Satisfying Sexual Events as Outcome Measures in Clinical Trial of Female Sexual Dysfunction

Sheryl A. Kingsberg, PhD*† and Stanley E. Althof, PhD†‡

*OB/GYN Behavioral Medicine, University Hospitals Case Medical Center Cleveland, OH, USA; †Case Western Reserve University School of Medicine, Cleveland, OH, USA; ‡Center for Marital and Sexual Health of South Florida, West Palm Beach, FL, USA

DOI: 10.1111/j.1743-6109.2011.02447.x

ABSTRACT

Introduction. Assessing the sexual response in women with female sexual dysfunctions (FSDs) in clinical trials remains difficult. Part of the challenge is the development of meaningful and valid end points that capture the complexity of women's sexual response.

Aim. The purpose of this review is to highlight the shortcomings of daily diaries and the limitations of satisfying sexual events (SSEs) as primary end points in clinical trials of women with hypoactive sexual desire disorder (HSDD) as recommended by the Food and Drug Administration (FDA) in their draft guidance on standards for clinical trials in women with FSD.

Methods. Clinical trials in women with HSDD using SSEs as primary end points were reviewed.

Main Outcome Measures. The agreement between three outcome measures (SSEs, desire, and distress) was assessed to illustrate to what degree improvements in SSEs were in agreement with improvements in sexual desire and/or personal distress.

Results. Nine placebo-controlled randomized trials in women with HSDD were reviewed: seven with transdermal testosterone and two with flibanserin. In four trials, all using transdermal testosterone 300 μg/day had agreement between changes in SSEs, desire, and distress. In five studies (testosterone 300 μg/day, n = 2; testosterone 150 μg/day, n = 1; flibanserin n = 2), changes in SSEs did not correlate with changes in desire and/or distress and vice versa. It should be noted that in the flibanserin trials, SSEs did correlate with desire assessed using the Female Sexual Function Index but not when it was assessed using the eDiary.

Conclusions. Findings in the literature do not uniformly support the recommendations from the FDA draft guidance to use diary measures in clinical trials of HSDD as primary end points. Patient-reported outcomes appear to be better suited to capture the multidimensional and more subjective information collected in trials of FSD. **Kingsberg SA and Althof SE. Satisfying sexual events as outcome measures in clinical trial of female sexual dysfunction. J Sex Med **;**: **–**.**

Key Words. Hypoactive Sexual Desire Disorder; Testosterone; Flibanserin; Diary

Introduction

Female sexual dysfunctions (FSDs) are complex conditions with clinical assessment continue to be a challenge. The only guidance on standards for clinical trials in women with FSD was issued in 2000 by the US Food and Drug Administration (FDA) and is currently still only available in draft form [1]. The FDA draft guidance states, among other recommendations, that the number of satisfactory sexual events (SSEs), collected using daily diaries, should be the primary end point in clinical trials of FSD, while patient-reported outcomes (PROs) are recommended as secondary end points. These recommendations have received much criticism from experts in the FSD field [2,3]. The concerns focus mainly on five areas [3]: (i) SSEs are being recommended as primary end points, although they are not part of the criteria for a FSD, recognized by experts or the DSM-IV-TR

(Diagnostic and Statistical Manual, 4th ed, Text Revision). (ii) A DSM-IV–recognized FSD symptom should be selected as a primary end point, e.g., improvement in sexual desire, accompanied by a reduction in distress, in women with hypoactive sexual desire disorder (HSDD). (iii) The emphasis on daily recording of symptoms, e.g., sexual desire in women with HSDD, should be replaced with a longer recall period, as it has been shown that women with HSDD find a 1- to 4-week recall meaningful [4], and there is the potential for measurement contamination from daily assessments. (iv) Psychometric concerns, i.e., the concept of SSEs is several steps removed from the components being studied, such as desire or arousal.

Objective

The purpose of this review is to assess the limitations of daily diaries and of SSEs as primary end points in clinical studies of women with HSDD and to recommend the use of PROs as primary end points.

Daily Diaries vs. Self-Administered Questionnaires

Daily diaries or event logs have historically been used in clinical trials of conditions with well-defined end points, such as overactive bladder [5], irritable bowel syndrome [6], and sexual dysfunction [7,8]. Use of daily diaries is appropriate when events such as frequency of orgasms, incontinence episodes, or bowel movements are counted; they are much less suited for collecting subjective data. Simply counting is an unsophisticated form of assessing a complex and multidetermined construct such as desire.

It has been suggested that PROs, such as the Sexual Function Questionnaire [9] or the Female Sexual Function Index (FSFI), [10] are better suited to obtain multidimensional and more subjective information. This is supported by a trial in women with HSDD using transdermal testosterone that noted disagreement between sexual desire results obtained from diary measures and a validated PRO, the Brief Index of Sexual Function [11]. More recently, in two separate trials of flibanserin, the sexual desire score obtained from the diary disagreed with that from the FSFI in both trials; only when data from the studies were pooled did the two measures agree (Tables 1 and 2) [21]. Furthermore, a study assessing the sensitivity of different types of outcome measures (event logs, PROs, physiological measures of arousal, and self-reported changes in subjective sexual arousal in a laboratory setting) for detecting treatment-induced changes in women with female sexual arousal disorder found that the FSFI was the only instrument to demonstrate treatment response [22].

SSEs are determined by asking women to record their subjective experience and frequency of sexual activity by paper or electronic diary, such as the Sexual Activity Log (SAL), which records the number of intercourse and nonintercourse sexual events, number of orgasms, level of sexual desire, and satisfying sexual activity experienced (Table 3) [23]. However, questions about intensity and frequency of sexual desire collected daily may not be conceptually relevant to HSDD; neither the daily time frame nor the measurement of intensity is closely linked to the HSDD construct. Recent evidence showed that women with HSDD did not find a 24-hour recall, the most appropriate time frame for assessing their perceived desire, and instead preferred a period of 1–4 weeks [4].

Besides determining the number of events, the use of SSEs as an end point also requires the determination of a woman's perception of success or satisfaction, a highly subjective matter. Women might experience improved desire but choose not to engage in sexual activities or may not perceive the activity as satisfactory for reasons not related to their desire, such as still being upset over an argument with their partner. Alternatively, women might describe a sexual encounter as successful or satisfactory despite not experiencing improved desire. Moreover, a satisfying event does not necessarily motivate women to want to have another sexual encounter. In women with HSDD, the construct of low desire is only indirectly related to the number of sexual events because most sexual events in women with HSDD are initiated by the partner [3]. Many women agree to lovemaking out of a sense of obligation or love for their partner, not because they feel sexual desire.

Thus, the concept of an SSE appears to be several steps removed and not necessarily related to the sexual response component being examined in a trial, such as desire or arousal.

Male vs. FSD Trials

There are noteworthy discrepancies between the end points utilized in clinical trials of male sexual function, such as erectile dysfunction and premature ejaculation, and trials of FSD with

Table 1 Summary of clinical trials using SSEs as an outcome measure

Citation	Primary/secondary end points (measured by)	Summary of study	Highlights
Braunstein. Safety of testosterone treatment in postmenopausal women. *Arch Intern Med.* 2005;165:1582–9 [12].	Coprimary end points: Sexual desire (PFSF) Frequency of SSEs (SAL)	24-week, randomized, double-blind, placebo-controlled, parallel-group, multicenter trial in women (aged 24–70 years) who developed HSDD after bilateral salpingo-oophorectomy and hysterectomy receiving oral estrogen therapy randomized to placebo (n = 119) or testosterone patches 150 μg/day (n = 107), 300 μg/day (n = 110), or 450 μg/day (n = 111) twice weekly for 24 weeks. SSEs/week at baseline: ~0.75 (all groups) SSEs/week after 24 weeks: placebo ~1.0; testosterone: 150 μg ~1.0; 300 μg ~1.25 ($P < 0.05$), 450 μg ~1.1 hormone replacement therapy; Significantly greater increases from baseline in frequency of SSEs (79% vs. 43%; $P = 0.049$) Increase in PFSF sexual desire domain: placebo, 8.4; T300, 13.7 ($P = 0.05$); Significantly greater increases from baseline in sexual desire (67% vs. 48%; $P = 0.05$)	Both frequency of SSEs and sexual desire domain were significantly increased with T300 vs. placebo but not with T150 or T450. Changes with testosterone met or exceeded MID for respective end points (SSEs: >1/week; desire domain, ≥8.9); changes with placebo approached these values. Decreases in distress were not statistically different from placebo.
Buster et al. Testosterone patch for low sexual desire in surgically menopausal women: a randomized trial. *Obstet Gynecol.* 2005;105:944–52 [13].	Primary end point: Change in frequency of SSEs (SAL) Secondary end points: Sexual desire (PFSF) Personal distress (PDS) INTIMATE SM1	24-week, multicenter, double-blind, placebo-controlled trial in 533 women with HSDD with previous hysterectomy and bilateral oophorectomy randomly assigned placebo or testosterone patch twice weekly. Total SSEs significantly improved in testosterone group vs. placebo after 24 weeks (mean change from baseline, 1.56 vs. 0.73 episodes/4 weeks; $P = 0.001$). Testosterone also significantly improved sexual desire (mean change, 10.57 vs. 4.29; $P < 0.001$).	SSEs, desire, and distress all significantly improved with T300; changes exceeded MID.
Davis et al. Efficacy and safety of a testosterone patch for the treatment of hypoactive sexual desire disorder in surgically menopausal women: a randomized, placebo-controlled trial. *Menopause.* 2006;13:387–96 [14].	Coprimary end points: Sexual desire (PFSF) Frequency of SSEs (SAL)	24-week, randomized, double-blind, placebo-controlled trial in Europe and Australia. Women with HSDD after oophorectomy, receiving transdermal estrogen were randomly allocated to placebo (n = 40) or T300 (n = 37). The frequency of SSEs at baseline was 0.80 ± 0.1 (placebo) and 0.52 ± 0.08 (testosterone) and increased at week 24 by 0.28 ± 0.15 and 0.77 ± 0.15, respectively, but was not statistically significant ($P = 0.06$). Testosterone treatment resulted in significantly greater change from baseline in sexual desire domain (16.43 vs. 5.98; $P = 0.02$) vs. placebo.	SSEs did not improve significantly with testosterone treatment. Sexual desire and personal distress improved significantly.
Davis et al. Testosterone for low libido in postmenopausal women not taking estrogen. *N Engl J Med.* 2008;359:2005–17 [15].	Primary end point: Change in frequency of SSEs (SAL) Secondary end points: Sexual desire (PFSF) Personal distress (PDS) APHRODITE	52-week double-blind, placebo-controlled trial in 814 women with HSDD randomly assigned to testosterone patch 150 or 300 μg/day or placebo. At 24 weeks, increase in 4-week frequency of SSEs significantly greater in T300 vs. placebo groups (increase of 2.1 episodes vs. 0.7; $P < 0.001$) but not in T150-μg group (1.2 episodes; $P = 0.11$). Both doses of testosterone were associated with significant increases in desire (300 μg, $P < 0.001$; 150 μg, $P = 0.04$).	SSEs, desire, and distress are significantly increased with T300, but with T150, only desire and distress are improved.
Clayton et al. Flibanserin: a potential treatment for Hypoactive Sexual Desire Disorder in premenopausal women. *Women's Health* 2010;6:639–53 [16].	Coprimary end points: Change in number of SSEs Change in sexual desire score (eDiary) VIOLET, DAISY	Pooled data from two randomized, double-blind studies (VIOLET, DAISY) in premenopausal women with HSDD randomized to placebo (n = 693) or flibanserin 100 mg qhs (n = 685) for 24 weeks. The mean change from baseline to week 24 in SSE/month was 1.7 for flibanserin vs. 1.0 for placebo ($P < 0.0001$). For eDiary desire score, the change from baseline to week 24 was 9.0 for flibanserin vs. 7.1 for placebo ($P < 0.05$).	In VIOLET and DAISY, sexual desire scores per eDiary were not significantly improved; SSEs significantly improved.
Shifren et al. Testosterone patch for the treatment of hypoactive sexual desire disorder in naturally menopausal women: results from the INTIMATE NM1 Study. *Menopause.* 2006;13:770–9 [17].	Primary end point: Change in frequency of SSEs (SAL) Secondary end points: Sexual desire (PFSF) Personal distress (PDS) INTIMATE NM1	Multicenter, double-blind, placebo-controlled study in naturally menopausal women with HSDD receiving estrogen with or without progestin (n = 549) randomized to placebo or testosterone patch (300 μg) for 24 weeks. Women with SHBG ≤ 160 nmol/L were included in primary analysis. SSEs were significantly increased with testosterone vs. placebo (mean change from baseline, 1.9 vs. 0.5; $P < 0.0001$). Testosterone significantly decreased distress and increased sexual desire.	SSEs, desire, and distress all significantly improved with T300.
Simon et al. Testosterone patch increases sexual activity and desire in surgically menopausal women with hypoactive sexual desire disorder. *J Clin Endocrinol Metab.* 2005;90:5226–33 [18].	Primary end point: Change in frequency of SSEs (SAL) Secondary end points: Sexual desire (PFSF) Personal distress (PDS) INTIMATE SM2	Women (aged 26–70 years) with HSDD after bilateral salpingo-oophorectomy receiving concomitant estrogen therapy received placebo (n = 279) or T300 (n = 283). At 24 weeks, there was an increase from baseline in frequency of total SSEs of 2.10 episodes/4 weeks in testosterone group, which was significantly greater than change of 0.98 episodes/4 weeks in placebo group ($P = 0.0003$).	SSEs, desire, and distress all significantly improved with T300.
Panay et al. Testosterone treatment of HSDD in naturally menopausal women: the ADORE study. *Climacteric* 2010; 13:121–32 [19].	Primary end point: Change in frequency of SSEs (SAL) Secondary end points: Sexual desire (PFSF) Personal distress (PDS) ADORE	Two hundred seventy-two naturally menopausal women with HSDD predominantly not using HRT were randomized to a transdermal testosterone patch or placebo. After 6 months, the testosterone group demonstrated significant improvements in SSEs ($P = 0.0089$), sexual desire ($P = 0.0007$), and reduced personal distress ($P = 0.0024$) vs. placebo.	SSEs, desire, and distress all significantly improved with T300.

PFSF = Profile of Sexual Function; SSEs = satisfactory sexual events; SAL = Sexual Activity Log; HSDD = hypoactive sexual desire disorder; T300/150/450 = testosterone 300/150/450 μg/day; PDS = Personal Distress Scale; INTIMATE = Investigation of Natural Testosterone In Menopausal Women Also Taking Estrogen; SM = surgically menopausal; MID = minimally important difference; APHRODITE = A Phase III Research Study of Female Sexual Dysfunction in Women on Testosterone Patch Without Estrogen; VIOLET = Evaluation of the Impact on Sexuality with Evening Treatment; DAISY = Dose Ascending Study Over Half a Year; NM = naturally menopausal; qhs = every night at bedtime; SHBG = sex hormone binding globulin; ADORE = A Study in Women with Low Sexual Desire to Evaluate the Efficacy and Safety of Transdermal Testosterone Therapy in Naturally Menopausal Women Receiving Transdermal Estrogen Therapy; HRT = hormone replacement therapy.

Table 2 Summary of changes in outcomes measures in clinical trials using SSEs, desire, and distress

	Mean change from baseline			
	SSEs MID ≥ 1	Desire MID		Distress
Braunstein et al. [20]				
T300	0.58*	13.7*		NS†
PBO	—†	8.4		—†
INTIMATE SM1 [13]				
T300	1.56***	10.5***		−16.05***
PBO	0.73	4.98		−22.7
INTIMATE SM2 [18]				
T300	2.1***	11.06***		−15.07**
PBO	0.98	5.94		−22.77
INTIMATE NM1 [17]				
T300	1.92****	9.79****		−20.5****
PBO	0.54	4.00		−11.5
APHRODITE [15]				
T300	2.1***	~twofold† ↑***		~twofold† ↓***
PBO	0.7	—†		—†
Davis 2006 [14]				
T300	0.77	5.98*		−22.8**
PBO	0.28	16.4		−3.49
Panay 2010 [19]				
T300	1.69***	25.98***		−20.95**
PBO	0.53	16.4		−10.4
Flibanserin [16,21]		eDiary	FSFI desire	
VIOLET				
Flibanserin	1.0**	9.1	0.9****	−8.9***
PBO	0	6.9	0.5	−4.9
DAISY				
Flibanserin	1.9*	8.5	0.9****	−7.8***
PBO	1.1	6.8	0.6	−6.7
Pooled				
Flibanserin	1.7****	9.0**	0.9***	−8.0****
PBO	1.0	7.1	0.5	−4.8

*$P < 0.05$, **$P < 0.01$, ***$P \leq 0.001$, ****$P < 0.0001$, all vs. PBO.
†Numeric value not provided in original source.
Minimal important differences have been established for SSEs (increase in frequency >1), sexual desire (increase in score ≥ 8.9), and personal distress (decrease in score ≥ 20) [13,18].
SSE = satisfactory sexual events; T300 = testosterone 300 μg; PBO = placebo; NS = nonsignificant change; INTIMATE = Investigation of Natural Testosterone In Menopausal Women Also Taking Estrogen; SM = surgically menopausal; APHRODITE = A Phase III Research Study of Female Sexual Dysfunction in Women on Testosterone Patch Without Estrogen; FSFI = Female Sexual Function Index; VIOLET = Evaluation of the Impact on Sexuality with Evening Treatment; DAISY = Dose Ascending Study Over Half a Year.

respect to primary end points. Many clinical trials of phosphodiesterase type 5 inhibitors used questions 3 and 4 of the International Index of Erectile Function (IIEF) or the erectile function domain (EF) of the IIEF [7,8,24] and sexual encounter profile questions 2 and 3 (SEP2, SEP3) as primary end points [7,8]. Clinical studies of premature ejaculation assess the intravaginal ejaculation latency time (IELT), an objective time measurement, as well as the Index of Premature Ejaculation, a validated self-administered questionnaire, and/or the premature ejaculation profile, a validated self-reported outcome measure [25]. Thus, clinical studies of male sexual dysfunction tend to use both diary and validated PROs as primary end points, although the aspect of sexual function assessed, such as EF and IELT, are discrete and readily measured end points.

Both male and female hypoactive sexual desire are a challenging concept to measure. It would seem that subjective aspects of hypoactive sexual desire are the most important areas to assess. As previously discussed, PROs are better suited to obtain multidimensional and more subjective information than diary measures. There has been no FDA guidance necessitating the use of SSE to measure male HSDD as there has been for female HSDD; it is unclear why the guidance is different for women than for men with HSDD.

Table 3 Questions for Sexual Activity Log, FSFI desire domain, and FSDS Q13

Question	Responses	Rating
eDiary—Sexual Activity Log (daily recall)		
Q1. Indicate your most intense level of sexual desire in the last 24 hours?	No desire	0
	Low desire	1
	Moderate desire	2
	Strong desire	3
Q2. How distressed have you felt about your level of sexual desire in the last 24 hours?	Not at all	0
	A little bit	1
	Moderately	2
	Quite a bit	3
	Extremely	4
Q3. Did you have sex in the last 24 hours?	No	0
	Yes	1
Q4. How many times did you have sex in the last 24 hours?		
Q5. Was the sex satisfying for you?	No	0
	Yes	1
Q6. Did you have an orgasm?	No	0
	Yes	1
FSFI desire domain (4-week recall)		
Q1. Over the past 4 weeks, how often did you feel sexual desire or interest?	Almost always or always	5
	Most times	4
	Sometimes	3
	A few times	2
	Almost never or never	1
Q2. Over the past 4 weeks, how would you rate your level (degree) of sexual desire or interest?	Very high	5
	High	4
	Moderate	3
	Low	2
	Very low or none at all	1
FSDS (4-week recall)		
Q13. How often did you feel bothered by low sexual desire?	Always	4
	Frequently	3
	Occasionally	2
	Rarely	1
	Never	0

FSFI = Female Sexual Function Index; FSDS = Female Sexual Distress Scale.

Placebo Effect

The use of daily diaries in general is associated with a placebo effect irrespective of the condition that is treated. For example, placebo effects have been demonstrated in trials of overactive bladder [5], depression [26], and irritable bowel syndrome [6] but are also common in trials of neurologic/psychiatric conditions. Parkinson's disease, epilepsy, pain, and depression all share a strong "top-down" or cortically-based regulation mechanism that is thought to involve a dopaminergic reward system.

High placebo responses can make it difficult to demonstrate significant treatment effects and can severely limit the assay sensitivity of a clinical trial. For example, in the flibanserin trials, placebo responses for SSEs and sexual desire, both daily eDiary measures, were 37% and 62%, respectively. In comparison, placebo responses for the FSFI desire score and the Female Sexual Distress Score (FSDS) Item 13 were 26% and 16%, respectively (13% for total FSFI and 17% for total FSDS). Compared with measures from the FSFI and FSDS, the two diary outcomes were the least robust in these trials [3,21].

Additionally, placebo effects can be explained by other, more psychological factors. For example, women enrolled in FSD trials have a desire to improve their sex lives and take an active role in seeking help; additionally, expectancies for enhanced sexual desire would increase a woman's perception of having desire. Frequent diary entries and increased communication with their partners about their sexual experiences keep their sexual activity at the forefront of their minds. Thus, it is not surprising that placebo effects have also been demonstrated in trials of FSD [28], including those with flibanserin [21], testosterone [13,17,18,20], and sildenafil [27], which have all shown substantial placebo effects for primary/secondary end points. A review of placebo responses in the treatment of FSD in 16 clinical studies found that postmenopausal women and women with HSDD may be more likely to respond to placebo treatment [28].

Clinical Trials Using SSEs as End points

The first clinical FSD trials using SSEs as primary end point were published in 2005 and most used testosterone for the treatment of HSDD in postmenopausal women.

Table 1 gives an overview of all FSD trials, their design, end points, and results, and Table 2 summarizes changes in three treatment-related outcomes, SSEs, desire, and distress.

Overall, there were seven clinical trials assessing transdermal testosterone patches in naturally or surgically menopausal women [13–15,17–20], including three Investigation of Natural Testosterone In Menopausal Women Also Taking Estrogen (INTIMATE) trials [13,17,18], the A Phase III Research Study of Female Sexual Dysfunction in Women on Testosterone Patch Without Estrogen (APHRODITE) trial [15], and the A Study in Women with Low Sexual Desire to Evaluate the Efficacy and Safety of Transdermal Testosterone Therapy in Naturally Menopausal Women Receiving Transdermal Estrogen Therapy (ADORE) study [19]. In all seven trials, the primary end point was the change in frequency of SSEs measured using a SAL with a 7-day recall period, in other words, not a daily diary. Secondary end points included sexual desire and personal distress measured by the Profile of Sexual Function (PFSF) and Personal Distress Scale (PDS); both were self-administered questionnaires. Most testosterone trials lasted for 24 weeks, with the exception of the APHRODITE study (52 weeks). All but one trial [14] using 300 μg testosterone resulted in SSEs being significantly improved compared with placebo at the end of the study. Additionally, 300 μg testosterone significantly improved sexual desire in all trials and decreased distress in all but one study [20]. Another trial using SSEs as end points assessed the use of testosterone as transdermal spray in premenopausal women; sexual desire and personal distress were not measured [29]. After 16 weeks, the mean number of SSEs was significantly greater and achieved the minimal important difference (MID) with the intermediate dose of testosterone (90 μL) compared with placebo.

Two recent trials using SSEs as coprimary end points involved the assessment of flibanserin in women with generalized HSDD [16]. Pooled data from both trials revealed that the number of SSEs was increased significantly over placebo after 24 weeks. Using pooled data, changes in sexual desire (measured using an eDiary as well as the FSFI desire score) and personal distress (measured using item 13 of the FSDS; see Table 3) were significant in the flibanserin group vs. placebo. However, changes in the eDiary sexual desire score were not significant in the individual trials.

Clinical Relevance

Ultimately, the critical questions for any treatment are whether they are considered beneficial and safe for women. Studies looking at the clinical benefit answer one of these two questions. The INTIMATE surgically menopausal trials included two clinical relevance studies [30,31]. One study by DeRogatis et al. [30] established the MIDs in a subset of 132 surgically postmenopausal women from two INTIMATE studies [13,18]. MIDs for SSEs (increase in frequency >1/month) were achieved in five of the seven above studies [13,15,17–19], MIDs for sexual desire scores (increase of ≥8.9) were achieved in all seven studies, and MIDs for personal distress (decrease of ≥20.0) were achieved in six studies, with one trial not showing the change in distress [20].

Another analysis of the same subset of 132 women from INTIMATE studies by Kingsberg et al. [31] also assessed the clinical relevance of benefits associated with testosterone treatment, and reported that 52% of women receiving testosterone reported a meaningful treatment benefit, and that in those women the mean change from baseline in 4-week frequency of SSEs was 4.4, in desire score was 21.0 (moving from "seldom" to "sometimes" feeling desire), and in distress score was –26.5 (moving from "often" to "seldom" being distressed). The same scores were notably lower in women who did not report a meaningful treatment benefit (mean change from baseline in 4-week frequency of SSEs, 0.5; desire score was 2.9; distress score was –8.8). Analyses from both of these studies demonstrate that surgically menopausal women with HSDD received clinically significant and meaningful treatment benefits from testosterone patch treatment. Furthermore, they support the three end points being used (SSEs, PFSF, PDS) as corresponding to clinical relevant improvement in sexual function and demonstrate that these end points are able to discriminate between women with and without perceived benefit.

In the flibanserin trials, patient perception of treatment benefit was assessed using the Patient's Global Impression of Improvement (PGI-I; "How is your condition—meaning decreased sexual desire and feeling bothered by it—today compared with when you started study medication?"), a seven-point scale ranging from 1 (very much improved) to 7 (very much worse), and the patient benefit evaluation (PBE; "Overall, do you believe that you have experienced a meaningful benefit from the study medication?" Yes/No) [16]. The number of women reporting that their overall condition was "very much improved," "much improved," or "minimally improved" on the PGI-I at week 24 was 48.3% for flibanserin vs. 30.3% for placebo ($P < 0.0001$). Similarly, more women in the flibanserin (40.5%) than in the placebo group (25.2%) reported on the PBE that they had received a meaningful benefit from the study medication ($P < 0.0001$). Thus, premenopausal women with generalized acquired HSDD who received flibanserin for 24 weeks were significantly more likely to report a meaningful improvement in their condition than women who received placebo.

Correlations Between e-Diary SSE and Other End-Point Measures

Pearson correlations between standardized SSEs and psychometrically validated PROs with desire PROs were determined using data from two flibanserin trials (Table 4). The data shows moderate correlations between SSEs and the e-Diary desire score (0.49) and the FSFI desire domain (0.47). Correlations between the FSFI desire domain and the Female Sexual Distress Scale-Revised and the PGI-I were –0.56 and –0.67, respectively. In comparison, correlations between SSEs and the Female Sexual Distress Scale–Revised and the PGI-I were –0.35 and –0.44, respectively. This correlational data supports our contention of the somewhat weaker relationship between SSEs and data derived from validated sexual desire questionnaires. Quantitatively speaking, SSEs appear to measure something only slightly related to sexual desire as assessed by a validated questionnaire. Similarly, SSEs are only moderately related to the daily diary desire score also suggesting it is likely to be assessing something different than sexual desire. While SSEs may demonstrate improvement in some studies, it appears that they account for only a small portion of the variance in PROs measures of desire and distress. Perhaps, this is statistically analogous to the saying that a high tide floats all boats, and the improvements seen in SSEs may be because of improvements in sexual desire and diminished distress.

Summary and Conclusions

To summarize, the concerns regarding SSEs, which are derived from daily diaries in clinical trials of FSD, include the following: (i) their use as primary

Table 4 Pearson correlations of standardized SSE with other patient-reported outcomes (PROs) in flibanserin trials

		Pearson correlation with SSEs		Pearson correlation with eDiary desire		Pearson correlation with FSFI desire domain	
		Change from baseline to week 24					
Trial	Measure	r	r^2	r	r^2	r	r^2
VIOLET	SSE	1.00	1.00	0.49	0.23	0.48	0.23
	eDiary desire	0.49	0.24	1.00	1.00	0.59	0.35
	FSDS-R total	–0.38	0.14	–0.46	0.21	–0.57	0.32
	FSFI total	0.48	0.23	0.59	0.17	1.00	1.00
	PGI Improvement	–0.46	0.21	–0.54	0.29	–0.67	0.45
DAISY	SSE	1.00	1.00	0.49	0.24	0.47	0.22
	eDiary desire	0.49	0.24	1.00	1.00	0.63	0.40
	FSDS-R total	–0.34	0.212	–0.45	0.20	–0.55	0.30
	FSFI total	0.47	0.22	0.63	0.40	0.61	0.37
	PGI Improvement	–0.43	0.19	–0.55	0.30	–0.67	0.45
Pooled	SSE	1.00	1.00	0.49	0.24	0.47	0.23
	eDiary desire	0.49	0.24	1.00	1.00	0.62	0.38
	FSDS-R total	–0.35	0.13	–0.45	0.20	–0.56	0.31
	FSFI total	0.47	0.23	0.62	0.38	1.00	1.00
	PGI Improvement	–0.44	0.19	–0.54	0.30	–0.67	0.45

PROs = patient-reported outcomes; SSE = satisfactory sexual event; FSFI = Female Sexual Function Index; r = correlation coefficient; r^2 = squared correlation coefficient ; VIOLET = Evaluation of the Impact on Sexuality with Evening Treatment; FSDS-R = Female Sexual Distress Scale-Revised; PGI = Patient's Global Impression; DAISY = Dose Ascending Study Over Half a Year.

end points although they are not part of the DSM-IV-TR diagnostic criteria for FSD; (ii) the call for a DSM-IV–recognized FSD symptom as primary end point, e.g., improvement in sexual desire accompanied by a reduction in distress; (iii) the emphasis on daily recording of symptoms and potential measurement contamination from daily assessments; and (iv) the concept of SSEs being too far downstream from desire or arousal. Women's sexuality, particularly desire, is highly complex and may best be evaluated by measuring multiple end points addressing the multifactorial nature of desire. To that effect, SSEs, which are considered single-item scales, are too simplistic and fail to account for the subtle multidimensional subjective issues that contribute to the construct of sexual desire. Additionally, daily diaries have not been validated and are plagued with compliance issues [4,11].

In all INTIMATE studies, the primary/secondary end points (frequency of SSEs, sexual desire, personal distress) were significantly improved with testosterone vs. placebo, and the changes observed exceeded the minimally important difference (SSEs, >1/week; sexual desire, ≥8.9; personal distress, ≥20) [30]. In other trials, increases in the frequency of SSEs were not always accompanied by increases in desire or vice versa (Table 2) [14,15,21,32] In the flibanserin studies, for example, there was disagreement between changes in SSE frequency and the desire score from the eDiary in the two individual trials; only after data were pooled were these two outcomes in agreement. In contrast, when desire was assessed using the FSFI, changes in the frequency of SSEs were accompanied by a change in the same direction of the desire score. Distress resulting from decreased sexual desire was improved in all but one trial after treatment [20]. It is noteworthy that SSEs and sexual desire correlated in all trials that assess desire using a validated questionnaire (PFSF in testosterone studies; FSFI in flibanserin studies) but not when desire was assessed using a diary (flibanserin studies). These findings do not uniformly support the recommendations from the FDA draft guidance to use diary measures in clinical trials as primary end points and to designate PROs as secondary end points. Counting events is too simplistic to capture the complexities of female sexual disorders. Further, the variability between the different measures—SSEs, sexual desire, and personal distress—makes it difficult to interpret clinical results. PROs are more suitable as primary end points because of their ability to reliably capture these changes and their likelihood to reflect the outcome for multidimensional assessments; measuring the clinical relevance will help validate the end points chosen. Furthermore, outcome measures either not appropriate for a particular end point or lacking sensitivity may produce false negative results, increasing the possibility of discounting an effective drug.

Corresponding Author: Stanley E. Althof, PhD, Center for Marital and Sexual Health of South Florida, West Palm Beach, FL 33401, USA. Tel: 561-822-5454; Fax: 561-822-5458; E-mail: stanley.althof@case.edu

Conflict of Interest: S. Althof: Boehringer-Ingelheim—advisory board, principal investigator, speaker's bureau; Endoceutics—principal investigator; Johnson & Johnson—principal investigator, consultant; Lilly—advisory board, consultant; Neurohealing—advisory board; Palitan—consultant; Shinonogi—advisory board, consultant. Neither I nor any family members have a financial or material interest. S. Kingsberg: consulting—Boehringer-Ingelheim, Pfizer, Trimel Biopharm; advisory board—Viveve; clinical investigator—BioSante, Boehringer-Ingelheim.

Statement of Authorship

Category 1

(a) Conception and Design
Sheryl A. Kingsberg; Stanley E. Althof

(b) Acquisition of Data
Sheryl A. Kingsberg; Stanley E. Althof

(c) Analysis and Interpretation of Data
Sheryl A. Kingsberg; Stanley E. Althof

Category 2

(a) Drafting the Article
Sheryl A. Kingsberg; Stanley E. Althof

(b) Revising It for Intellectual Content
Sheryl A. Kingsberg; Stanley E. Althof

Category 3

(a) Final Approval of the Completed Article
Sheryl A. Kingsberg; Stanley E. Althof

References

1 US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Draft guidance for industry: Female sexual dysfunction: Clinical development of drug products for treatment. 2000; Available at: http://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm133202.htm (accessed August 1, 2011).

2 Althof SE, Rosen RC, DeRogatis L, Corty E, Quirk F, Symonds T. Outcome measurement in female sexual dysfunction clinical trials: Review and recommendations. J Sex Marital Ther 2005;31:153–66.

3 Clayton AH, Dennerstein L, Fisher WA, Kingsberg SA, Perelman MA, Pyke RE. Standards for clinical trials in sexual dysfunction in women: Research designs and outcomes assessment. J Sex Med 2010;7:541–60.

4 Revicki D, Margolis M, Bush E, Hanes V. Content validity of the Female Sexual Function Index in pre- and postmenopausal women with hypoactive sexual desire disorder. Presented at the 3rd International Consultation on Sexual Medicine. Paris, France, July 10–13, 2009.

5 Nitti VW, Dmochowski R, Sand PK, Forst HT, Haag-Molkenteller C, Massow U, Wang J, Brodsky M, Bavendam T. Efficacy, safety and tolerability of fesoterodine for overactive bladder syndrome. J Urol 2007;178:2488–94.

6 Kellow J, Lee OY, Chang FY, Thongsawat S, Mazlam MZ, Yuen H, Gwee KA, Bak YT, Jones J, Wagner A. An Asia-Pacific, double blind, placebo controlled, randomised study to evaluate the efficacy, safety, and tolerability of tegaserod in patients with irritable bowel syndrome. Gut 2003;52:671–6.

7 Keating GM, Scott LJ. Spotlight on vardenafil in erectile dysfunction. Drugs Aging 2004;21:135–40.

8 Seftel AD, Buvat J, Althof SE, McMurray JG, Zeigler HL, Burns PR, Wong DG. Improvements in confidence, sexual relationship and satisfaction measures: Results of a randomized trial of tadalafil 5 mg taken once daily. Int J Impot Res 2009;21:240–8.

9 Quirk FH, Heiman JR, Rosen RC, Laan E, Smith MD, Boolell M. Development of a sexual function questionnaire for clinical trials of female sexual dysfunction. J Womens Health Gend Based Med 2002;11:277–89.

10 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.

11 Shifren JL, Braunstein GD, 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.

12 Braunstein GD. Safety of testosterone treatment in postmenopausal women. Fertil Steril 2007;88:1–17.

13 Buster JE, Kingsberg SA, Aguirre O, Brown C, Breaux JG, Buch A, Rodenberg CA, Wekselman K, Casson P. Testosterone patch for low sexual desire in surgically menopausal women: A randomized trial. Obstet Gynecol 2005;105:944–52.

14 Davis SR, van der Mooren MJ, van Lunsen RH, Lopes P, Ribot C, Rees M, Moufarege A, Rodenberg C, Buch A, Purdie DW. Efficacy and safety of a testosterone patch for the treatment of hypoactive sexual desire disorder in surgically menopausal women: A randomized, placebo-controlled trial. Menopause 2006;13:387–96.

15 Davis SR, Moreau M, Kroll R, Bouchard C, Panay N, Gass M, Braunstein GD, Hirschberg A, Rodenberg C, Pack S, Koch H, Moufarege A, Studd J; the APHRODITE Study Team. Testosterone for low libido in postmenopausal women not taking estrogen. N Engl J Med 2008;359:2005–17.

16 Clayton AH, Pyke R, Sand M. Flibanserin: A potential treatment for Hypoactive Sexual Desire Disorder in premenopausal women. Women's Health 2010;6:639–53.

17 Shifren JL, Davis SR, Moreau M, Waldbaum A, Bouchard C, DeRogatis L, Derzko C, Bearnson P, Kakos N, O'Neill S, Levine S, Wekselman K, Bouch A, Rodenberg C, Kroll R. Testosterone patch for the treatment of hypoactive sexual desire disorder in naturally menopausal women: Results from the INTIMATE NM1 Study. Menopause 2006;13:770–9.

18 Simon J, Braunstein G, Nachtigall L, Utian W, Katz M, Miller S, Waldbaum A, Bouchard C, Derzko C, Buch A, Rodenberg C, Lucas J, Davis S. Testosterone patch increases sexual activity and desire in surgically menopausal women with hypoactive sexual desire disorder. J Clin Endocrinol Metab 2005;90:5226–33.

19 Panay N, Al-Azzawi F, Bouchard C, et al. Testosterone treatment of HSDD in naturally menopausal women: The ADORE study. Climacteric 2010;13:121–31.

20 Braunstein GD, Sundwall DA, Katz M, Shifren JL, Buster JE, Simon JA, Bachman G, Aguirre OA, Lucas JD, Rodenberg C, Buch A, Watts NB. Safety and efficacy of a testosterone patch for the treatment of hypoactive sexual desire disorder in sur-

gically menopausal women: A randomized, placebo-controlled trial. Arch Intern Med 2005;165:1582–9.
21 Jolly E, Clayton A, Thorp J, Nappi R, Kimura T, Sand M, Pyke R. Efficacy of flibanserin 100 mg qhs as a potential treatment for Hypoactive Sexual Desire. Presented at the 12th Congress of the European Society for Sexual Medicine. Lyon, France, November 15–18, 2009.
22 Rellini A, 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;3:283–91.
23 DeRogatis L, Rust J, Golombok S, Kuznicki J, Rodenberg C, McHorney C. A patient based diary to measure sexual activity in menopausal women with HSDD. Presented at the Annual Meeting of the International Society for the Study of Women's Sexual Health. Atlanta, GA, October 28–31, 2004.
24 Cappelleri JC, Stecher VJ. An assessment of patient-reported outcomes for men with erectile dysfunction: Pfizer's perspective. Int J Impot Res 2008;20:343–57.
25 Althof SE, Symonds T. Patient reported outcomes used in the assessment of premature ejaculation. Urol Clin North Am 2007;34:581–9. vii.
26 Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KR, Shiovitz T, Weihs K. An 8-week multicenter, parallel-group, double-blind, placebo-controlled study of sertraline in elderly outpatients with major depression. Am J Psychiatry 2003;160:1277–85.
27 Basson R, McInnes R, Smith MD, Hodgson G, Koppiker N. Efficacy and safety of sildenafil citrate in women with sexual dysfunction associated with female sexual arousal disorder. J Womens Health Gend Based Med 2002;11:367–77.
28 Bradford A, Meston CM. Placebo response in the treatment of women's sexual dysfunctions: A review and commentary. J Sex Marital Ther 2009;35:164–81.
29 Davis S, Papalia MA, Norman RJ, O'Neill S, Redelman M, Williamson M, Stuckey BG, Wlodarczyk J, Gard'ner K, Humberstone A. Safety and efficacy of a testosterone metered-dose transdermal spray for treating decreased sexual satisfaction in premenopausal women: A randomized trial. Ann Intern Med 2008;148:569–77.
30 DeRogatis LR, Graziottin A, Bitzer J, Schmitt S, Koochaki PE, Rodenberg C. Clinically relevant changes in sexual desire, satisfying sexual activity and personal distress as measured by the profile of female sexual function, sexual activity log, and personal distress scale in postmenopausal women with hypoactive sexual desire disorder. J Sex Med 2009;6:175–83.
31 Kingsberg S, Shifren J, Wekselman K, Rodenberg C, Koochaki P, Derogatis L. Evaluation of the clinical relevance of benefits associated with transdermal testosterone treatment in postmenopausal women with hypoactive sexual desire disorder. J Sex Med 2007;4:1001–8.
32 Nijland EA, Weijmar Schultz WC, Nathorst-Boos J, Helmond FA, van Lunsen RH, Palacios S, Norman RJ, Mulder RJ, Davis SR. Tibolone and transdermal E2/NETA for the treatment of female sexual dysfunction in naturally menopausal women: Results of a randomized active-controlled trial. J Sex Med 2008;5:646–56.