. Jeffery DD, Tzeng JP, Keefe FJ, Porter LS, Hahn EA, Flynn KE, Reeve BB, Weinfurt KP. Initial report of the cancer Patient-Reported Outcomes Measurement Information System (PROMIS) sexual function committee: review of sexual function measures and domains used in oncology. Cancer. 2009; 115:1142–1153. [PubMed: 19195044]

7. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007; 45:S3–S11. [PubMed: 17443116]

8. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. Med Care. 2007; 45:S12–S21. [PubMed: 17443114]

9. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Lai JS, Cella D. 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:S22–S31. [PubMed: 17443115]

10. Flynn KE, Jeffery DD, Keefe FJ, Porter LS, Shelby RA, Fawzy MR, Gosselin TK, Reeve BB, Weinfurt KP. Sexual functioning along the cancer continuum: focus group results from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)). Psychooncology. 2011; 20:378–386. [PubMed: 20878833]

11. Fortune-Greeley AK, Flynn KE, Jeffery DD, Williams MS, Keefe FJ, Reeve BB, Willis GB, Weinfurt KP. Using cognitive interviews to evaluate items for measuring sexual functioning across cancer populations: improvements and remaining challenges. Qual Life Res. 2009; 18:1085–1093. [PubMed: 19672697]

12. Litwin MS, Hays RD, Fink A, Ganz PA, Leake B, Brook RH. The UCLA Prostate Cancer Index: development, reliability, and validity of a health-related quality of life measure. Med Care. 1998; 36:1002–1012. [PubMed: 9674618]

13. Rosen R, Brown C, Heiman J, Leiblum S, Meston C, Shabsigh R, Ferguson D, D'Agostino R Jr. The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. J Sex Marital Ther. 2000; 26:191–208. [PubMed: 10782451]

14. Wilkinson, G. Wide Range Achievement Test (WRAT 3). Wide Range Inc; Wilmington: 1993.

15. Link CL, Lutfey KE, Steers WD, McKinlay JB. Is abuse causally related to urologic symptoms? Results from the Boston Area Community Health (BACH) survey. Eur Urol. 2007; 52:397–406. [PubMed: 17383083]

16. Seyfried LS, Wallner LP, Sarma AV. Psychosocial Predictors of Lower Urinary Tract Symptom Bother in Black Men: The Flint Men's Health Study. J Urol. 2009; 182:1072–1077. [PubMed: 19616799]

17. Rosen RC, Riley A, Wagner G, Osterloh IH, Kirkpatrick J, Mishra A. The international index of erectile function (IIEF): a multidimensional scale for assessment of erectile dysfunction. Urology. 1997; 49:822–830. [PubMed: 9187685]

18. McDonald, RP. Test theory: A unified treatment. Erlbaum; Mahwah: 1999.

19. Samejima, F. Psychometrika Monographs. 1969. Estimation of latent ability using a response pattern of graded scores.

20. Samejima, F. Graded Response Model. In: van der Linden, W.; Hambleton, R., editors. Handbook of Modern Item Response Theory. Springer; New York: 1997. p. 85-100.

21. Reise SP, Henson JM. A discussion of modern versus traditional psychometrics as applied to personality assessment scales. J Pers Assess. 2003; 81:93–103. [PubMed: 12946916]

22. Zumbo, BD. Directorate of Human Resources Research and Evaluation. Department of National Defense; Ottawa, ON: 1999. A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-type (Ordinal) Item Scores.

23. Weinfurt, K. Validity of 1-month Recall for Components of Sexual Function (poster presentation). Presented at the 18th Annual International Society for Quality of Life Research Conference; Denver CO. 2011.

24. PROMIS Sexual Function Domain Group. PROMIS Sexual Function User Manual. 2011.
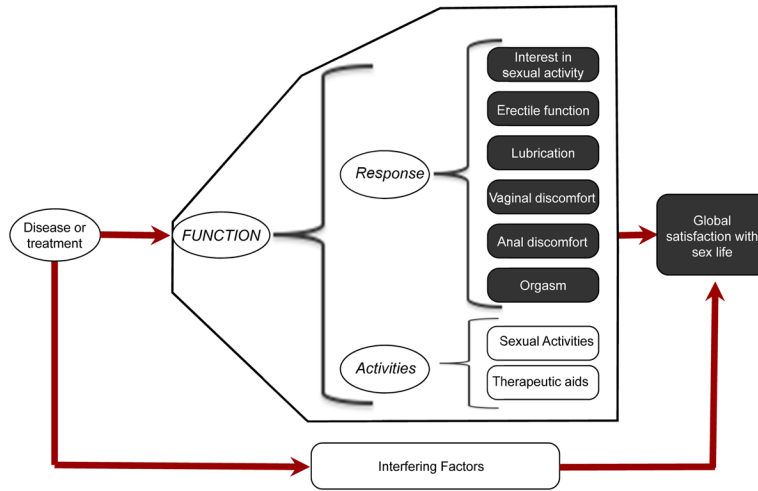
Figure. Conceptual Measurement Model of Sexual Function And Related Domains
Adapted from Flynn KE, Jeffery DD, Keefe FJ, et al. Sexual functioning along the cancer continuum: focus group results from the Patient-Reported Outcomes Measurement Information System (PROMIS). *Psycho-Oncology.* Apr 2011;20(4):378–386.

Table 1

Characteristics of Patients in Item Testing Sample (n = 819)

Characteristic	Total (n = 819)	Men (n = 389)	Women (n = 430)
Age, mean ± SD, y	58.5 ± 11.8	61.4 ± 10.8	55.9 ± 12.2
Age group, No. (%)			
40 years	59 (7)	14 (4)	45 (10)
41 to 50 years	127 (16)	29 (7)	98 (23)
51 to 64 years	377 (46)	197 (51)	180 (42)
65 to 79 years	232 (28)	134 (35)	98 (23)
80 years	21 (3)	13 (3)	8 (2)
Race, No. (%)			
Black or African American	80 (10)	39 (10)	41 (10)
American Indian/Alaska Native	10 (1)	6 (2)	4 (<1)
Asian	12 (1)	2 (<1)	10 (2)
Native Hawaiian/Other Pacific Islander	10 (1)	1 (<1)	1 (<1)
White	705 (87)	338 (87)	367 (86)
Multiple races or other	2 (< 1)	1 (<1)	5 (1)
Hispanic or Latino ethnicity, No. (%)	21 (3)	10 (3)	11 (3)
Educational attainment, No. (%)			
Less than high school	21 (3)	14 (4)	10 (2)
High school graduate/GED	100 (12)	35 (9)	65 (15)
Some college	255 (31)	122 (31)	133 (31)
College degree	229 (28)	114 (29)	115 (27)
Advanced degree (MA, PhD, MD)	211 (26)	104 (27)	107 (25)
Treatment status in past month, No. (%)			
None (ie, posttreatment follow-up)	526 (64)	290 (75)	236 (55)
Undergoing treatment	290 (36)	98 (25)	192 (45)
Radiation therapy	29 (10)	15 (4)	14 (3)
Hormonal therapy (eg, tamoxifen, anastrozole, leuprolide)	140 (48)	28 (7)	112 (26)
Chemotherapy (injection or oral)	116 (40)	46 (12)	70 (16)
Immunotherapy (eg, interferon)	9 (3)	2 (<1)	7 (2)
Other	36 (12)	17 (4)	19 (4)
Recurrence of cancer, No. (%)	151 (18)	68 (17)	83 (19)
Cancer spread to lymph nodes, No. (%)	202 (25)	68 (17)	134 (31)
Cancer spread to another area, No. (%)	134 (16)	57 (15)	77 (18)
Primary cancer diagnosis, No. (%)			
Bone/muscle cancer	14 (2)	6 (2)	8 (2)
Brain cancer	4 (< 1)	3 (<1)	1 (<1)
Breast cancer	252 (35)	1 (<1)	251 (62)
Colorectal	98 (13)	57 (18)	41 (10)

Characteristic	Total (n = 819)	Men (n = 389)	Women (n = 430)
Esophageal or stomach cancer	17 (2)	13 (4)	4 (<1)
Gynecologic cancer	29 (4)	--	29 (7)
Head/neck cancer	9 (< 1)	5 (2)	4 (<1)
Hodgkin lymphoma	23 (3)	7 (2)	16 (4)
Leukemia	20 (3)	12 (4)	8 (2)
Liver cancer	3 (< 1)	1 (<1)	2 (<1)
Lung cancer	56 (8)	35 (11)	21 (5)
Melanoma	4 (< 1)	2 (<1)	2 (<1)
Multiple Myeloma	2 (< 1)	2 (<1)	--
Non-Hodgkin lymphoma	12 (2)	7 (2)	5 (1)
Pancreatic cancer	5 (< 1)	3 (<1)	2 (<1)
Prostate cancer	146 (20)	146 (45)	--
Urologic cancer	23 (3)	16 (5)	7 (2)
Unknown or Other	9 (1)	6 (2)	3 (<1)

Table 2

Fit Indices for Confirmatory Factor Analysis of Calibrated Subdomains

Instrument Name	# of items	CFI	TLI	RMSEA
Global Satisfaction with Sex Life	7	0.983	0.976	0.168
Interest in Sexual Activity	4	0.998	0.995	0.129
Lubrication	8	0.985	0.979	0.187
Vaginal Discomfort	10	0.993	0.991	0.124
Erectile Function	8	0.988	0.986	0.134

Abbreviation: CFI, comparative fit index; TLI, Tuker-Lewis index; RMSEA, root-mean-square error of approximation.

Table 3

Correlations between PROMIS Sexual Function and Satisfaction Full Bank and Brief Profile Subdomains and Corresponding Measures

Women (n = 430)		Men (n = 388)	
Measures	**r**[*]	**Measures**	**r**[*]
PROMIS Interest in Sexual Activity	0.84 (0.82)	PROMIS Interest in Sexual Activity	0.82 (0.79)
FSFI[†] Desire		IIEF[‡] Desire	
FSFI Arousal	0.71 (0.68)		
PROMIS Lubrication	0.92 (0.90)	PROMIS Erectile Function	0.81 (0.69)
FSFI Lubrication		IIEF Erectile Function	
PROMIS Vaginal Discomfort	0.90 (0.84)		
FSFI Pain			
PROMIS Orgasm	0.78 (0.78)	PROMIS Orgasm	0.62 (0.62)
FSFI Orgasm		IIEF Orgasmic Function	
PROMIS Global Satisfaction with Sex Life	0.76 (0.62)	PROMIS Global Satisfaction with Sex Life	0.82 (0.66)
FSFI Satisfaction		IIEF Overall Satisfaction	
		IIEF Intercourse Satisfaction	0.75 (0.68)

Note: Correlations in parentheses are for the PROMIS Sexual Function and Satisfaction Brief Profile scores.

[*] Pearson correlation coefficients

[†] Female Sexual Function Index

[‡] International Index of Erectile Function

Table 4

Effect Sizes Discriminating Askers From Non-Askers[a] (n = 806)

Domain	PROMIS Sexual Function and Satisfaction Full Bank		PROMIS SexFS Brief Profile		Legacy Measures		
	Effect Size	P Value[b]	Effect Size	P Value[b]	Domain	Effect Size	P Value[b]
Interest in Sexual Activity (women)	0.18	0.16	0.16	0.19	FSFI Desire	−0.05	0.69
					FSFI Arousal	−0.13	0.35
Interest in Sexual Activity (men)	0.22	0.04	0.21	0.05	IIEF Desire	0.14	0.20
Erectile Function	−0.45	<.0001	−0.49	<.0001	IIEF Erectile Function	−0.22	0.04
Lubrication	−0.75	<.0001	−0.66	<.0001	FSFI Lubrication	−0.67	<.0001
Vaginal Discomfort	0.75	<.0001	0.60	<.0001	FSFI Pain	−0.60	<.0001
Orgasm (women)	−0.40	<.0001	−0.40	0.003	FSFI Orgasm	−0.11	0.41
Orgasm (men)	−0.55	<.0001	−0.55	<.0001	IIEF Orgasmic Function	−0.55	<.0001
Global Satisfaction with Sex Life (women)	−0.24	0.05	−0.27	0.03	FSFI Satisfaction	0.08	0.54
Global Satisfaction with Sex Life (men)	−0.15	0.15	−0.18	0.09	IIEF Overall Satisfaction	−0.16	0.14

Abbreviation: PROMIS, Patient Reported Outcomes Measurement Information System; FSFI, Female Sexual Function Index; IIEF, International Index of Erectile Function.

[a] Difference in the means between people who answered "Yes" (n = 237) and "No" (n = 569) to the question, "Have you ever asked an oncology doctor or nurse about problems with your sex life?" divided by the common standard deviation.

[b] From t-test comparing PROMIS measure to legacy measure.

Table 5

Reliability of Calibrated Subdomains

Measures	Women		Men	
	Cronbach's alpha (n = 430)	ICC PROMIS SexFS (Brief Profile) (n = 101)	Cronbach's alpha (n = 388)	ICC PROMIS SexFS (Brief Profile) (n = 101)
PROMIS Interest in Sexual Activity	0.89	0.77 (0.72)	0.87	0.71 (0.65)[*]
PROMIS Lubrication	0.95	0.87 (0.87)		
PROMIS Vaginal Discomfort	0.94	0.80 (0.75)		
PROMIS Erectile Function			0.92	0.87 (0.77)
PROMIS Global Satisfaction with Sex Life	0.93	0.75 (0.69)	0.92	0.74 (0.66)

[*] Numbers reflect the deletion of a single outlier with a "5" at first administration and a "1" at second administration. Inclusion of the outlier results in ICCs of 0.55 and 0.54 for the full bank and brief profile measures, respectively.