

ORIGINAL RESEARCH—OUTCOMES ASSESSMENT**Methods to Determine the Minimum Important Difference for a Sexual Event Diary Used by Postmenopausal Women with Hypoactive Sexual Desire Disorder**

Tara Symonds, PhD,* Cathie Spino, ScD,[†] Melanie Sisson, AuD,[‡] Paresh Soni, MD, PhD,[‡] Mona Martin, RN, MPH,[§] Lacey Gunter, MS,[¶] and Donald L. Patrick, PhD, MSPH**

*Pfizer Ltd., Kent, UK; [†]Pfizer Inc., Ann Arbor, MI, USA; [‡]Pfizer Inc, New London, CT, USA; [§]Health Research Associates, Seattle, WA, USA; [¶]University of Michigan, Ann Arbor, MI, USA; **University of Washington, Seattle, WA, USA

DOI: 10.1111/j.1743-6109.2007.00562.x

ABSTRACT

Introduction. Recently, there has been much discussion in the literature about how to determine the meaningfulness of results generated from a patient-reported outcome measure. A number of reviews have shown that there are two main approaches: anchor- and distribution-based approaches for determining the minimum important difference (MID) for a new measure. There are issues with calculating an MID using each method: Will the two approaches give the same estimate? If the estimates differ, how do you decide on one estimate? Would asking patients directly be more beneficial?

Aim. A case study was presented to address these issues based on a newly developed diary assessing number of satisfactory sexual events (SSEs) per week in women with hypoactive sexual desire disorder (HSDD).

Methods. Anchor- and distribution-based estimates were generated from data gathered in two double-blind, placebo-controlled, parallel group trials for the treatment of HSDD (N = 788). A novel interview study was used to ask women directly about an MID for SSEs (N = 77).

Main Outcome Measures. Defining the MID for an SSE diary in women with HSDD.

Results. The estimates varied, producing a range of mean MID estimates between 0.04 and 0.46 SSEs per week.

Conclusion. We recommend that rather than defining the MID, a range should be selected from the set of estimates formed by the limits of the 95% confidence intervals. **Symonds T, Spino C, Sisson M, Soni P, Martin M, Gunter L, and Patrick DL. Methods to determine the minimum important difference for a sexual event diary used by postmenopausal women with hypoactive sexual desire disorder. J Sex Med 2007;4:1328–1335.**

Key Words. Hypoactive Desire Disorder; Female Sexual Dysfunction; Minimum Important Difference

Introduction

Patient-reported outcomes (PROs) have become widely accepted as important assessments of patients' self-reported change in health status or change in quality of life after some therapeutic intervention. However, changes on these PROs are often difficult to interpret: Is an average change of five points on a 0–100 scale in a group of patients meaningful? The problem is often com-

plicated by the fact that the end point has not been widely used and, therefore, makes understanding and interpreting what a meaningful change is difficult. This has led researchers to consider ways of interpreting such changes in the absence of wider experience. Recently, a number of reviews [1–3] have identified various approaches that could be useful to interpret the minimum important difference (MID). The methods outlined in these reviews are grouped together under two general

methodologies: anchor- and distribution-based approaches.

Anchor-based methods relate magnitudes of change in an end point with an independent measure of change (anchor) that provides meaning to the degree of change. A suitable anchor is one that is interpretable and has an appreciable correlation with the target measure. Traditionally, the anchor-based approach has been used to determine the MID, from a patient perspective as defined by Juniper et al. [4]: “the smallest difference in score . . . which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management.” Asking patients to determine if there has been any change in their condition (e.g., dyspnea, fatigue, incontinence) allows the magnitude of this change to then be anchored to change in the end point under investigation; those individuals who indicate that they have had a slight improvement are grouped together to determine average change on the end point. This MID would be considered patient derived because the change has been anchored to a patient self-report of change. The anchor is traditionally asked at the end of the treatment period, which allows only a cross-sectional assessment of the change in the outcome with the patients’ assessment of their condition (subject to recall bias) at this point in time. However, a longitudinal approach is thought to be better than a cross-sectional approach [5] whereby the change in the anchor is determined prospectively with assessments at baseline (pretreatment) and at the end of treatment (EOT). The preference for use of an anchor to define the magnitude of change and, thus, generate an MID rather than asking patients to directly define the MID for a particular end point has not been specifically discussed in the literature. However, this preference may be rooted in the fact that this would be a very difficult cognitive task for patients and that expectations may overinflate the estimate, making it difficult to demonstrate that any treatments have resulted in a minimum important change.

Distribution-based methods examine magnitudes of between-group change in the end point relative to some measure of its variability (i.e., on a standardized scale). Researchers have suggested different effect sizes (ES) as indicators of “true” effects beyond chance. The most cited is Cohen’s 0.2, 0.5, and 0.8 to indicate small, moderate, and large ES. [6]. Sloan et al. [2] and Norman et al. [7] have suggested that an ES of 0.5 could be consid-

ered a good estimate of MID for any end point, although they acknowledge that they are not focused on finding a minimal estimate as much as an estimate of an effect that is not ignorable.

Often the goal in the development of a new PRO is to calculate the MID (i.e., a single estimate); however, there is no guidance in the literature to suggest which approach is preferable, or even if more than one approach is used, which estimate should be chosen. The aim of this article was to identify the MID for a patient diary that assesses the number of weekly satisfactory sexual events (SSEs) for women with hypoactive sexual desire disorder (HSDD), an important and key component for any measure but specifically for a female sexual dysfunction (FSD) measure [8]. The validity of an event-driven PRO for assessing changes in sexual function has been questioned [8,9], with one recent study showing the superiority of a questionnaire for assessing FSD outcome compared to a diary [10]. However, given the Food and Drug Administration (FDA) Draft Guidance document, “Guidance for Industry: FSD: Clinical Development of Drug Products for Treatment” stipulating that the primary end point for an FSD study should be SSEs, assessed using a daily diary [11], the MID for this end point was investigated using a diary rather than a questionnaire-based measure. This case study allowed us to address several key questions: First, would a longitudinal anchor-based approach (assumed to be a patient-generated MID) produce a similar estimate to actually asking patients directly what the MID is for them? Second, would an anchor-based approach produce a similar estimate to a distribution-based approach? Finally, if the estimates are different, how should one MID estimate be calculated or chosen?

Method

Subjects

Seven hundred eighty-eight women were recruited into two double-blind, randomized placebo-controlled trials for the treatment of HSDD (randomized controlled trial [RCT] group). These were global studies with women recruited across Europe, United States, Australia, and Canada. The studies were approved by the institutional review board, and each woman provided informed consent. All women were postmenopausal. Diagnosis of HSDD was based on the American Foundation of Urologic Diseases [12]: “persistent or recurrent deficiency (or absence) of sexual fantasies/thoughts, and/or desire for or

receptivity to sexual activity, which causes personal distress” and was made by trained interviewers using a structured diagnostic method (SDM) [13]. Data from these studies were used to calculate the anchor- and distribution-based MID estimates.

A novel patient interview study was also used as a means of obtaining a direct estimate of MID. The patient interview study used a total of 77 women, across three U.S. sites (qualitative group). The women were all postmenopausal and diagnosed as having HSDD using the same SDM process as for the two clinical trials described earlier. This study was also IRB approved, and women provided informed consent. No treatment was given in this observational study.

Procedure

RCT Group

Women were screened for eligibility primarily based on postmenopausal status, diagnosis of HSDD, and particular aspects of their medical history and sexual activity. If eligible, women were given a diary to take home to record baseline levels of satisfaction with sexual events (see Appendix). The diary was based on the FDA Draft Guidance [11]; women had to record whether a sexual event occurred or not each day, whether a recorded event was satisfactory or not, and if satisfactory, what type of activity it was (e.g., sexual intercourse not resulting in your orgasm, sexual intercourse resulting in your orgasm, self-masturbation resulting in your orgasm, or oral sex resulting in your orgasm). As part of the diagnostic process and subsequent assessment of efficacy, women mailed in completed diary pages (each page contained a week’s worth of data) every 2 weeks. Women were also asked to complete a number of questionnaires at baseline: Female Sexual Distress Scale (FSDS) [14], Sexual Function Questionnaire (SFQ) [15], and Sexual Quality of Life-Female measure [16]. The women returned after 6–8 weeks and were given another diary for completion. The women were then treated with daily oral therapy with an experimental drug or placebo for their HSDD. Further study visits occurred at 6 weeks, 3 and 6 months. Diaries were distributed at each visit, and follow-up questionnaires (SFQ and FSDS) were administered at the 6-month visit.

Qualitative Group

Women for this study were recruited based on the same main criteria as the RCT group. Upon meeting the entry criteria, women were asked to complete the same diary as in the clinical trials for

a 6-week period; again women were asked to return the diary pages every 2 weeks. Following completion of the 6-week diary, each woman returned for an interview. The standardized interview guide was administered by one of five interviewers. The women were sequentially (i) asked to supply an ideal number of SSEs, (ii) made aware of the actual number of SSEs from analysis of the 6-week diary, (iii) asked at what point a change would *not* be meaningful (i.e., where a change would not be perceived as noticeable and important), and (iv) asked what a minimally important change would be. As the interview progressed, revisions to the ideal, no change, and minimally important difference values could be made. A guide was provided to help the women quantify sexual events over time: for example, one in 1 year, one in 8 months, one in 6 weeks, two in 1 week, one in 1 day. These values were then converted into weekly values (e.g., one in 1 year = 0.02 per week and one in 3 weeks = 0.33 per week).

Analyses

Anchor-Based Methods

The anchor used for determining MID was defined as Question 34 of the SFQ (SFQ34: Over the last 4 weeks, taking the whole of your sexual life into account, how satisfied have you been? Response options: not, slightly, moderately, very, and extremely satisfied). This anchor was chosen because it had face validity with the primary end point of improving satisfaction with sexual events. This question, being part of the SFQ, allowed both a baseline measurement and end-of-treatment assessment; therefore, a longitudinal analysis could be conducted. MID was defined as the mean change from baseline of average weekly SSEs between subjects who improved and those subjects who reported no change on SFQ34. Improvement was defined as the change from baseline to month 6 of exactly one category on SFQ34 (e.g., not satisfied to slightly satisfied, slightly satisfied to moderately satisfied). Stable or no change referred to those who stayed in the same category from baseline to month 6 on SFQ34 (e.g., not satisfied to not satisfied, slightly satisfied to slightly satisfied).

In calculating MID, observations were pooled without regard for treatment when computing the estimate.

Distribution-Based Methods

Standard effect size (pooled baseline standard deviation (SD) of the end point [6]) was used.

Table 1 Summary statistics for weekly total sexual events and satisfactory sexual events (SSEs)—qualitative group (N = 77)

	Ideal SSEs	Total sexual events	Baseline SSEs	No change SSEs	MID SSEs
Mean (SD)	3.14 (3.52)	1.42 (1.14)	0.98 (0.73)	0.23 (0.16)	0.35 (0.32)
95% CI	2.3, 3.9	1.2, 1.7	0.8, 1.1	0.2, 0.3	0.3, 0.4
Median (IQR)	2.00 (1.25)	1.17 (1.08)	0.83 (1.00)	0.20 (0.21)	0.25 (0.33)
Minimum, maximum	0.04, 21.00	0.33, 7.17	0.00, 3.5	0.00, 1.00	0.02, 2.00

Ideal SSEs = average weekly ideal number of SSEs (theoretical construct).

Total sexual events = total number of sexual events reported in the 6-week diary period, transformed to a weekly average.

Baseline SSEs = number of SSEs reported in the 6-week diary period, transformed to a weekly average.

No change SSEs = average weekly number of SSEs that indicates “no change,” defined as a change not perceived as noticeable and important (theoretical construct).

MID SSEs = average weekly number of SSEs that reflect a minimally important difference (theoretical construct).

SD = standard deviation; IQR = interquartile range; MID = minimum important difference; CI = confidence interval.

Cohen suggested different ES (hereafter called MID-ES) ranging from 0.2 (small) to 0.5 (moderate) to 0.8 (high) [6]. For example,

$$\frac{\text{MID}}{\text{Measure of variability}} \approx 0.5$$

would provide an estimate of the MID to be one half of the SD. The MID would be approximately 0.5 SD. In our work, both 0.2 ES and 0.5 ES were analyzed. Ninety-five percent confidence intervals (CIs) for the distribution-based MID were based on the standard 95% CI for the population variance σ^2 [17].

Qualitative Study

The ideal number of SSEs, total number of sexual events, total number of SSEs, change not perceived as noticeable and important, and the minimally important change were summarized descriptively. All diary data were reported per 1-week period.

Spearman correlation coefficients were calculated between MID and baseline SSEs (both measured on a weekly basis) to explore how a woman’s experience influenced her assessment of MID.

Results

Demographics

RCT Group

The mean age was 54.8 years (range 45–75 years); 93% were Caucasian. On average, the women were 7 years postmenopausal. Mean baseline SSE was 0.58 (SD = 0.51) and mean “total sexual events” were 1.14 (SD = 0.79) per week.

Qualitative Group

Overall, the mean age was 55 years, ranging between 48 and 74 years. The majority was Caucasian (78%). Mean years since menopause ranged

from 7 to 8 years. Mean baseline SSEs, and mean total sexual events per week can be seen in Table 1.

SSE MID Estimates from Anchor- and Distribution-Based Methods for the RCT Group

Table 2 summarizes the results for the anchor- and distribution-based estimates of the MID. The longitudinal approach gave a similar estimate to the 0.5 ES estimate (0.33 vs. 0.26, respectively). The 0.2 ES estimate was somewhat smaller with an estimate of 0.10.

SSE MID Estimate from the Qualitative Study

General descriptive summary statistics for weekly SSEs can be seen in Table 1. The mean MID for the total sample (N = 77) was 0.35 SSEs per week (95% CI 0.3, 0.4), and the median was 0.25 SSEs per week. This discrepancy between the mean and the median reflected that there were several women with extremely high MID estimates, resulting in a positively skewed distribution of MIDs. The relationship between MID and baseline SSEs was a moderate-to-large positive association; if a woman’s average weekly number of SSEs at baseline was high, she was more likely to report a high MID (Spearman’s $r = 0.42$, $P < 0.001$).

Calculating a Single Estimate for MID

To facilitate deciding on an estimate of MID, a single estimate of MID for average weekly SSEs

Table 2 Weekly SSE MID estimates and 95% CIs for the anchor-based and distribution-based methods—RCT group

	N	MID (95% CI)
Anchor-based method		
Longitudinal	328	0.33 (0.20, 0.46)
Distribution-based method		
ES = 0.2	726	0.10 (0.10, 0.11)
ES = 0.5		0.26 (0.24, 0.27)

ES = effect size; CI = confidence interval; N = number of subjects; MID = minimum important difference; SSE = satisfactory sexual events; RCT = randomized controlled trial.

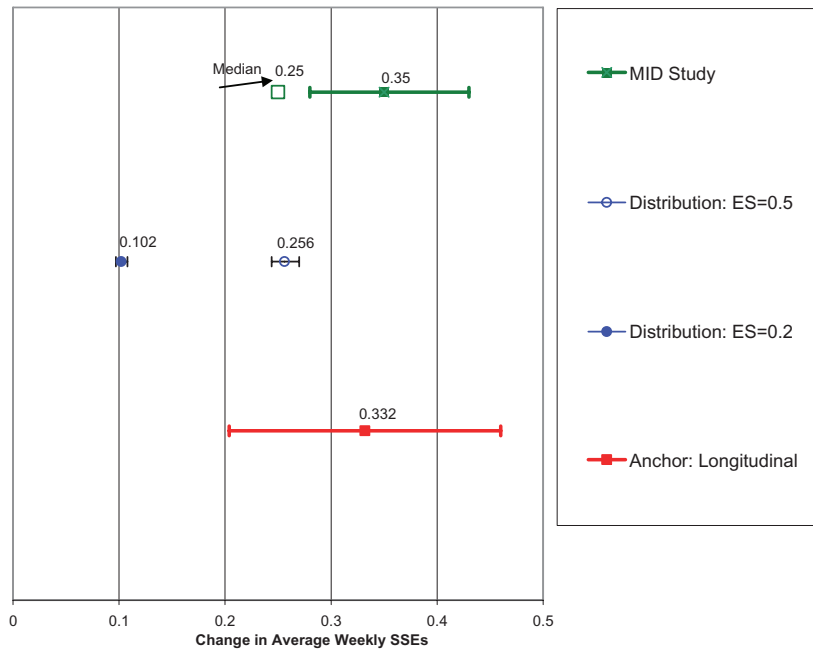


Figure 1 Minimum important difference (MID) estimates (with 95% confidence interval) using anchor-based, distribution-based, and -patient interview study (MID study) methods. ES = effect size; SSE = satisfactory sexual event.

could be selected based on an interpretable value that was in the “middle” of the set of estimates formed by the limits of the 95% CIs. By looking at all of the estimates combined (see Figure 1), a middle value of 0.2 SSEs per week might be proposed as the MID. However, a range for the MID might be more appropriate given the wealth of information generated by use of the different methods.

Discussion and Conclusions

We set out to address three questions regarding how best to define an MID for a new measure, and we used the case study of SSEs per week in postmenopausal women with HSDD:

1. Would asking patients directly what an MID was be different from that estimated from an anchor-based approach?
2. Would the anchor- and distribution-based approaches produce similar estimates?
3. If the estimates are different, how should the MID estimate be made?

The longitudinal approach provided a slightly higher estimate than the distribution-based approach), but the 95% CIs did overlap. When looking at clinical trial data, perhaps the use of the longitudinal approach is more appropriate given that this is looking at change over time; therefore, the MID estimate should reflect this also. However, it is known that the type of anchor ques-

tion used can influence the estimate generated [18] and is not always valid [19], so care should be taken in choosing the most appropriate anchor. One way of determining if an anchor is relevant for estimating the MID is to look at its relationship to the end point under investigation. The anchor should be negatively correlated with the end point (in this case SSEs) at baseline and positively correlated with the end point at the EOT. We had a 0.23 correlation at baseline and 0.50 at EOT. Hayes et al. [20] recently stated that if the correlation between the anchor and health-related quality of life measure is zero, then the anchor will be poor in determining the MID. They recommended a correlation above 0.37 (based on Cohen’s 0.8 ES) as a correlation threshold that indicates a large association between the two measures. This would indicate that the anchor used in this study is appropriate and valid. However, a more specific question directly asking about satisfaction with sexual events might be more appropriate than the “taking the whole of your sexual life into account, how satisfied have you been” question that was used in this study. Further research to investigate if this would produce a different MID estimate would be interesting.

Being aware that using a single item to estimate MID is problematic, an approach asking patients directly was explored. This approach gave an estimate similar to that calculated from the longitudinal anchor-based approach, which increases our confidence in the anchor used. The direct-patient

approach is novel in that we asked patients what an MID would be for them in terms of SSEs, by working through an exercise based on their recent experience of SSEs in the past 6 weeks rather than from an anchor question asked during a study. Redelmeir et al. [21] used a similar approach by asking patients directly, albeit using a hypothetical situation where patients had to determine if they were better or worse than the person they were partnered with in the study. Subjects were not asked directly about an MID for changes in dyspnea, fatigue, emotion, or coping. A direct approach may have previously been avoided because of unrealistic expectations by the patient leading to unrealistically high MIDs. Also, it is cognitively difficult to comprehend the concept of MID (i.e., *minimum* important difference). However, the approach we used seems to have some validity because there was concordance with the anchor-based estimate.

The anchor- and distribution-based estimates were similar, although the 0.2 ES estimate was somewhat smaller than the anchor-based approach. The 0.2 ES may be a closer estimate to a cross-sectional approach estimate because this is not dissimilar to findings by Kulkarni [22] who used a cross-sectional anchor-based approach and 0.2 ES and found similar estimates across the two approaches for the Hydrocephalus Outcome Questionnaire. Based on this research, the 0.5 ES and direct questioning approach may reflect more the longitudinal-based approach estimate.

Given that there was spread in the estimates, defining a range might be more appropriate, as proposed by other researchers, for example, Guyatt et al. [1], Hayes et al. [20], and Marquis et al. [23]. In using the CIs, an idea would be to use the anchor-based upper bound and the lower bound of the distributional approach (0.5 SD), this would give a range of estimates of 0.24–0.46 SSEs per week. If we used the lower bound of the 0.2 SD estimate, this would result in a range of 0.10–0.46.

The aim of this article was to explore the concordance of different methods to define the MID of SSEs in women with HSDD; therefore, further work is needed to investigate the MID of SSEs in other FSD disorders (e.g., female sexual arousal disorder) or different diary instruments. Furthermore, because global assessment of change data was not collected in the RCTs, it would be of particular interest for future work to use a global assessment of change as the anchor (the most common approach), in addition to or instead of,

the SFQ34 that was used in the anchor-based analysis for the RCTs.

There are no easy answers when trying to define the MID for a new measure. Experience and, therefore, time will help. The methodology used to ask patients directly seemed to be successful, but it is resource intense and did not result in a markedly different estimate than those produced by the anchor- and distribution-based approaches. We recommend that researchers use both anchor-based (cross-sectional and longitudinal) and distribution-based (0.5 ES) approaches as well as the 95% CI to help bound the MID estimate range.

Corresponding Author: Tara Symonds, PhD, World-wide Outcomes Research, Pfizer Ltd. (IPC160), Sandwich, Kent, CT13 9NJ, UK. Tel: +44-1304-641553. Fax: +44-1304-658823; E-mail: tara.symonds@pfizer.com

Conflict of Interest: Drs. Symonds, Soni, Sisson, and Spino are all Pfizer employees. Dr. Patrick is a paid consultant.

Statement of Authorship

Category 1

(a) Conception and Design

Tara Symonds; Cathie Spino; Melanie Sisson; Paresh Soni; Mona Martin; Lacey Gunter; Donald L. Patrick

(b) Acquisition of Data

Tara Symonds; Cathie Spino; Mona Martin; Melanie Sisson

(c) Analysis and Interpretation of Data

Tara Symonds; Cathie Spino; Melanie Sisson; Paresh Soni; Mona Martin; Lacey Gunter; Donald L. Patrick

Category 2

(a) Drafting the Article

Tara Symonds; Cathie Spino; Melanie Sisson; Paresh Soni; Mona Martin; Lacey Gunter; Donald L. Patrick

(b) Revising It for Intellectual Content

Tara Symonds; Cathie Spino; Melanie Sisson; Paresh Soni; Mona Martin; Lacey Gunter; Donald L. Patrick

Category 3

(a) Final Approval of the Completed Article

Tara Symonds; Cathie Spino; Melanie Sisson; Paresh Soni; Mona Martin; Lacey Gunter; Donald L. Patrick

References

- 1 Guyatt GH, Osoba D, Wu AW, Wyrich KW, Norman GR, Clinical Significance Consensus Meeting Group. Methods to explain the clinical

- significance of health status measures. *Mayo Clin Proc* 2002;77:371–83.
- 2 Sloan J, Symonds T, Varga-Chanes D, Fridley B. Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Inf J* 2003;37:23–31.
 - 3 Wyrwich K, Bullinger M, Aaronson N, Hays R, Patrick D, Symonds T, the Clinical Significance Consensus Group. Estimating clinically significant differences in quality of life outcomes. *Qual Life Res* 2005;14:285–95.
 - 4 Juniper EF, Guyatt GH, Willam A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994;47:81–7.
 - 5 Samsa G, Edelman D, Rothman M, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: A general approach with illustration to the Health Utilities Index Mark II. *Pharmacoecon* 1999;15:141–55.
 - 6 Cohen J. *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press Inc.; 1977.
 - 7 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health related quality of life: The remarkable universality of half a standard deviation. *Med Care* 2003;41:582–92.
 - 8 Althof SE, Dean J, Derogatis LR, Rosen RC, Sisson M. Current perspectives on the clinical assessment and diagnosis of female sexual dysfunction and clinical studies of potential therapies: A statement of concern. *J Sex Med* 2005;3(suppl 2):146–53.
 - 9 Althof SE, Rosen RC, Derogatis L, Corty E, Quirk F, Symonds T. Outcomes measurement in FSD clinical trials: Review and recommendations. *J Sex Marital Ther* 2005;31:153–66.
 - 10 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.
 - 11 FDA Guidance Document. Guidance for Industry: FSD: Clinical development of drug products for treatment. 2000. Available at: <http://www.fda.gov/cder/guidance/3312dfthtm> (accessed July 11, 2007).
 - 12 Basson R, Berman J, Burnett A, Derogatis L, Ferguson D, Fourcroy J, Goldstein I, Graziottin A, Heiman J, Lann E, Leiblum S, Padma-Nathan H, Rosen R, Segraves K, Segraves RT, Shabsigh R, Sipski M, Wagner G, Whipple B. Report of the International Consensus Development Conference on Female Sexual Dysfunction: Definitions and classifications. *J Urol* 2000;163:888–93.
 - 13 Utian WH, MacLean DB, Symonds T, Symons J, Somayaji V, Sisson M. A methodology study to validate a structured diagnostic method used to diagnose female sexual dysfunction and its subtypes in postmenopausal women. *J Sex Marital Ther* 2005;31:271–83.
 - 14 Derogatis L, Rosen R, Leiblum S, Burnett A, Heiman J. The Female Sexual Distress Scale (FSDS): Initial validation of a standardized scale for assessment of sexually related personal distress in women. *J Sex Marital Ther* 2002;28:317–30.
 - 15 Quirk FH, Heiman JR, Rosen RC, Laan E, Smith MD, Boolell M. Development of a sexual functioning questionnaire for clinical trials of female sexual dysfunction. *J Womens Health Gender-based Med* 2002;11:277–89.
 - 16 Symonds T, Boolell M, Quirk F. Development of a questionnaire on sexual quality of life in women. *J Sex Marital Ther* 2005;31:385–97.
 - 17 Ott L. *An introduction to statistical methods and data analysis*. Belmont, CA: Wadsworth Publishing Company, Inc.; 1977.
 - 18 Barber BL, Santanello NC, Epstein RS. Impact of the global on patient perceivable change in an asthma specific QOL questionnaire. *Qual Life Res* 1996;5:117–22.
 - 19 Norman GR, Stratford P, Rehehr G. Methodological problems in the retrospective computation of responsiveness to change: The lesson of Cronbach. *J Clin Epidemiol* 1997;50:869–79.
 - 20 Hayes RD, Brodsky M, Johnston MF, Spritzer KL, Hui KK. Evaluating the statistical significance of health-related quality of life change in individual patients. *Eval Health Prof* 2005;28:160–71.
 - 21 Redelmeir DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: A comparison of two techniques. *J Clin Epidemiol* 1996;49:1215–9.
 - 22 Kulkarni AV. Distribution-based and anchor-based approaches provided different interpretability estimates for the Hydrocephalus Outcome Questionnaire. *J Clin Epidemiol* 2006;59:176–84.
 - 23 Marquis P, Chassany O, Abetz L. A comprehensive strategy for the interpretation of quality of life data based on existing methods. *Value Health* 2004;7:93–104.

Appendix

Daily Diary

Weekly Diary. To be completed everyday regardless of whether or not sexual activity took place.

- Q1. Did you take part in sexual activity
Yes/No If No, stop here.
- Q2. If you had a sexual experience, please check one of the following:
Unsatisfactory/Satisfactory/Very Satisfactory

Q3. Only if question 2 was Sat. (satisfactory) or Very Sat. (very satisfactory), check *all* that apply.

- Sexual intercourse not resulting in your orgasm
- Sexual intercourse resulting in your orgasm

- Self-masturbation resulting in your orgasm
- Partner masturbation resulting in your orgasm
- Oral sex resulting in your orgasm
- Other sex resulting in your orgasm