

ORIGINAL RESEARCH—OUTCOMES RESEARCH

The Sensitivity of Event Logs, Self-Administered Questionnaires and Photoplethysmography to Detect Treatment-Induced Changes in Female Sexual Arousal Disorder (FSAD) Diagnosis

Alessandra Rellini, MA, and Cindy Meston, PhD

Department of Psychology, University of Texas—Austin, Austin, TX, USA

DOI: 10.1111/j.1743-6109.2005.00153.x

ABSTRACT

Introduction. The literature provides little evidence on what type of endpoints should be used to assess treatment-induced improvement in female sexual function.

Aim. The main goal of this study was to provide empiric evidence on the sensitivity of different types of measures for detecting treatment-induced changes in female sexual dysfunction diagnosis.

Methods. The measures investigated in this study included event logs, self-administered questionnaires (Female Sexual Functioning Index; FSFI), vaginal photoplethysmography, and continuous subjective sexual arousal measured during exposure to erotic videos. Participants were 24 women with female sexual arousal disorder (FSAD) who received sex therapy, placebo, or ginkgo biloba in a four-arm double-blind placebo-controlled clinical trial. FSAD was diagnosed utilizing a semistructured interview administered at pre- and post-treatment. Those women who did not meet FSAD criteria at post-treatment (N = 10) were labeled as “improved,” while women who still met FSAD criteria (N = 14) were categorized as “not improved” even if they showed signs of improvements.

Main Outcome Measure. Change scores from pre- to post-treatment on the FSFI, event logs, vaginal photoplethysmography, and continuous subjective levels of sexual arousal were used to predict whether women improved at post-treatment. Results were checked with exact logistic regression to control for the small sample size.

Results. The FSFI was the only measure to significantly predict whether women improved at post-treatment. The findings from this study lend support for the use of validated questionnaires as endpoint criteria in detecting treatment-induced changes in women’s sexual dysfunction. **Rellini A, and Meston C. The sensitivity of event logs, self-administered questionnaires and photoplethysmography to detect treatment-induced changes in female sexual arousal disorder (FSAD) diagnosis. J Sex Med 2006;3:283–291.**

Key Words. Design; Methodology of Clinical Trials; Psychological Assessment of Sexual Dysfunction History; Physical; Examination Diagnostic Testing; Psychophysiological Studies of Sexual Function

Introduction

In 2003, the International Consultation on Urological Diseases (ICUD), an organization registered with the World Health Organization (WHO), met in Paris and provided an in-depth review of published evidence on the efficacy of treatments for women’s sexual dysfunction. Evidence from this 2nd International consultation on

sexual dysfunctions revealed that the field of female sexuality is lacking in empirically supported treatments for almost all types of sexual dysfunction except primary female orgasmic disorder [1]. One of the impediments of conducting clinical trials on female sexuality is the controversy regarding what endpoints most effectively assess treatment-induced changes [2]. Currently, there are a variety of measures that can be used as end-

points, specifically, self-administered questionnaires, event logs or diaries, physiological measures of sexual arousal, and clinician-administered interviews. While each method has pros and cons that need to be considered, it is important to acknowledge that this discussion is often framed within the context provided by the Food and Drug Administration (FDA) [3] guidelines for clinical trials on female sexuality. These guidelines recommend using diaries or event logs as the primary endpoints while self-administered questionnaires are accepted only as secondary outcome variables:

Primary endpoints for trials of drug products to treat FSD should be clinically meaningful and specifically related to the component or components of FSD being studied in the trials. These endpoints should be based on the number of successful and satisfactory sexual events or encounters over time. The determination of *successful* and *satisfactory* should be made by the woman participating in the trial, as opposed to her partner. [3]

Researchers have argued that although event logs and diaries can provide meaningful information on the patient's condition, self-administered questionnaires can be equally or more informative than event logs.

The characteristics necessary for an outcome measure are reliability, validity, and sensitivity to detect treatment-induced changes. Reliability is usually assessed within items and within participants over a period of time. Within-item correlations provide information on whether the instrument is assessing a consistent construct, while test-retest reliability allows researchers to assess whether scores are consistent across time. Validity tests include construct, concurrent, divergent, and discriminatory validity. Construct validity is tested statistically using principal component analyses to investigate the consistency of scores for items intended to address the same domain. Convergent and divergent validity tests how similar or different scores of the target measure are as compared with other measures known to address related constructs. This analysis is important to assess whether the instrument is providing additional information as compared with existing scales and whether it correlates with constructs that are supposed to be associated. The discriminatory validity of an outcome measure is the ability of the measure to distinguish between women with and without a clinical diagnosis. Often this is tested by conducting a clinical interview and comparing group differences on the scores from the measure studied. Sensitivity is defined as the abil-

ity of the measure to detect treatment-induced changes. This analysis can be conducted using a logistic or linear regression depending on the nature of the dependent variable (i.e., categorical or continuous).

The available measures for female sexual function are self-administered questionnaires, event logs, physiological measures of sexual arousal, and interviews. In a review of validated self-administered questionnaires, Meston and Derogatis [4] cited five instruments developed to assess sexual function. The advantage of self-administered questionnaires for female sexual function over other types of measures is the large amount of empiric support for their reliability and validity. The Female Sexual Function Index (FSFI) [5] is one of the most currently and widely used questionnaires in clinical and research settings and is composed of 19 items divided into six domains: desire, arousal, lubrication, orgasm, satisfaction, and pain. This scale was first developed on a sample of 18- to 65-year-old heterosexual women in relationships with and without a diagnosis of female sexual arousal disorder (FSAD). The FSFI takes only 15 minutes to complete, which makes it appealing for large multisite clinical trials. One of the limitations of the FSFI is the lack of information on the sensitivity of this instrument.

Event logs, according to the FDA guidelines, are to be completed by the participant after each sexual event and they are to assess frequency of *successful* and *satisfactory* sexual events. Currently, there are no published psychometrics on any event log, although several are in the making [2]. One of the benefits of using event logs in clinical trials is the generalizability of the results to the life of the individual. Being able to become aroused more easily may be useful but it becomes meaningless if taken outside of the context of the sexual relationship of the patient. On the other hand, there are some limitations to event logs that need careful considerations. One of the main concerns raised by Althof et al. [6] is the lack of concordance between event logs and self-administered questionnaires [7], which is accompanied by a lack of sensitivity to treatment-induced changes, and a lack of compliance [7,8]. Althof et al. [6] make the case that asking women to indicate whether the sexual event was successful and satisfactory is based on a subjective judgment that may be better captured by a questionnaire which addresses a variety of dimensions of satisfaction. A sexual event also may not be satisfactory for reasons other than levels of desire, sexual

arousal, and orgasm, thus may be removed from the primary objectives of the treatment. Further concerns were raised on the preference for a measure based on two categories (satisfactory event or nonsatisfactory event) as compared with a measure based on a more sophisticated scale such as ratio scale (Likert scale). However, Althof et al. [6] point to a lack of sufficient objective evidence to either retain or reject event logs as primary endpoints.

Physiological measures of sexual arousal are currently used in a variety of laboratories. The most commonly used method to assess physiological sexual arousal is vaginal photoplethysmography, an indirect index of vaginal engorgement during exposure to erotic stimuli. This measure was first invented by Sintchak and Geer [9] and has since been used in a number of studies on female physiological sexual arousal. Vaginal pulse amplitude (VPA), the unit of physiological sexual arousal, has been found to increase specifically after exposure to erotic stimuli [10]. Exposure to anxiety-provoking videos was not found to increase VPA, thus testing to the sensitivity of the measure. VPA does not have an absolute zero, thus changes in VPA to erotic stimuli have been measured in comparison with the VPA during exposure to a neutral video. One major limitation of this instrument is the large unexplained variability between participants. While some women may show a baseline VPA of 1 mV, others may have 50 mV or larger baselines. Attempts to minimize variance within participants have been addressed by using a placement bar to help standardize the device's position. Studies have consistently failed to find a meaningful difference in VPA response between women with and without FSAD [11,12]. Recently, Brotto et al. [13] found that when women were divided into subcategories of FSAD (i.e., subjective, genital, and combined) as recommended by an international panel of experts [14], significant differences between women with genital FSAD vs. controls and women with subjective or combined FSAD emerged. Meston et al. (unpublished data) also recently noted significant differences between subgroups of women with FSAD in VPA responses to erotic stimuli. This suggests that perhaps VPA has discriminant validity when considering a women with FSAD that is primarily of physical (i.e., genital) origin. One of the appeals of VPA is the objectivity of the measure, which is also one of its limitations because a physiological change induced by treatment may be meaningless

if not accompanied by a subjective reduction in distress. Other limitations of photoplethysmography are its invasive nature which may not be appealing to all women; the time burden associated with the procedure; and the financial cost of running physiological visits. On the other hand, photoplethysmography could provide invaluable data on the mechanisms of change which could contribute to developing empirically validated treatments for women with FSAD in particular.

Interviews are potentially rich methods to acquire information on the condition of the patient. To our knowledge, only one structured interview has been published for the assessment of female sexual dysfunction [15]. Although interviews provide probably the most in-depth and accurate information on the patient's sexual function, they are also time-consuming and intrainterviewer reliability can become problematic in large multisite clinical trials [2].

The purpose of this study was to determine which assessment method, for example, event logs, self-administered questionnaires, or psychophysiological measures, best predicts treatment outcome. To do this, we re-analyzed the data from a four-arm double-blind placebo-controlled clinical trial that compared the efficacy of sex therapy, placebo, and ginkgo biloba. Detailed results from this study are presented elsewhere (C.M. Meston & A.H. Rellini, unpublished data). Clinical interviews conducted at pre- and post-treatment were used as the gold standard to detect changes in FSAD diagnosis after treatment. Event logs, a validated self-administered questionnaire (FSFI), and both self-report and physiological data during exposure to erotic stimuli in the laboratory were compared as to their ability to detect changes in diagnosis at post-treatment.

Method

Participants

The original study enrolled 122 women with a variety of sexual dysfunctions. For the purpose of this study we selected only those women with a diagnosis of FSAD (N = 44) and of these, we used data only for women who completed the event logs at pre- and post-treatment (N = 24). Participants were recruited through advertisements placed in a free community newspaper and fliers posted in Ob/Gyn clinics. During an initial phone call, potential participants were asked a series of questions to assess their qualification for the study and were explained the aims and logistics of the

study. Women were included in the study if they were over 18 years of age, currently in a committed relationship, not taking any beta-blockers, reported subjective distress in relation to a problem with sexuality, and denied any abusive sexual encounter during the previous 2 years. Exclusion criteria included presence of domestic violence, menopausal or peri-menopausal status, problems with severe nerve damage, any Axis I diagnosis (*Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, Text Revised [DSM-IV-TR]) [16], drug or alcohol abuse, and organic psychological problems (e.g., schizophrenia or psychoses). After data collection, participants were divided into a group of women who at post-treatment no longer met DSM-IV-TR [16] criteria for FSAD ($N = 14$) or a second group that still met diagnostic criteria for FSAD at post-treatment ($N = 10$).

Measures

Interview

A standardized interview was used to assess changes in DSM-IV-TR [16] diagnosis for FSAD. The interviewer was an advanced clinical student with over 300 hours of experience conducting clinical interviews in the area of sexual functioning. The interviewer explained that for the sake of the study "arousal" was defined as any physical or mental experience of being "turned on." Examples of physical sensations provided to the participants included changes in vaginal lubrication, feeling flushed, experiencing increased heart beat or breathing. The interviewer continued by asking whether the participant "experienced difficulties in becoming sexually aroused or 'turned on,' when receiving the appropriate amount of stimulation," and whether "this lack of arousal or difficulty in becoming aroused or 'turned on' was creating problems for her or for her relationship." The interviewer used a number of prompts to further investigate the nature and severity of the dysfunction. In case the participant reported that the sexual arousal problem was due to sexual dysfunction experienced by the partner or relational difficulties her data were no longer considered for this study. Participants who reported a subjective distress associated with the lack of, or difficulties with, sexual arousal were classified as women with FSAD. If the participant continued to report distress associated with her difficulty to become sexually aroused at post-treatment she was categorized as "no change" even if she reported that her condition had improved.

Self-Report Questionnaire

The self-administered questionnaire selected for this study was the FSFI [5]. The FSFI is a 19-item questionnaire that has shown good reliability, validity, and most importantly has been shown to have adequate sensitivity and specificity in the discrimination of women with and without different sexual dysfunction. The scale is divided into six domains: desire, arousal, lubrication, orgasm, satisfaction, and pain. The arousal subscale has shown to discriminate women with no sexual dysfunction from women with FSAD [5], female orgasmic disorder (FOD), or hypoactive sexual desire disorder (HSDD) [17,18]. Difference scores were computed between the FSFI arousal and FSFI satisfaction domains completed before and during the post-treatment visit.

Event Log

The event log used in this study was developed in agreement with the FDA guidelines [3] and comprised questions regarding sexual satisfaction (How satisfied were you with your sexual arousal during this sexual activity?) and ability to become mentally and physically sexually aroused (How easy was for you to experience physical sexual arousal from the sexual activity? How easy was for you to feel mentally "turned on" during this sexual activity?). The potential answers consisted of a six-point Likert scale and, for the sake of following the FDA conceptualization of event logs, a score of "moderately satisfied" or "somewhat easy" or higher was calculated as a positive experience. Participants were asked to complete the log after each sexual encounter. They were asked to indicate whether the sexual activity included the partner or not. Only sexual activities with a partner were used in the analysis. The ratio of satisfying sexual encounters, mental sexual arousal, and physical sexual arousal was computed for 2 weeks prior to the treatment and for the last 2 weeks of treatment. The difference scores between the ratio at post- and pretreatment in satisfaction, mental sexual arousal, and physical sexual arousal were used as predictors for the regression on changes in sexual dysfunction at post-treatment.

Physiological Sexual Arousal

Changes in VPA between exposure to a neutral and erotic video were calculated. The VPA was measured with a vaginal photoplethysmograph the participant inserted on her own while alone in a private room. A placement bar was clipped on the photoplethysmograph to ensure standardization

of the depth and the position of the light sensor in relation to the vaginal walls. Participants were instructed to try to remain still during the exposure to the videos. The video sequence started after a 10-minute habituation period and lasted 3 minutes (neutral video) and 10 minutes (erotic video). The photoplethysmograph was connected to an analog/digital data signal coder and signal samples were selected at a rate of 80 samples/second. Movement artifacts were edited after data collection according to the procedure commonly used with this type of data [11,19]. The pulse amplitude was measured in distance between peak and trough (mV) of each pulse and averaged across neutral and erotic video. Percentage of VPA change from neutral to erotic videos was used as a measure of physiological sexual response during that session. The percentage change in VPA during the post-treatment visit was subtracted from the percentage of VPA change during the pretreatment visit and this variable was used as an indication of changes in physiological sexual arousal from pre- to post-treatment.

Subjective Sexual Arousal

Self-reported sexual arousal to the erotic videos was measured continuously using a device termed "arousometer" that was developed by the Female Sexual Psychophysiology Laboratory at the University of Texas at Austin [20]. The arousometer is composed of a computer mouse mounted on a wooden track that the participant moves from a score of "0" or "not aroused at all," to "7" or "very aroused." This technique has been shown to be a valid measure of subjective levels of sexual arousal during laboratory exposure to erotic videos in that it correlates highly with levels of mental sexual arousal as measured by a self-administered questionnaire completed post videos, $\beta = 0.01$, $t = 5.88$, $P < 0.006$ [20]. Moreover, levels of mental sexual arousal have been found to be significantly lower in women with FSAD and FOD as compared with women with no sexual dysfunction (C.M. Meston, et al., unpublished data). For each psychophysiological assessment session the average scores during the erotic video were subtracted from the average scores during the neutral video. Change scores were operationalized as the difference between the subjective sexual response during the post-treatment and the pretreatment visits.

Procedure

Participants completed two pretreatment and one post-treatment assessment visits during which

time they completed questionnaires, interviews and a physiological (i.e., vaginal photoplethysmography) assessment. The videos used for the physiological assessment were selected from a video library composed of six videos which have shown to comparably increase physiological and subjective sexual arousal in a sample of 66 women (more information on the videos' validation is available from the authors). To reduce potential sample bias, pre- and post-treatment videos were counterbalanced between participants. Participants were provided with event logs and instructed to complete a log each time they engaged in a sexual activity for 2 consecutive weeks pretreatment and during the last 2 weeks of treatment (post-treatment). At the end of each week they were contacted by the experimenter and asked to send the completed event logs to the Sexual Psychophysiology Laboratory. Treatment lasted 8 weeks and consisted of either: (i) 300 mg of ginkgo biloba daily; (ii) placebo pill daily; (iii) weekly sex therapy; or (iv) weekly sex therapy plus 300 mg of ginkgo daily.

Results

Participants

Women who showed post-treatment improvement in FSAD and women who continued to have FSAD post-treatment did not differ in pretreatment measures of sexual dysfunction, relationship satisfaction, and demographic characteristics (see Table 1).

Validity of Instruments

Interview Validity

The interview was originally administered to 41 women diagnosed with FSAD and 80 women diagnosed with either HSDD or FOD. To confirm that the interviews were valid, we used the FSFI scores in the arousal and lubrication domains to assess discriminant validity. Results from these t -tests showed a significant difference in the arousal ($t(120) = 2.55$, $P < 0.01$), lubrication ($t(120) = 3.05$, $P < 0.01$), and satisfaction ($t(120) = 3.18$, $P < 0.01$) domains between women with FSAD and women with HSDD or FOD. Moreover, participants diagnosed with FSAD showed means very similar to the ones reported by Rosen et al. [5] and the means of women with other types of diagnoses (HSDD or FOD) showed means more similar to the norms shown in women with HSDD and FOD diagnoses [17]. Table 2 illustrates the means observed among

Table 1 Demographics divided by the presence/absence of FSAD at post-treatment

	FSAD at post-treatment (N = 14) Mean (SD)	No FSAD at post-treatment (N = 10) Mean (SD)
Age	28.7 (10.3)	26.7 (4.7)
Length of relationship	2.6 (1.3)	3.3 (0.9)
Education	3.1 (1.3)	3.6 (1.2)
	N (%)	N (%)
Marital status		
Single	7 (50.0)	5 (50.0)
Married	4 (28.6)	4 (40.0)
Divorced	3 (21.4)	1 (10.0)
Sexual orientation		
Bisexual	2 (14.3)	1 (10.0)
Heterosexual	12 (85.7)	9 (90.0)
Ethnicity		
Caucasian	11 (78.6)	4 (40.0)
African American	0 (0.0)	1 (10.0)
Asian American	0 (0.0)	1 (10.0)
Hispanic	1 (7.1)	4 (40.0)
Other	1 (7.1)	0 (0.0)
More than one	1 (7.1)	0 (0.0)

Length of relationship is scored on a point Likert scale (0 = 0–6 months; 1 = 6–12 months; 2 = 1–2 years; 3 = 3–5 years; 4 = 5–10 years; 5 = more than 10 years). Education was scored on a 0–6 point Likert scale (0 = some HS; 1 = HS diploma or GED; 3 = 2 years of college, 4 = 4 years of college; 5 = advanced degree). HS = high school; GED = general education diploma.

women in this study and women used for the standardization of the FSFI.

Event Logs Validity

In an attempt to provide some initial validity on the event logs, we used the data from all women with FSAD diagnoses in the sample that completed the diaries at pretreatment (N = 42). A Pearson correlation coefficient was computed between each of the domains of the FSFI and the measures from the event logs at pretreatment to test for convergent validity. We found a significant correlation between the question on sexually satisfying encounters and the FSFI arousal ($r(41) = 0.461$, $P < 0.01$), lubrication ($r(41) = 0.393$, $P < 0.01$), and orgasm ($r(41) = 0.467$, $P < 0.01$) domains. The questions in the event logs

about physical sexual arousal were moderately associated with the FSFI arousal domain ($r(41) = 0.328$, $P < 0.05$), and the event logs question on mental sexual arousal was strongly associated with the FSFI arousal domain ($r(41) = 0.459$, $P < 0.001$) and moderately associated with the FSFI lubrication domain ($r(41) = 0.300$, $P < 0.05$), orgasm domain ($r(41) = 0.313$, $P < 0.05$), and the total FSFI score ($r(41) = 0.338$, $P < 0.05$). This suggests that the questions on sexual satisfaction, and mental and physical sexual arousal from the event logs were moderately and significantly correlated with the domains of sexual arousal, lubrication, and orgasm as measured with the FSFI. The test–retest reliability of the event logs remains unknown at this point. It is conceivable that, given the simplicity of the questions, the instrument has good face validity.

Predicting Treatment-Induced Changes in FSAD Diagnosis

To test whether the different types of measures were able to predict changes in FSAD from pre- to post-treatment, the chosen outcome variable of a series of logistic regressions was change in FSAD diagnoses from pre- to post-treatment (1 = improved; 0 = not improved) assessed via the clinical interview. Given that the goal of this study was to assess how changes in clinical diagnoses were detected through changes in different measures of sexual function, we used as predictors the change scores from pre- to post-treatment in event logs dimensions (satisfaction, mental sexual arousal, and physical sexual arousal), FSFI arousal domain, VPA, and continuous subjective arousal. As the sample size was small and logistic regression is sensitive to sample size, after computing the analyses with regular logistic regression we checked the results with a new analytical technique that controls for small samples size, exact logistic. This technique uses a Monte Carlo simulation, a method of analysis based on artificially recreating a chance process, which has been shown to be

Table 2 Mean and SD for the FSFI domains in women diagnosed with FSAD or with FOD and HSDD

	FOD+HSDD Mean (SD)	FOD and HSDD (Meston, 2003) ¹⁹ Mean range (SD range)	FSAD Mean (SD)	FSAD (Rosen et al., 2000) ⁵ Mean (SD)
Desire	5.2 (2.2)	5.2–5.9 (1.4–2.0)	4.4 (2.1)	4.7 (2.1)
Arousal	11.5 (4.5)	12.0–12.7 (4.0–4.1)	9.4 (4.1)	9.7 (4.8)
Lubrication	14.2 (4.7)	14.5–14.6 (5.4–5.4)	11.4 (4.8)	10.9 (5.5)
Orgasm	6.7 (3.6)	7.3–8.5 (3.6–4.1)	5.7 (3.3)	7.1 (4.1)
Satisfaction	9.1 (3.4)	9.7–9.7 (3.0–3.4)	7.2 (2.5)	8.2 (3.6)
Pain	12.4 (3.5)	13.0–13.1 (2.8–3.1)	11.2 (3.4)	10.1 (4.6)

FSFI scores of women diagnosed with FOD+HSDD and with FSAD by the interview utilized in the study were compared with FSFI norms published on women with HSDD and FOD (Meston, 2003)¹⁹ and women with FSAD (Rosen et al., 2000).⁵

Table 3 Results for individual logistic regressions used to predict treatment-induced change in FSAD diagnosis at post-treatment

	Correct FSAD predictions at post-treatment [†]		β	SE(B)	Wald χ^2
	FSAD N (%)	No FSAD N (%)			
Event logs	11 (78.6)	3 (30.0)			1.45
Satisfaction			0.818	1.21	0.03
Mental arousal			0.833	1.30	0.02
Physical arousal			3.976	1.30	1.13
FSFI arousal	22 (88.0)	6 (42.9)	0.189	1.21	4.61*
Psychophysiology	12 (80.0)	8 (66.7)			6.35*
VPA			-0.057	1.04	0.30
Continuous subjective			-0.213	0.94	3.26

* $P < 0.05$.[†]The frequencies and percentages reported in the table correspond to the number of participants that the model correctly detected. All frequencies and percentages are based on the entire FSAD (N = 14) and No FSAD (N = 10) sample.

accurate with samples of 20 participants or more. The results found using the regular statistics were confirmed by the exact logistic algorithm; thus we report here only the results from the standard logistic regressions (see Table 3).

First, we conducted a single logistic regression that included as separate predictors event logs (satisfaction, mental sexual arousal, and physical sexual arousal), FSFI arousal domain, VPA, and continuous subjective sexual arousal. We combined VPA in the same regression model with continuous subjective sexual arousal because the two measures were measured in the laboratory concurrently. Of the three models, the FSFI arousal domain and the VPA and continuous subjective sexual arousal measures were shown to predict changes in FSAD diagnosis above and beyond chance. As illustrated in Table 3, changes in FSFI scores accurately predicted women who did not change diagnoses in 88% of the cases and women who did change in 42.9% of the cases, for a total of 71.8% of correct prediction. The odds ratio to correctly predict change in diagnoses by considering only the arousal domain of the FSFI were 1.2, meaning that the odds of being classified correctly were 20% greater than been classified incorrectly after considering FSFI arousal domain. The VPA and continuous subjective model correctly predicted 80% of participants who still had FSAD and 67% of women who no longer had FSAD at post-treatment. However, the VPA data failed to converge which is problematic with the Monte Carlo simulation; thus the results for VPA need to be interpreted with caution.

Given that the FDA proposed event logs as the gold standard on which to base outcome measures, we conducted a series of hierarchical logistic regressions where we compared the performance of event logs to other measures (i.e., FSFI, VPA,

and continuous subjective sexual arousal) on their ability to correctly predict changes in FSAD diagnosis. To conduct this test we used a two-step hierarchical regression where event logs change scores in satisfaction, mental sexual arousal, and physical sexual arousal were the predictors used in step one to predict changes in FSAD diagnoses. In step two, either FSFI or VPA and continuous subjective sexual arousal were added to the model (Table 4). A significant increase in the Wald statistics from step one to step two indicates that the variables included in step two predict the criterion above and beyond variables in step one. The FSFI arousal domain was the only measure that was shown to outperform the event log change scores in the prediction of FSAD changes from pre- to post-test ($\chi^2 = 5.42$, $P < 0.05$).

Discussion

To the best of our knowledge, this is the first study to empirically evaluate whether event logs are a better measure of changes in clinical FSAD diagnosis compared with self-administered questionnaires, physiological measures of sexual arousal, and self-reported changes in subjective sexual arousal during laboratory exposure to erotic videos. The results from this study suggest that change scores in the FSFI arousal domain were able to significantly predict changes in clinical FSAD diagnosis at post-treatment. The frequency in satisfying sexual events did not predict changes in FSAD diagnosis above chance. Moreover, when compared with event logs, the FSFI arousal domain was the only measure shown to predict changes in diagnoses more accurately than the event logs.

There are several potential explanations for this finding. One viable interpretation of the findings

Table 4 Results for the hierarchical logistic regression testing whether the FSFI arousal domain or the VPA plus continuous subjective model predict treatment-induced changes in FSAD more accurately than event logs

	Correct FSAD predictions at post-treatment [†]		β	SE(B)	Wald χ^2
	FSAD, N (%)	No FSAD, N (%)			
<i>Step one</i>	11 (78.6)	3 (30.0)			1.45
Event logs					
Satisfaction			0.818	1.21	0.028
Mental arousal			0.833	1.30	0.020
Physical arousal			3.976	1.30	1.13
<i>Step two</i>	12 (85.7)	6 (60.0)			4.84*
Event logs					
Satisfaction			-0.83	0.44	1.87
Mental arousal			-0.63	0.53	0.18
Physical arousal			1.21	3.35	0.80
FSFI arousal			0.23	1.26	3.05**
<i>Step one</i>	11 (78.6)	3 (30.0)			1.45
Event logs					
Satisfaction			0.82	1.21	0.028
Mental arousal			0.83	1.30	0.020
Physical arousal			3.98	1.30	1.13
<i>Step two</i>	5 (41.7)	5 (80.0)			5.58
Event logs					
Satisfaction			-3.14	0.04	0.58
Mental arousal			0.59	1.81	0.05
Physical arousal			4.50	89.98	0.77
VPA			-0.04	0.97	0.01
Continuous subjective			-0.23	24.95	0.93

* $P < 0.05$; ** $P < 0.06$.[†]The frequencies and percentages reported in the table correspond to the number of participants that the model correctly detected. All frequencies and percentages are based on the entire FSAD (N = 14) and No FSAD (N = 10) sample.

is that event logs may not be able to capture the complex dimensions of women's sexual functioning. An increase in frequency of satisfying sexual events may not be associated with the overall sense of sexual satisfaction or sexual functioning. For example, increasing the number of sexually satisfying experiences may not compensate for the times when the sexual experience remains deeply unsatisfying and causes high levels of distress. Another potential explanation is that the time frame used for the event logs (2 weeks) may not be long enough to capture a woman's sense of sexual functioning. Indeed, the FDA guidelines suggest a 4-week baseline. This is a potential limitation for the sensitivity test of the event logs. However, the correlation between FSFI domains and event logs observed at pretreatment provides at least partial evidence that the event logs were able to accurately detect levels of sexual functioning. It is also feasible that using a different question in the event logs may have better captured the changes in sexual functioning among the participants. The questions in the event logs used in this study were selected because they directly addressed the guidelines published by the FDA (2000).

Conclusion

In conclusion, this study provides initial evidence that frequency of satisfying sexual encounters or frequency of sexual encounters when arousal is achieved may not be a sensitive index of treatment-induced changes in FSAD diagnoses. The findings from this study also provide initial evidence for the sensitivity of the FSFI to detect changes in sexual functioning for the arousal domain.

Acknowledgment

This publication was made possible by Grant Number 5 RO1 AT00224-02 from the National Center for Complementary and Alternative Medicine to Cindy Meston. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Center for Complementary and Alternative Medicine.

Corresponding Author: Alessandra Rellini, MA, University of Texas at Austin—Psychology, One University Station A8000, Austin, TX 78712, USA. Tel: (512) 232-4805; Fax: (512) 471-5935; E-mail: arellini@mail.utexas.edu

Conflict of Interest: None.

References

- 1 Meston CM. Female orgasmic disorder: Treatment strategies and outcome results. In: Goldstein I, Meston C, Davis S, Traish A, editors. *Women's sexual function and dysfunction: Study, diagnosis and treatment*. UK: Taylor & Francis (in press).
- 2 Rosen R, Barsky J. Psychological assessment and self-report questionnaires. In: Goldstein I, Meston C, Davis S, Traish A, editors. *Women's sexual function and dysfunction: Study, diagnosis and treatment*. UK: Taylor & Francis (in press).
- 3 FDA guidance document, guidance for industry: Female sexual dysfunction: Clinical development of drug products for treatment. 2000. Available at: <http://www.fda.gov/cder/guidance/3312dft.htm>.
- 4 Meston CM, Derogatis L. Validated Instruments for assessing female sexual function. *J Sex Marital Ther* 2002;28:155–64.
- 5 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.
- 6 Althof S, Rosen R, Derogatis L, Cortic E, Frances Q, Symonds T. Outcome measurement in female sexual dysfunction clinical trials. *J Sex Marital Ther* 2005;31:153–66.
- 7 Shifren J, Glenn BD, Simon JA, Casson PR, Buster JE, Redmond GP, Burki RE, Ginsburg ES, Rosen RC, Leiblum SR, Caramelli KE, Mazer NA. Transdermal testosterone treatment in women with impaired sexual function after oophorectomy. *N Engl J Med* 2000;343:682–8.
- 8 Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003;24:182–99.
- 9 Sintchak G, Geer JH. A vaginal plethysmograph system. *Psychophysiology* 1975;12:113–5.
- 10 Laan E, Everaerd W, Evers A. Assessment of female sexual arousal: Response specificity and construct validity. *Psychophysiology* 1995;32:476–85.
- 11 Meston CM, Gorzalka BB. The differentiated effects of sympathetic activation on sexual arousal in sexually functional and dysfunctional women. *J Abnorm Psychol* 1996;105:582–91.
- 12 Morokoff PJ, Heiman JR. Effects of erotic stimuli on sexually functional and dysfunctional women: Multiple measures before and after sex therapy. *Behav Res Ther* 1980;18:127–37.
- 13 Brotto LA, Basson R, Gorzalka BB. Psychophysiological assessment in premenopausal sexual arousal disorder. *J Sex Med* 2004;1:266–77.
- 14 Basson R, Leiblum S, Brotto L, DeRogatis L, Rourcroy J, Fugl-Meyer K, Graziottin A, Heiman JR, Laan E, Meston C, Schover L, van Lankveld J, Schultz WW. Definitions of women's sexual dysfunction reconsidered: Advocating expansion and revision. *J Psychosom Obstet Gynaecol* 2003;24:221–9.
- 15 Derogatis LR. The Derogatis interview for sexual functioning (DISF/DISF-R): An introductory report. *J Sex Marital Ther* 1997;23:291–304.
- 16 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th edition, text revised. Washington, DC: American Psychiatric Association; 2000.
- 17 Meston CM. 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* 2003;29:39–46.
- 18 Wiegel M, Meston CM, Rosen R. The female sexual function index (FSFI): Cross-validation and development of clinical cutoff scores. *J Sex Marital Ther* 2005;31:1–20.
- 19 Meston CM, Heiman JR. Ephedrine-activated physiological sexual arousal in women. *Arch Gen Psychi* 1998;55:652–6.
- 20 Rellini AH, McCall KM, Randall PK, Meston CM. The relationship between women's subjective and physiological sexual arousal. *Psychophysiology* 2005;42:116–24.

Copyright of *Journal of Sexual Medicine* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.