

A new measure of sexual function in women with pelvic floor disorders (PFD): the Pelvic Organ Prolapse/Incontinence Sexual Questionnaire, IUGA-Revised (PISQ-IR)

R. G. Rogers · T. H. Rockwood · M. L. Constantine ·
R. Thakar · D. N. Kammerer-Doak · R. N. Pauls ·
M. Parekh · B. Ridgeway · S. Jha · J. Pitkin · F. Reid ·
S. E. Sutherland · E. S. Lukacz · C. Domoney · P. Sand ·
G. W. Davila · M. E. Espuna Pons

Received: 6 August 2012 / Accepted: 1 December 2012 / Published online: 30 April 2013
© The International Urogynecological Association 2013

Abstract

Introduction and hypothesis The objective of this study was to create a valid, reliable, and responsive sexual function measure in women with pelvic floor disorders (PFDs) for both sexually active (SA) and inactive (NSA) women.

Methods Expert review identified concept gaps and generated items evaluated with cognitive interviews. Women underwent Pelvic Organ Prolapse Quantification (POPQ) exams and completed the Incontinence Severity Index (ISI), a prolapse question from the Epidemiology of Prolapse and Incontinence

Questionnaire (ISI scores), the Pelvic Floor Distress Inventory-20 (PFDI-20), and the Female Sexual Function Index (FSFI). Principle components and orthogonal varimax rotation and principle factor analysis with oblique rotation identified item grouping. Cronbach's alpha measured internal consistency. Factor correlations evaluated criterion validation. Change scores compared to change scores in other measures evaluated responsiveness among women who underwent surgery. **Results** A total of 589 women gave baseline data, 200 returned surveys after treatment, and 147 provided test-retest

A related editorial can be found at doi: [10.1007/s00192-012-1952-3](https://doi.org/10.1007/s00192-012-1952-3).

Electronic supplementary material The online version of this article (doi:10.1007/s00192-012-2020-8) contains supplementary material [the Pelvic Organ Prolapse/Incontinence Sexual Questionnaire, IUGA-Revised (PISQ-IR)], which is available to authorized users.

R. G. Rogers (✉)
University of New Mexico Health Sciences Center, Albuquerque,
NM, USA
e-mail: rrogers@salud.unm.edu

T. H. Rockwood
University of Minnesota, Minneapolis, MN, USA

M. L. Constantine
University of Michigan, Ann Arbor, MI, USA

R. Thakar
Mayday University Hospital, Surrey, UK

D. N. Kammerer-Doak
Women's Pelvic Specialty Care, Albuquerque, NM, USA

R. N. Pauls
Division of Urogynecology and Reconstructive Surgery, Good
Samaritan Hospital, Cincinnati, OH, USA

M. Parekh
Prime Care of SE Ohio, Genesis Health System,
Zanesville, OH, USA

B. Ridgeway
The Cleveland Clinic, Cleveland, OH, USA

S. Jha
Royal Hallamshire Hospital, Department of Urogynaecology,
Sheffield Teaching Hospitals NHS Foundation Trust,
Sheffield, UK

J. Pitkin
Northwick Park & St Marks Hospital, Middlesex, UK

F. Reid
The Warrell Unit, Whitworth Park,
Manchester, UK

data. For SA women, 3 subscales each in 2 domains (21 items) and for NSA women 2 subscales in each of 2 domains (12 items) emerged with robust psychometric properties. Cronbach's alpha ranged from .63 to .91. For SA women, correlations were in the anticipated direction with PFDI-20, ISI, and FSFI scores, POPQ, and EPIQ question #35 (all $p < .05$). PFDI-20, ISI, and FSFI subscale change scores correlated with Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire International Urogynecological Association-revised (PISQ-IR) factor change scores and with mean change scores in women who underwent surgery (all $p < .05$). For NSA women, PISQ-IR scores correlated with PFDI-20, ISI scores, and with EPIQ question #35 (all $p < .05$). No items demonstrated differences between test and retest (all $p \geq .05$), indicating stability over time.

Conclusions The PISQ-IR is a valid, reliable, and responsive measure of sexual function.

Keywords Sexual function questionnaire · Pelvic organ prolapse · Urinary and/or anal incontinence

Introduction

Sexual health is a state of physical, emotional, mental, and social well-being in relation to sexuality; it is not merely the absence of disease, dysfunction, or infirmity [1]. Pelvic floor disorders (PFD), including urinary (UI) and anal incontinence (AI) and pelvic organ prolapse (POP), are common presenting gynecological complaints and adversely affect quality of life, including sexual health [2, 3]. Up to 60 % of sexually active women attending urogynecology clinics report sexual dysfunction [3]. Despite the prevalence, only a minority of urogynecologists consistently screen patients for sexual complaints. Lack of time, uncertainty about therapeutic options, and older age of the patient have been cited as potential reasons for failing to address sexual function as part of routine gynecological history [4].

Questionnaires play an integral role in the evaluation of female sexual function. A number of general sexual function questionnaires have been developed and utilized in the evaluation of women with PFD, including the Female Sexual

Function Index (FSFI) [5], the Profile of Female Sexual Function (PFSF) [6], and the McCoy Female Sexuality Questionnaire (MFSQ) [7]. While these measures are validated, they were not developed to focus on PFD and its impact on sexual health, and because they are not condition-specific for PFD, they may not be sensitive enough to detect differences in sexual function due to PFD. The Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ) and its short form version, the PISQ-12, are the only current validated condition-specific female sexual function questionnaires purposefully developed to assess sexual function in women with UI and/or POP [8, 9].

Questionnaires utilized to assess sexual health among women with PFDs share common limitations [10]. None of the questionnaires have been validated in a population of women with AI, nor have they been developed within a framework that intentionally included women without a partner or those who did not consider themselves to be sexually active. This may underestimate PFD impact on sexual function since women with severe PFD may elect to become sexually inactive. A validated sexual measure that evaluates not only the impact of PFDs on sexual function but also PFD impact on sexual activity is needed. The questionnaire should have proven responsiveness in evaluating treatment outcomes as well as validity and reliability. This multicenter international study, sponsored by the International Urogynecological Association (IUGA), was designed to address the above limitations in current sexual function measures. We aimed to build upon prior work and establish the validity, reliability, and responsiveness of a revised PISQ, the PISQ- IUGA-revised or PISQ-IR.

Materials and methods

Twenty-three international experts in urogynecology, female sexual function, and survey design and validation convened to evaluate existing sexual function questionnaires and elucidate shortcomings. Three fundamental domains were identified: *sexual inactivity*; *sexual response*; and *quality, satisfaction, and desire*. The PISQ-12 was selected as the foundation for developing a new condition-specific instrument. In addition to the 12 items from the PISQ-12, 48

S. E. Sutherland

Metro Urology, Centers for Continence Care and Female Urology, Center for Pelvic Floor Disorders, Department of Urologic Surgery, University of Minnesota School of Medicine, Minneapolis, MN, USA

E. S. Lukacz

UC San Diego Health System, San Diego, CA, USA

C. Domoney

Chelsea and Westminster Hospital, London, UK

P. Sand

Evanston Continence Center, Evanston Hospital, Northwestern University, Feinberg School of Medicine, Evanston, IL, USA

G. W. Davila

Cleveland Clinic Florida, Department of Gynecology, Section of Urogynecology, Weston, FL, USA

M. E. Espuna Pons

Department of Obstetrics and Gynecology, Hospital Clinic, University of Barcelona, Villarroel 170, Barcelona 08036, Spain

additional items were adopted, revised, or developed for consideration. The basic construct (face) validity of this 60-item pool was evaluated using 31 cognitive interviews [11, 12] conducted at three sites: University of Minnesota ($n=16$), University of New Mexico ($n=5$), and at the Cleveland Clinic, Florida ($n=10$). Based on the results of the cognitive interviews, 18 items were eliminated and 42 items were selected for inclusion in the final item pool.

To increase the diversity of the population used to develop the questionnaire, 12 sites from across the USA and 5 sites from the UK participated in the validation study. Institutional Review Board approval was obtained from all sites, and all women gave written consent. To participate, women were 18 years or older, not pregnant, able to read/write and understand English, and seeking treatment for UI and/or AI and/or POP. Exclusion criteria were a diagnosis of vulvodynia, painful bladder syndrome, or chronic pelvic pain, defined as pelvic pain for greater than 6 months. Since this was a study to evaluate both sexual activity status as well as sexual function, women did not need to be sexually active to participate.

Women were given a survey packet and asked to complete it at home or in the clinic. Due to cost and follow-up capabilities the test-retest packets were only distributed in the USA. To increase response rates, the study followed the principles of the tailored design method [13] in conducting follow-up mailings. To establish the responsiveness of the new questionnaire, a portion of the women who completed the baseline validation survey were mailed a following questionnaire 4–6 months after initial enrollment (Fig. 1).

After written consent, women underwent a physical examination, including the Pelvic Organ Prolapse Quantification Scale (POPQ) [14], muscle strength using the Oxford grading scale [15] and muscle tone. If the woman had these measures assessed within the past month without treatment, she did not undergo repeat examination. Baseline characteristics and past medical history were collected. If the subject was scheduled for surgical treatment, the anticipated surgery date was recorded. Clinicians indicated one or more PFD diagnoses based on assessment of the physical examination findings, history, and any other clinical data available. Definitions conformed to IUGA/International Continence Society (ICS) recommendations (Fig. 1 and Table 1).

In addition to the 42-item validation version of the questionnaire, women also completed the Incontinence Severity Index (ISI) [16], a single question evaluating prolapse and its bother (question #35) from the Epidemiology of Prolapse and Incontinence Questionnaire (EPIQ) [17], the Pelvic Floor Distress Inventory-20 (PFDI-20) [18–20], and the FSFI [5, 21]. The first step in the analysis was to describe the basic distribution of patient and clinical characteristics. Comparisons were made between the UK and US populations using chi-square or *t* tests as appropriate. The test-retest reliability of each item was assessed using Student's *t* tests.

Basic psychometric analytic tools [22–25] were used to guide the evaluation of the instrument. Briefly, we evaluated item distribution and test-retest reliability and then progressed into bi- and multivariate evaluations. In the bi- and multivariate analysis, correlational and factor analysis were the primary statistical tools used. Two independent investigators conducted analyses, using both principle components (PCA) and principle factor analysis (PFA) methods that included orthogonal (varimax) and oblique rotation (promax, Harris-Kaiser) [22–24]. Both analytic methods are commonly used to evaluate the interrelatedness of questions to determine if particular items form coherent and valid subscales.

Factor analysis is a common method used in scale development for instruments such as the PISQ-IR. When a group of items within a scale demonstrates strong relationships to one another then factor analysis identifies this grouping as a factor or subscale. Eigenvalue statistics are used to evaluate the strength of these relationships. A high eigenvalue indicates that an underlying subscale exists; low values indicate that the underlying subscale might not exist. For example, if POP/UI/AI uniquely contributes to sexual life then some items which deal with condition-specific impacts (e.g., I feel sexually inferior because of my incontinence and/or prolapse) should emerge as a subscale. If the condition-specific subscale does not emerge, or if it has a low eigenvalue, then it is assumed that POP/UI/AI does not have a unique role relative to sexual life. If a subscale's eigenvalue meets the threshold for retention, the next step in the analysis is to evaluate how items relate to a particular factor using factor loading scores. A factor loading score represents the strength of the relationship of an item with a particular subscale. In factor analysis each item has a loading score for all candidate subscales. High factor loading scores indicate that an item is strongly associated with that factor; low scores indicate that an item is not strongly associated with that particular factor. Standard criteria for selection and retention of factors (eigenvalue values greater than 1) and items within factors (factor loading scores) were used [22–24]. The majority of items were retained in a subscale if the item had a factor loading of 0.60 and did not load at greater than 0.40 on other factors. For some items, an alternative criterion was used; loading of the item on the primary factor had to be at least 0.20 greater than the items loading on any other factor. If an item meets either the 0.60/0.40 or 0.20 difference criteria then the item was retained as part of a subscale [22–24]. Scale development was iterative and included the full study team who discussed if items in each scale were coherent, made sense, and were clinically useful [26].

In the criterion validity and responsiveness evaluation, correlations served as the primary evaluation technique, but regression was also used to evaluate how our new measure correlated with other proved measures of PFD and sexual function. We used the POPQ, Oxford grading and pelvic floor tone (clinical exam measures), the PFDI-20, ISI, EPIQ

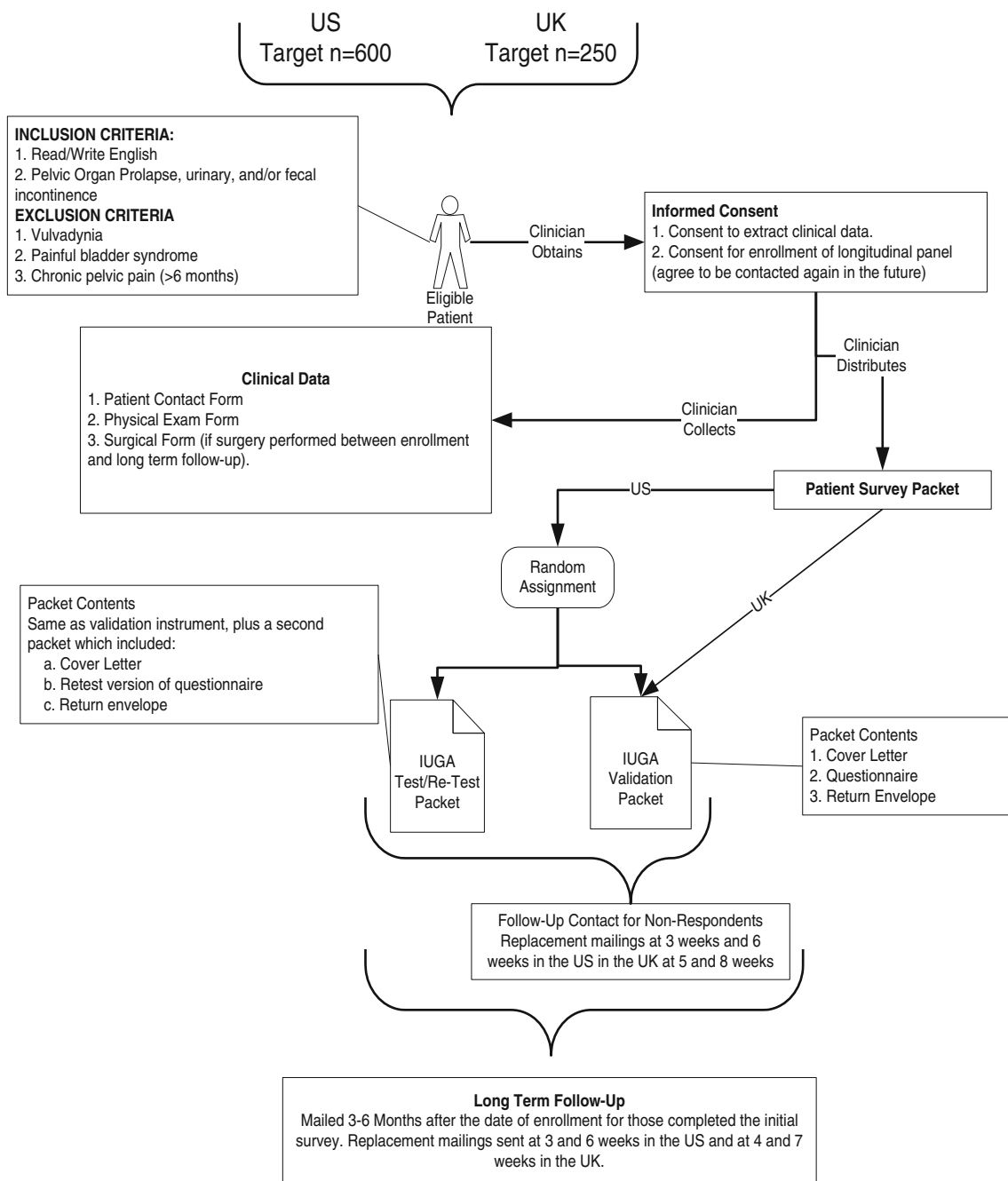


Fig. 1 Study recruitment flow sheet

question #35 (self-reported indicators of condition including severity and impact), and the FSFI (self-reported measure of sexual function). Since the value of criterion analyses was based upon a priori established comparisons, five experts determined which comparisons would be made. Subscales which had condition-specific or related items were compared against all clinical exams and self-reported indicators. For FSFI comparisons, the evaluation focused on the items with sufficient overlap between the PISQ-IR and FSFI

subscales that the reviewers felt were measuring conceptually similar constructs.

Our responsiveness evaluation was based on comparison of the change scores in our new questionnaire to change scores in the PFDI-20, ISI, EPIQ question #35, and FSFI following surgical treatment using correlations. We compared the change scores for women who had surgery during the course of the study with those who did not have surgery (difference of difference test using Student's *t*).

Table 1 Enrollment and follow-up

| | UK, <i>n</i> (%) ^a | USA, <i>n</i> (%) ^a | Total, <i>n</i> |
|--------------------------------------------|-------------------------------|--------------------------------|-----------------|
| Total enrolled | 239 (27) | 638 (73) | 887 |
| Clinical data provided | 234 (98) | 612 (96) | 846 |
| Patient returned completed baseline survey | 162 (68) | 427 (67) | 589 |
| Retest pairs | NA | 147 | 147 |
| Follow-up paired data | 42/91 (46) | 158/305 (52) | 200 |

NA not applicable

^a Figure in percent represents the response rate among those eligible for follow-up mailing; not all women enrolled received follow-up mailings due to the study ending prior to them reaching the required time window for the follow-up survey

Sample size was determined based on the sample size needed to conduct the psychometric analysis, including responsiveness. The basic rule of thumb in psychometric analysis is that ten subjects per item are needed for analyses [23]. For sexually active women we included 32 items, requiring a sample size of at least 320; for women who were not sexually active we included 14 items, requiring a sample of at least 140. To evaluate the responsiveness of the instrument, we estimated that we would need 350 respondents who completed both the baseline as well as a 4- to 6-month follow-up survey. This is based on using an alpha of .05, a targeted power of .80, and an assumed change of 20 % in the score between baseline and follow-up. Adjusting for expected response rates, the targeted enrollment for the study was 850 subjects (600 US and 250 UK subjects). All analyses were conducted using SAS© (v. 9.12). This trial was registered with ClinicalTrials.gov, NCT00952406.

Results

A total of 877 women were enrolled for the study, and 589 returned a completed baseline survey (67 % response rate); 200/288 (70 %) women gave data after surgery and the response rate for the test-retest was 147 (54 %). The majority of women were married, middle-aged (55 ± 12.1 years), and Caucasian. As expected, most women had UI and/or POP; 11 % of the population had AI. Sixty-eight percent of the study population reported sexual activity. A number of differences between the UK and the US populations were noted (Table 2).

The univariate distribution of responses to all items demonstrated acceptable distributions, except for two items with skewed distributions, ceiling effects in items Q12c (shame) and Q12d (fear) (Appendix A). Further evaluation of these items demonstrated that the ceiling effect is driven by the POP-only respondents, but the items demonstrated particular sensitivity in women with AI. As a result of this

sensitivity, the decision was made to retain the items for the rest of the psychometric evaluation. For an evaluation of other issues, such as item nonresponse in relationship to the scaling of the PISQ-IR, see the companion to this article [27].

The test-retest evaluation demonstrated that 3 of the 42 items had significantly different responses between the test and retest administrations (paired *t* test, all $p < 0.05$). We reran the analysis using a correction for the number of tests to determine whether or not the observed difference was secondary to chance (step-down Bonferroni). When corrected for the number of tests none of the items demonstrated significant differences between test and retest. We then evaluated the mean and maximum difference in responses for these three items and the number of respondents who selected a response option that was at least two categories different between test and retest. We found only a few women fell into this category. Finally, we identified and eliminated all respondents who had surgery between the test and retest administrations, assuming that a change in scores following surgery would be expected; when these cases were removed, none of the items demonstrated a significant difference between test and retest. Based on this analysis, all items were retained.

Psychometric analysis

Our original study goal was to pool the data for both sexually active and inactive women for the quality, satisfaction, and desire construct of the questionnaire. During the psychometric analysis, it became clear that the underlying structure and nature of this construct was distinct between groups and pooling was not possible. Thus, the results for the sexually active and inactive groups are presented separately.

Sexually inactive

Two subscales emerged in each of two domains. In the sexual inactivity domain one subscale captured the contribution of PFDs and personal health [not sexually active–condition-specific (NSA-CS), three items], and the other captured partner and personal interest as to why a person is not sexually active [NSA-partner-related (NSA-PR), two items] (see Table 3). The quality and satisfaction domain is composed of two subscales: a global rating of sexual quality [NSA–global quality (NSA-GQ), four items] and a condition-specific subscale [NSA-condition impact (NSA-CI), three items]. Each of the subscales within the two domains demonstrated sound psychometric properties and the items, as demonstrated by their factor loadings, revealed a conceptual grouping that conformed to the underlying assumptions used when developing the instrument (Table 3).

Table 2 Clinical and demographic characteristics of study population

| | % UK participants | % US participants | % Total participants | <i>p</i> |
|-----------------------------------------------|-------------------|-------------------|----------------------|----------|
| Enrolled (number of patients) | 239 | 638 | 877 | |
| Completed baseline survey(number of patients) | 162 | 422 | 584 | |
| Age (mean ± SD) | 52.3±11.7 | 55.8±12.1 | 54.9±12.1 | 0.004 |
| Education—some college or more | 30 | 40 | 38 | 0.01 |
| Relationship status | | | | 0.58 |
| Married/marriage-like relationship | 77 | 76 | 76 | |
| Separated | 2 | 2 | 2 | |
| Divorced | 15 | 14 | 14 | |
| Widowed | 4 | 7 | 6 | |
| Never married | 3 | 2 | 2 | |
| Postmenopausal | 61 | 64 | 63 | 0.37 |
| Caucasian | 70 | 80 | 77 | 0.008 |
| Primary language English | 90 | 96 | 94 | <0.001 |
| BMI (mean±SD) | 27.4±4.4 | 28.3±4.8 | 28.0±4.7 | 0.03 |
| Gravidity | 2.6±1.4 | 3.0±1.6 | 2.9±1.6 | <0.001 |
| Parity | 2.3±1.3 | 2.5±1.2 | 2.4±1.3 | 0.17 |
| All vaginal deliveries | 90 | 85 | 86 | 0.15 |
| Past medical history | | | | |
| Diabetes | 5 | 9 | 8 | 0.06 |
| Neurologic | 1 | 6 | 5 | 0.003 |
| Depression | 5 | 27 | 21 | <0.001 |
| Past surgical history | | | | |
| Hysterectomy and/or BSO | 21 | 34 | 30 | <0.001 |
| Prior prolapse surgery | 15 | 23 | 14 | 0.59 |
| Prior incontinence surgery | 10 | 15 | 13 | 0.06 |
| Not sexually active | 32 | 32 | 32 | 1.0 |
| Underwent surgery | 31 | 34 | 33 | 0.37 |
| Clinical diagnosis | | | | <0.001 |
| POP only | | 35 | 14 | 20 |
| UI only | 36 | 42 | 41 | |
| FI only | 1 | 1 | 1 | |
| POP/UI | 21 | 28 | 26 | |
| POP/FI | 1 | 1 | 1 | |
| UI/FI | 3 | 5 | 5 | |
| POP/UI/FI | 2 | 6 | 5 | |
| POPQ stage | | | | |
| Stage 0 | 55 | 64 | 61 | <0.001 |
| Stage I | 2 | 2 | 2 | |
| Stage II | 17 | 17 | 17 | |
| Stage III | 12 | 15 | 14 | |
| Stage IV | 13 | 2 | 5 | |
| Oxford grading | | | | 0.08 |
| No contraction | 5 | 4 | 5 | |
| Flicker | 8 | 13 | 12 | |
| Weak | 42 | 31 | 34 | |
| Moderate | 30 | 34 | 33 | |
| Good | 14 | 14 | 14 | |
| Strong | 1 | 3 | 3 | |

Table 2 (continued)

| | % UK participants | % US participants | % Total participants | <i>p</i> |
|----------------------|-------------------|-------------------|----------------------|----------|
| Pelvic floor muscles | | | | |
| Normal | 79 | 76 | 77 | 0.59 |
| Overactive | 2 | 1 | 1 | |
| Underactive | 15 | 19 | 18 | |
| Nonfunctioning | 4 | 4 | 4 | |

Percentages are given unless otherwise indicated

BMI body mass index, *BSO* bilateral salpingo-oophorectomy, *FI* fecal incontinence

Sexually active

For the sexual response domain three subscales emerged that included sexual arousal and orgasm [sexually active-arousal, orgasm (SA-AO), four items], partner-related issues [SA-partner-related (SA-PR), three items], and condition-specific issues [SA-condition-specific (SA-CS), three items]

(see Table 4). The quality, satisfaction, and desire domain included one subscale on global quality [sexual activity-global quality (SA-GQ), four items], one on sexual desire [SA-desire (SA-D), three items], and a condition-specific subscale [SA-condition-specific (SA-CS), four items]. Two items did not meet the 0.6/0.4 criteria, but each met the 0.20 criterion and were retained in the subscale.

Table 3 Final factor loadings and internal consistency for scales among women who report sexual inactivity

| Dimension: Sexual inactivity | Final factor loadings | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------|
| Item | Scale #1: not sexually active-condition-specific (NSA-CS) | Scale #2: not sexually active-partner-related (NSA-PR) |
| <u>Condition-specific reasons for sexual inactivity</u> | | |
| The following are a list of reasons why you might not be sexually active, for each one, please indicate how strongly you agree or disagree with it as a reason you are not sexually active | | |
| Due to bladder or bowel problems or due to prolapse (a feeling of or a bulge in the vaginal area) | 0.82 | |
| Because of my other health problems | 0.84 | |
| Pain | 0.83 | |
| <u>Partner-related reasons for sexual inactivity</u> | | |
| No partner | | 0.87 |
| No interest | | 0.56 |
| Dimension: Quality and satisfaction | Scale #3: not sexually active-global quality (NSA-GQ) | Scale #4: not sexually active-condition impact (NSA-CI) |
| <u>Global quality</u> | | |
| Please circle the number between 1 and 5 that best represents how you feel about your sex life | | |
| What best represents how you feel about your sex life: | | |
| satisfied...unsatisfied | 0.92 | |
| adequate...inadequate | 0.87 | |
| How strongly do you agree/disagree: | | |
| "I feel frustrated by my sex life" | 0.82 | |
| Overall, how bothersome is it to you that you are not sexually active? | 0.79 | |
| <u>Condition impact</u> | | |
| How much does the fear of leaking urine and/or stool and/or a bulging in the vagina (either the bladder, rectum, or uterus falling out) cause you to avoid or restrict your sexual activity? | | 0.83 |
| How strongly do you agree/disagree: | | |
| "I feel sexually inferior because of my incontinence and/or prolapse" | | 0.89 |
| "I feel angry because of the impact that incontinence and/or prolapse has on my sex life" | | 0.83 |
| Scale summary statistics | Alpha | Mean score±SD |
| NSA-CS | 0.79 | 2.8±1.0 |
| NSA-PR | * | 2.2±1.0 |
| NSA-GQ | 0.91 | 2.9±1.2 |
| NSA-CI | 0.85 | 2.1±0.1 |
| Total score | 0.83 | 2.9±0.5 |

*Alpha not calculated because there are only two items in this scale

Table 4 Final factor loadings and internal consistency for scales measuring function among women who report sexual activity

| Dimension: Sexual response | | Final factor loadings | | |
|----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|----------------------------------------------------------|------------------------------------------------------|-----------------------------------------|
| Item | Scale #1: Sexually active-arousal, orgasm (SA-AO) | Scale #2: Sexually active-partner-related (SA-PR) | Scale #3: Sexually active-condition-specific (SA-CS) | |
| Arousal/orgasm | | | | |
| How often do you feel sexually aroused during sexual activity? | 0.83 | | | |
| When you are involved in sexual activity how often do you feel fulfilled? | 0.82 | | | |
| Compared to orgasms you have had in the past how intense are your orgasms now? | 0.67 | | | |
| How often do you feel pain during sexual intercourse? | 0.50 | | | |
| Partner-related | | | | |
| How often does your partner have a problem that limits your sexual activity? | | 0.85 | | |
| Does your partner have a positive or negative impact on your sexual desire? | | 0.79 | | |
| Does your partner have a positive or negative impact on the frequency of your sexual activity? | | 0.79 | | |
| Condition-specific | | | | |
| When you are involved in sexual activity, how often do you feel shame? | | | 0.82 | |
| When you are involved in sexual activity, how often do you feel fear? | | | 0.80 | |
| How often do you leak urine and/or stool with any type of sexual activity? | | | 0.68 | |
| Dimension: Quality, satisfaction, and desire | | Scale #4: Sexually active-global quality rating (SA-GQR) | Scale #5: Sexually active-condition impact (SA-CI) | Scale #6: Sexually active-desire (SA-D) |
| Global quality | | | | |
| What best represents how you feel about your sex life (adequate...inadequate) | 0.93 | | | |
| What best represents how you feel about your sex life (satisfied...dissatisfied) | 0.91 | | | |
| How do you feel about your sex life (confident...not confident) | 0.82 | | | |
| "I feel frustrated by my sex life" (strongly agree/strongly disagree) | 0.69 | | | |
| Condition impact | | | | |
| "I feel sexually inferior because of my incontinence and/or prolapse" | | | 0.88 | |
| "I feel angry because of the impact of incontinence and/or prolapse has on my sex life" | | | 0.86 | |
| How much does the fear of leaking urine, stool and/or bulging in the vagina (prolapse) cause you to avoid sexual activity? | | | 0.78 | |
| "I feel embarrassed about my sex life" | 0.41 | | 0.65 | |
| Desire | | | | |
| How frequently do you have sexual desire | | | | 0.81 |
| How would you rate your level (degree) of sexual desire or interest? | | | | 0.79 |
| When you are involved in sexual activity, how often do you feel that you want more? | | | | 0.79 |
| Scale summary statistics | | Alpha | Mean score±SD | |
| SA-AO | | 0.74 | 3.4±0.8 | |
| SA-PR | | 0.77 | 3.2±0.7 | |
| SA-CS | | 0.63 | 4.4±0.7 | |
| SA-GQ | | 0.91 | 3.0±1.1 | |
| SA-CI | | 0.85 | 3.0±0.9 | |
| SA-D | | 0.75 | 3.0±0.9 | |
| Total score | | | 3.3±0.6 | |

Cells in the table in which the factor loading scores are below 0.40 have not been populated

Criterion validity

Sexually inactive

None of the physical exam measures demonstrated significant correlations with the two condition-specific factors (NSA-CS, NSA-CI) for women who reported sexual inactivity. However, the PFDI-20, ISI, and EPIQ question #35 did demonstrate significant correlations in the anticipated direction.

Sexually active

Both condition-specific factors (SA-CS and SA-CI) in both domains demonstrated significant correlations in the anticipated direction with PFDI-20 and ISI scores. EPIQ question #35 and the POPQ correlated with the quality, satisfaction, and desire condition-specific factor (SA-CI) as anticipated. As shown in Table 5, all comparisons with FSFI subscales demonstrate significant correlations in the direction predicted.

This criterion analysis demonstrates that the subscales correlate with external criteria in a manner that was predicted for those who are sexually active. This correlation was found for physical exam measures, condition-specific validated questionnaires as well as the FSFI. For women with AI, strong correlations were found with the FSFI in all factors (Table 5).

Responsiveness

Sexually inactive

The condition impact factor in the quality and satisfaction domain (NSA-CI) demonstrated significant correlation with the EPIQ question #35 and ISI, but not with other measures used to test responsiveness (Table 6).

Sexually active

Improvement in the PFDI-20, ISI, and EPIQ question #35 scores were positively correlated with improvement

Table 5 Criterion validity: scale correlations with other measures

| Sexually active scale correlations | | | | | | | | | | | | |
|-----------------------------------------------------------|-------|---------------------------------------------------|-------|------------------------------------------------------|-------|---------------------------------------------------------|-------|----------------------------------------------------|-------|-----------------------------------------|-------|-------|
| Dimension: Sexual response | | | | | | Dimension: Quality, satisfaction, and desire | | | | | | |
| Scale #1: sexually active-arousal, orgasm (SA-AO) | | Scale #2: sexually active-partner-related (SA-PR) | | Scale #3: sexually active-condition-specific (SA-CS) | | Scale #4: sexually active-global quality (SA-GQ) | | Scale #5: sexually active-condition impact (SA-CI) | | Scale #6: sexually active-desire (SA-D) | | |
| All | AI | All | AI | All | AI | All | AI | All | AI | All | AI | |
| Physical exam measures | | | | | | | | | | | | |
| POPQ | | | | -0.12* | -0.03 | | | 0.17* | 0.02 | | | |
| Pelvic floor tone | | | | -0.02 | 0.18 | | | 0.08 | -0.08 | | | |
| Oxford scale | | | | -0.03 | -0.13 | | | 0.10 | 0.24 | | | |
| Self-report condition-specific measures | | | | | | | | | | | | |
| ISI | | | | 0.29* | NA | | | 0.12* | NA | | | |
| PFDI-20 | | | | 0.34* | 0.08 | | | 0.39* | 0.6* | | | |
| EPIC Q #35 | | | | 0.07 | NA | | | 0.40* | NA | | | |
| Self-report sexual function measures | | | | | | | | | | | | |
| FSFI-desire | 0.56* | 0.56* | | | | | 0.48* | 0.51* | | | 0.86* | 0.86* |
| FSFI-arousal | 0.79* | 0.79* | | | | | 0.58* | 0.63* | 0.42* | 0.42* | | |
| FSFI-lubrication | 0.50* | 0.29 | | | | | | | | | | |
| FSFI-orgasm | 0.75* | 0.77* | | | | | | | | | | |
| FSFI-satisfaction | | | 0.53* | 0.65* | 0.30* | 0.36* | 0.74* | 0.71* | | | | |
| FSFI-pain | 0.55* | 0.53* | | | 0.29* | 0.42* | | | 0.47* | 0.36* | | |
| FSFI-total score | | | | | | | | | | | | |
| Sexually inactive scale correlations | | | | | | | | | | | | |
| Dimension: Sexual inactivity | | | | | | Dimension: Quality and satisfaction | | | | | | |
| Scale #1: not sexually active-condition-specific (NSA-CS) | | | | | | Scale #4: not sexually active-condition impact (NSA-CI) | | | | | | |
| All | | AI | | All | | AI | | All | | AI | | |
| Physical exam measures | | | | | | | | | | | | |
| POPQ | | | | -0.12 | | | | -0.08 | | | | -0.32 |
| Pelvic floor tone | | | | 0.12 | | | | 0.02 | | | | 0.34 |
| Oxford scale | | | | 0.14 | | | | 0.06 | | | | -0.09 |
| Self-report condition-specific measures | | | | | | | | | | | | |
| ISI | | | | 0.21* | NA | | | 0.26* | NA | | | NA |
| PFDI-20 | | | | 0.28* | 0.57* | | | 0.22* | 0.38 | | | NA |
| EPIC Q #35 | | | | 0.27* | NA | | | 0.26 | NA | | | NA |

AI anal incontinence, POPQ Pelvic Organ Prolapse Quantification exam, ISI Incontinence Severity Index, PFDI-20 Pelvic Floor Distress Inventory short form (only the summary PFDI-20 score is used because this scale was developed for use across all PFD conditions and all analyses were conducted with all subjects), EPIC Q #35 analysis limited to only those with a prolapse diagnosis. Epidemiology of Prolapse and Incontinence, question #35, FSFI Female Sexual Function Index, NA not applicable

*Denotes Pearson correlation coefficient with $p < 0.05$

in both condition-specific subscales (SA-CI and SA-CS) in both domains (Table 6). In addition, all comparisons with the FSFI show a significant positive correlation. The two condition-specific subscales show significant differences in scores between women who did and did not have surgery.

Figure 2 illustrates the subscales that emerged from the item pool across three basic domains: sexual response, quality/satisfaction/desire, and sexual inactivity.

The sexual response dimension applies only to individuals who are sexually active and has three subscales; the sexual inactivity dimension applies only to those who are not active and has two subscales. There is overlap for the subscales assessing quality and satisfaction between those who are sexually active and those who are not, but the desire subscale (SA-D) emerged as a coherent subscale only in those who are sexually active.

Table 6 Responsiveness evaluation: change scores correlations

| Sexually Active | | | | | | | |
|--------------------------------------------------------------------------|---------------------------------------------------|---------------------------------------------------|------------------------------------------------------|--------------------------------------------------|----------------------------------------------------|-----------------------------------------|-------------|
| Dimension: Sexual response | | | | Dimension: Quality, satisfaction, and desire | | | Total score |
| Scale | Scale #1: sexually active-arousal, orgasm (SA-AO) | Scale #2: sexually active-partner-related (SA-PR) | Scale #3: sexually active-condition-specific (SA-CS) | Scale #4: sexually active-global quality (SA-GQ) | Scale #5: sexually active-condition impact (SA-CI) | Scale #6: sexually active-desire (SA-D) | |
| Condition-specific measures | | | | | | | |
| ISI | | | 0.36* | | 0.29* | | 0.34* |
| PFDI-20 ^a | | | 0.42* | | 0.47* | | 0.55* |
| EPIC question #35 | | | 0.04 | | 0.16 | | 0.10 |
| FSFI total and subscales | | | | | | | |
| FSFI-desire | 0.36* | | | 0.30* | | 0.75* | |
| FSFI-arousal | 0.49* | | | 0.27* | 0.40* | | |
| FSFI-lubrication | 0.50* | | | | | | |
| FSFI-orgasm | 0.30* | | | | | | |
| FSFI-satisfaction | | 0.48* | 0.33* | 0.41* | | | |
| FSFI-pain | 0.50* | | 0.27* | | 0.31* | | |
| FSFI-total score | | | | | | | 0.47* |
| Surgery | | | | | | | |
| Had surgery (mean difference between scale scores pre- and post-surgery) | 0.05 | 0.10 | 0.35* | 0.36* | 0.43* | 0.11 | 0.26 |

| Sexually Inactive | | | | | | | |
|--------------------------------------------------------------------------|-----------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------|---------------------------------------------------------|-------|---------|-------------|
| Dimension: Sexual inactivity | | | | Dimension: Quality and satisfaction | | | Total score |
| Scale | Scale #1: not sexually active-condition-specific (NSA-CS) | Scale #2: not sexually active-partner-related (NSA-PR) | Scale #3: not sexually active-global quality (NSA-GQ) | Scale #4: not sexually active-condition impact (NSA-CI) | | | |
| Condition-specific measures | | | | | | | |
| ISI | -0.18 | | | | 0.44* | | 0.21 |
| PFDI-20 | 0.17 | | | | 0.24 | | 0.08 |
| EPIC question #35 | -0.34 | | | | 0.60* | | 0.43 |
| Surgery | | | | | | | |
| Had surgery (mean difference between scale scores pre- and post-surgery) | 0.16 | | 0.37 | 0.29 | | 0.38 NS | 0.10 |

A positive correlation indicates improvement in both measures; a negative correlation indicates that the measures moved in opposite directions

ISI Incontinence Severity Index, PFDI-20 Pelvic Floor Distress Inventory-20, EPIC Epidemiology of Prolapse and Incontinence Questionnaire, FSFI Female Sexual Function Index

*Denotes Pearson coefficient correlation significant to $p < 0.05$

^aOnly the summary PFDI-20 score is used because this scale was developed for use across all PFD conditions and all analyses were conducted with all subjects

Questionnaire scoring

Different scoring approaches for all subscales were evaluated, including summated, variants of magnitude and mean scoring [22, 28, 29]. We recommend that one of two methods be used to score the PISQ-IR: mean calculation or a transformed sum. In the calculation of a score for a subscale the respondent must have answered more than one half of the items in each subscale. Missing values should not be imputed. Appendix B gives detailed instruction on how to calculate mean scores. In brief, mean subscale scores are calculated by summing the valid responses to items in the subscale and then divided by the number of items with valid responses. A transformed sum can also be used to score the PISQ-IR; transformed sums demonstrate more accuracy than mean calculation and are further described elsewhere [27]. For sexually inactive women four separate scores are calculated and for sexually active women six separate subscale scores. Total scores are not reported, since the subscales emerged as distinct in the psychometric analyses.

Discussion

We describe the validation and reliability testing of a condition-specific measure of sexual function for women with PFDs and its responsiveness to change. The PISQ-IR improves upon prior sexual function questionnaires used in women with PFDs because it evaluates the effect of PFDs on sexual inactivity and includes validation in women with AI. Although the absolute numbers of women with AI were low, significant findings and correlations were noted. We included a broad population of women with pelvic floor dysfunction in both the USA and UK. Although we did find some baseline differences between groups, we feel that the diversity between the US and UK populations strengthens the generalizability of the PISQ-IR. Finally, this measure is ideally poised for international translation into a variety of languages because the items tested in this validation study were initially vetted by an international panel. Ultimately, we hope that this work provides the groundwork for global validation of the PISQ-IR in a variety of languages and cultural contexts.

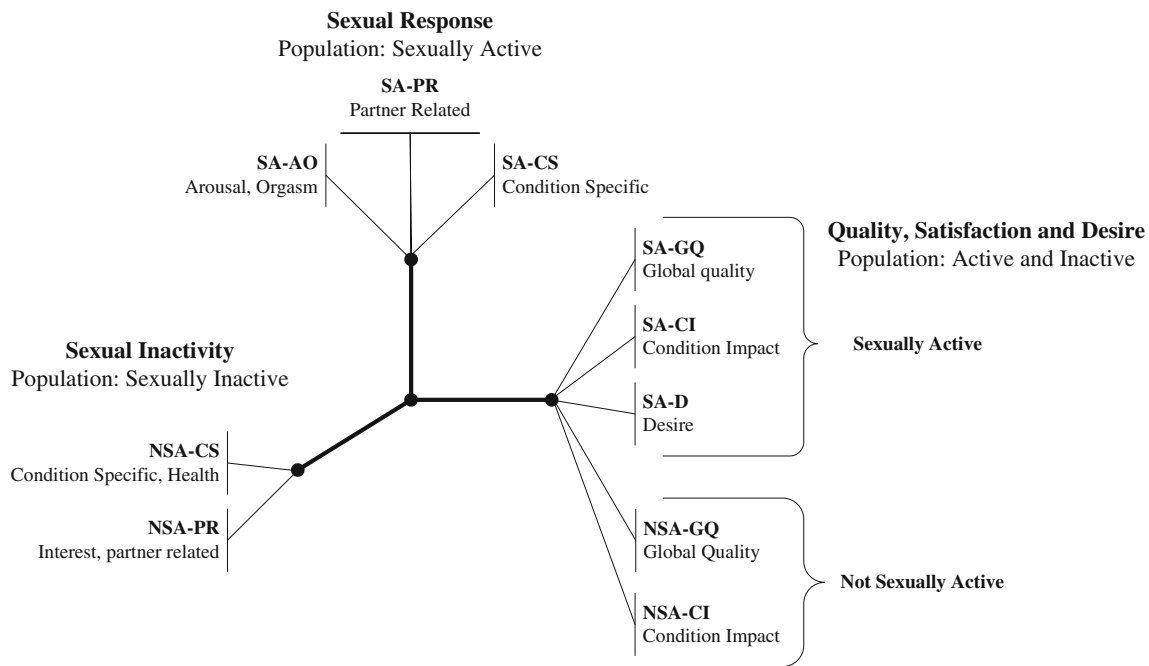


Fig. 2 Final conceptual framework for domains and scales for sexually active and inactive populations resulting from the factor analysis

We acknowledge that patient input into this new measure was relatively limited with 31 cognitive interviews. We did, however, reach saturation in the interviews and believe that we had adequate patient input for the new measure. The psychometric analysis of the PISQ-IR supported the framework that guided the development of the questionnaire (Fig. 2). For each of the domains, multi-item subscales emerged which demonstrated sound psychometric properties. The items in each subscale are internally consistent and form a coherent grouping that is differentiated from the content of the other subscales and between domains. Scaling demonstrated a robust measure in which specific subscales emerged, which assessed core dimensions that are related to general sexual function as well as the impact of PFDs on sexual function. Furthermore, these subscales captured the underlying intent of assessing multiple aspects for each domain including both condition-specific issues and non-condition-specific issues. Initially, the items in the quality, satisfaction, and desire domain were developed to be relevant to both sexually active and inactive women. The psychometric analysis partially supported this assumption, but identified some differences between groups and therefore separate subscales were developed for sexually active and inactive women. Since the domains were distinct for sexually active and inactive women, scores of women who change sexual activity status over time cannot be assessed. For all women, sexual activity status is arguably the most important measure of sexual function, and change of status is an important marker of improved or deteriorated sexual health and should be reported in all studies of sexual function as a separate outcome.

In addition to not being able to measure changes in sexual activity status, the PISQ-IR scoring does not support a single summary score. This is because on psychometric analysis the domains and subscales in the measure emerged as distinct. In order to create a summary score, more patient input is needed to accurately weight the relative importance of various items to respondents. Our new measure represents a significant departure from the original PISQ, although many of the items are similar. For this reason we do not feel that the PISQ-IR scores are comparable with the original questionnaire's scores. Finally, the minimally important difference (MID) for the PISQ-IR has not been established; testing of the instrument in women with pelvic floor dysfunction before and after interventions will be needed to determine the MID.

We found that in the criterion validity testing our new scale correlated significantly in the anticipated direction with other self-report measures. However, the correlations with physical exam measures, including the POPQ, Oxford grading scale, and assessment of pelvic floor muscle tone, were not consistent. We feel that this is representative of how women experience functional disorders which are imprecisely tied to physical exam findings.

We found that the PISQ-IR was responsive to change as measured by correlations with other self-report questionnaires and with changes in PISQ-IR scores following surgery for women who reported sexual activity. While we did not test responsiveness to nonsurgical management, interventions that result in changes similar to surgery are likely to result in similar responsiveness. Responsiveness in condition-specific scales is often evaluated post hoc, in which the scale is developed and

subsequent research is conducted to determine if the scale is responsive. The integration of the responsiveness into the design of this study had significant impact on scale development. When reducing an item pool through psychometric evaluation, items often demonstrated very similar properties. Responsiveness data was used to inform the choice between these items, and we were able to select items that demonstrated greater responsiveness. In addition, establishing responsiveness to change at the time of initial publication allows more confidence in the initial use of the questionnaire. We had aimed to include data from 350 women in our responsiveness evaluations and, despite recruiting the anticipated number of women from the UK and USA, were only able to obtain follow-up data from 200 women. Nonetheless, we were able to demonstrate significant correlations with changes in other self-report measures as well as changes in scores following surgery. In addition, we did not include responsiveness evaluation of women who underwent conservative treatment of pelvic floor dysfunction, since most women who chose conservative management underwent a variety of interventions of varied efficacy. It is possible, but unlikely, that women who respond to nonsurgical treatments and have similar responses in improvement of function will not demonstrate the responsiveness we found among women who underwent surgery.

The responsiveness of the domains evaluating women who reported sexual activity was more robust than the responsiveness of the questionnaire in the sexually inactive group. It was recognized at the start of this project that responsiveness in sexually inactive women would be difficult to evaluate. Nonetheless, we did see that the sexually inactive measures were responsive to change in scores on some of the condition-specific questionnaires.

In conclusion, we have presented the initial validation and reliability data as well as responsiveness of a condition-specific measure of sexual function in women with PFD and that assesses the impact of PFD on sexual inactivity and includes the evaluation of women with AI. Importantly, this measure is now available for international validation in a variety of languages and cultural contexts.

Acknowledgments This study was funded by the International Urogynecological Association. University of Minnesota, Grant Award Number CON000000021500, Todd H Rockwood, Ph.D., PI.

Conflicts of interest R.G. Rogers: consultant (data safety monitoring chair for TRANSFORM trial): AMS; paid travel and paid member: American College of Obstetrics and Gynecology Executive Board; paid travel and consultant: Food and Drug Administration Devices Panel; Editor: International Urogynecology Journal; Editorial Board Member: Journal of American College of Obstetrics and Gynecology. F. Reid: paid travel and consultant: Pfizer; paid travel: Astellas, Gynecare. E.S. Lukacz: consultant: Pfizer; paid travel and paid research: Renew Medical; paid research and consultant: Ethicon/J and

J. C. Domoney: paid travel: Pfizer, Astellas; honoraria: Novo Nordisk. P. Sand: paid speaker and consultant: Allergan, Astellas, Pfizer, Watson, Ferring, Merck, Uroplasty; consultant: Teva, Targacept; paid research: Boston Scientific, Contura, EM Kinetics. G.W. Davila: paid travel and consultant: AMS; consultant: Coloplast; paid travel, consultant and paid research: CL Medical; paid travel and consultant: Astellas; paid travel: Warner-Chilcott.

T.H. Rockwood, M. Constantine, R. Thakar, D.N. Kammerer-Doak, R.N. Pauls, M. Parekh, B. Ridgeway, S. Jha, J. Pitkin, S.E. Sutherland, M. Espuna: none.

References

1. Defining sexual health: report of a technical consultation on sexual health, 28–31 January 2002 (2006) World Health Organization, Geneva
2. Rogers RG, Villarreal A, Kammerer-Doak D, Qualls C (2001) Sexual function in women with and without urinary incontinence and/or pelvic organ prolapse. *Int Urogynecol J Pelvic Floor Dysfunct* 12(6):361–365
3. Pauls RN, Segal JL, Silva WA, Kleeman SD, Karram MM (2006) Sexual function in patients presenting to a urogynecology practice. *Int Urogynecol J Pelvic Floor Dysfunct* 17(6):576–580
4. Brock G, Nicolosi A, Glasser DB et al (2002) Sexual problems in mature men and women: results of a global study. *Int J Impot Res* 14(Suppl 3):S57–S58
5. Rosen R, Brown C, Heiman J et al (2000) The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther* 26(2):191–208
6. Rust J, Derogatis L, Rodenberg C, Koochaki P, Schmitt S, Golombok S (2007) Development and validation of a new screening tool for hypoactive sexual desire disorder: the Brief Profile of Female Sexual Function (B-PFSF). *Gynecol Endocrinol* 23(11):638–644
7. McCoy NL, Matyas JR (2000) McCoy female sexuality questionnaire. In: Davis CM, Yarber WL et al (eds) *Handbook of sexuality-related measures*. Sage, London, pp 249–251
8. Rogers RG, Kammerer-Doak D, Villarreal A, Coates K, Qualls C (2001) A new instrument to measure sexual function in women with urinary incontinence and pelvic organ prolapse. *Am J Obstet Gynecol* 184:552–558
9. Rogers RG, Coates KW, Kammerer-Doak D, Khalsa S, Qualls C (2003) A short form of the Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-12). *Int Urogynecol J Pelvic Floor Dysfunct* 14(3):164–168, discussion 168
10. Omotosho TB, Rogers RG (2009) Shortcomings/strengths of specific function questionnaires currently used in urogynecology: a literature review. *Int Urogynecol J Pelvic Floor Dysfunct* 20(Suppl 1):S51–S56
11. Willis GB (2005) *Cognitive interviewing: a tool for improving questionnaire design*. Sage, Thousand Oaks
12. Schwarz N, Sudman S (1996) *Answering questions: methodology for determining cognitive and communicative processes in survey research*. Jossey-Bass, San Francisco
13. Dillman DA (2007) *Mail and Internet surveys: the tailored design method*. Wiley, Hoboken
14. Bump RC, Mattiason A, Bø K, Brubaker LP, DeLancey JOL, Klarskov P et al (1996) The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol* 175:10–17
15. Laycock J (1995) *Pelvic floor dysfunction*. Ph.D. thesis, University of Bradford, UK
16. Sandvik H, Hunskaar S, Seim A, Hermstad R, Vanvik A, Bratt H (1993) Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. *J Epidemiol Community Health* 47(6):497–499

17. Lukacz ES, Lawrence JM, Buckwalter JG, Burchette RJ, Nager CW, Luber KM (2005) Epidemiology of prolapse and incontinence questionnaire: validation of a new epidemiologic survey. *Int Urogynecol J Pelvic Floor Dysfunct* 16(4):272–284
18. Barber MD, Kuchibhatla MN, Pieper CF, Bump RC (2001) Psychometric evaluation of 2 comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. *Am J Obstet Gynecol* 185(6):1388–1395
19. Barber MD, Walters MD, Bump RC (2005) Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). *Am J Obstet Gynecol* 193(1):103–113
20. Barber MD, Walters MD, Cundiff GW, PESSRI Trial Group (2006) Responsiveness of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) in women undergoing vaginal surgery and pessary treatment for pelvic organ prolapse. *Am J Obstet Gynecol* 194(5):1492–1498
21. Meston CM (2003) Validation of the Female Sexual Function Index (FSFI) in women with female orgasmic disorder and in women with hypoactive sexual desire disorder. *J Sex Marital Ther* 29:39–46
22. Nunnally JC, Bernstein IH (1994) *Psychometric theory*. McGraw-Hill, New York
23. Pett MA, Lackey NR et al (2003) *Making sense of factor analysis: the use of factor analysis for instrument development in health care research*. Sage, Thousand Oaks
24. Gorsuch RL (1974) *Factor analysis*. Saunders, Philadelphia
25. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16:197–334
26. Armstrong JS (1967) Derivation of theory by means of factor analysis or Tom Swift and his electric factor analysis machine. *Am Stat* 21(5):17–21
27. Rockwood TH, Constantine ML, Adegoke O, Rogers RG, McDermott E, Davila GW, Domoney C, Jha S, Kammerer-Doak D, Lukacz ES, Parekh M, Pauls R, Pitkin J, Reid F, Ridgeway B, Thakar R, Sand PK, Sutherland SE, Espuna-Pons M (2013) The PISQ-IR: considerations in scale scoring and development. *Int Urogynecol J*. doi:10.1007/s00192-012-2037-z
28. Spector PE (1992) *Summated rating scale construction: an introduction*. Sage, Newbury Park
29. Lodge M (1981) *Magnitude scaling, quantitative measurement of opinions*. Sage, Beverly Hills