The Sexual Event Diary (SED): Development and validation of a standardized questionnaire for assessing female sexual functioning over discrete sexual events

Introduction

Low sexual desire and arousal are the most common sexual complaints among women, and commonly cause sexual dissatisfaction and personal distress [1]. These conditions were classified in the DSM-IV Text Revision as Hypoactive Sexual Desire Disorder (HSDD) and Female Sexual Arousal Disorder (FSAD) [2] respectively, but have been merged in the DSM-5 as Female Sexual Interest/Arousal Disorder (FSIAD) [3].

The pharmacotherapeutic options for HSDD/FSIAD are limited, with only one approved drug on the market in the US [4]. This drug, flibanserin, is taken daily to increase overall sexual desire. There are other therapies in late stages of clinical development [5-8] that are not taken daily, but instead taken on-demand, i.e. when a woman with HSDD/FSIAD *wants to want* to have sex. These medications are not intended to increase sexual desire continuously, but only prior to and during sexual activity. Measuring the efficacy of such an on-demand drug necessitates a different approach.

Flibanserin's efficacy was assessed using the Female Sexual Function Index (FSFI). The FSFI assesses different dimensions of female sexual functioning over the preceding four weeks [9]. The efficacy of an on-demand drug for HSDD/FSIAD is best determined by assessing the quality of a sexual event during which the drug was taken. Assessing sexual functioning retrospectively over a longer period of time, for example over four weeks as does the FSFI, gives a more distal estimation of an on-demand drug's influence on sexual functioning than assessing sexual functioning over the actual events during which the drug was taken. However, in order to determine an on-demand drug's efficacy, an estimation of long-term effects is necessary. This can be operationalized by evaluating the change in number of satisfactory sexual events (SSE) from a baseline establishment period (BLE) to an active treatment period (ATP) during which the on-demand therapy was used. The primary endpoint in such trials is the difference between active treatment and placebo treatment arms in the change in number of SSEs from baseline to end of treatment, which is one of the U.S. Food and Drug Agency's preferred primary endpoints for the

indication HSDD/FSIAD [10]. For this, a standardized and validated sexual event questionnaire is necessary.

The aim of this research was to develop and validate a standardized event log for assessing sexual satisfaction and sexual functioning of a single sexual event. This patient reported outcome, the Sexual Event Diary (SED), underwent three cycles of development, starting with a 58-item version, followed by a 16-item version, and finally, an 11-item version. This patient-reported outcome instrument was developed to gather primary and (key) secondary endpoint data in clinical trials assessing the efficacy of on-demand drugs in women with HSDD/FSIAD.

Materials and methods

Questionnaire development

The first version of the SED included 58 items, which were selected based on literature review, expert opinion, and information from over 250 clinical interviews that were conducted at our laboratory with women having sexual problems. The items that were included were selected to provide a comprehensive representation of sexual functioning and sexual satisfaction of a sexual event. Three focus groups, two with five premenopausal women and one with five postmenopausal women, with (predominantly) sexual problems, were performed to discuss what constituted sexual satisfaction and whether the 58-item SED adequately measured satisfaction as well as all other relevant aspects of sexual functioning.

The Dutch pilot version of the 58-item SED was tested in 156 women with (n=89) and without (n=67) sexual problems. These data were used for the initial validation and item reduction. Aside from completing the SED regarding their last sexual event, the subjects were asked to select those 15 SED items that were most relevant to them in capturing sexual satisfaction and sexual functioning during an event. Principal Components Analysis was performed to determine the factors underlying the SED. Correlations of the items with global sexual satisfaction and Cronbach's alpha coefficients were calculated in order to assess internal consistency (reliability).

The goal was to develop a comprehensive and compact questionnaire that could adequately assess the quality of a sexual event without burdening the subject. Based on the gathered qualitative and psychometric assessments, the 58-item SED was reduced to a 16-item version. This 16-item version was subsequently translated into US English by a certified medical translation office in the Netherlands (Wilkens c.s., Medical Translations, Leiden, The Netherlands). Two female interviewers with experience in women’s sexual medicine (RTI Health

Solutions, Research Triangle Park, NC, USA) performed cognitive debriefing interviews with five native US-English speaking women to test the adequacy of the translated version.

All participants of the focus groups, debriefing interviews and observational study described above provided written informed consent.

The first versions of the SED were called the 'Satisfaction of an Event Questionnaire' (SSEQ), but it was later renamed to SED because this name covered the content of the questionnaire more adequately. In the present article, only the name SED is used (also to refer to prior versions) for sake of clarity.

Data

Clinical studies

Reliability, validity, and responsiveness of the 16-item US-version of the SED were assessed using data collected during two clinical studies in the United States. Study 1 (clinical trail identifier: NCT01432665) investigated the efficacy and safety of on-demand use of 4 doses of the combined administration of sublingual testosterone (0.25 mg, 0.5 mg) and sildenafil (25 mg, 50 mg), compared to placebo and monotherapies, in women with HSDD with low sensitivity for sexual cues. Study 2 (NCT01743235) investigated the efficacy and safety of on-demand use of 4 doses of the combined administration of sublingual testosterone (0.25 mg, 0.5 mg) and buspirone (5 mg, 10 mg), compared to placebo and monotherapies, in women with HSDD and dysfunctional over-activation of sexual inhibitory mechanisms [5]. Following subsequent qualitative and quantitative validation, the 16-item SED was modified into the 11-item SED (see Appendix). Finally, the reliability, validity, and responsiveness of the 11-item SED was assessed using data collected in a third US study (NCT02101203) [5], in which the efficacy and safety of on-demand use of sublingual testosterone (0.5 mg) combined with buspirone (10 mg) was compared to placebo, in women with HSDD and dysfunctional over-activation of sexual inhibitory mechanisms. A total of 5281 16-item SED's were filled out by 188 women who were in the Intention To Treat (ITT) population of study 1. A total of 4604 16-item SED's were filled out by 183 women who were in the ITT population of study 2. A total of 1074 11-item SED's were filled out by 50 women of study 3. The ITT population contained 52 women, but data from two women were excluded because one had only incomplete SEDs and one completed her SEDs too long after the sexual events occurred. All completed questionnaires were used for the statistical analyses.

All three studies consisted of a 4-week BLE, two 4-week placebo run-in periods (PRI), and two 4-week ATP. These studies investigated the effect of on-demand therapies' influence on

discrete sexual events. Because frequency of sexual events varies per individual, the number of sexual events for each subject in each study and each 4-week period varied, and thus the number of collected SEDs for each subject.

The SED was filled out on a secure web-based system (Viedoc™ Me, Pharma Consulting Group, Uppsala, Sweden) that the participants could access at home via computer or from their portable device.

All participants of the three trials provided written informed consent.

Levels of assessment

Reliability and validity of the 16- and 11-item SED were assessed at event level and at subject level. Analyses were performed on event level to establish reliability and validity of the questionnaire in its ability to assess sexual function of a discrete sexual event, and on subject level to assess sexual function of an individual. Thus subject level can be used to establish reliability and validity of the primary endpoint change in number of SSEs from BLE to ATP. For event level analyses, SED's filled out by the same woman were treated as independent events. For subject level analyses, validity and reliability of subject mean scores were assessed over 4-week periods. The SED mean scores in the BLE and those in the ATP period were calculated separately. The PRI mean scores were not included in the analyses at subject level.

For assessing known groups validity, SED mean scores in satisfying sexual events were compared with those in unsatisfying sexual events as reported by SED item 4 only on event level. For evaluating responsiveness on subject level, SED mean change scores from baseline in those subjects reporting study medication-dependent improvement could be compared with those reporting no study medication-dependent improvement. Medication-dependent improvement was assessed by the Subjective Evaluation of Gain questionnaire (SEG). This questionnaire was also administered in the 3 studies, and item 1 asked if subjects had experienced improvement in their sexual functioning attributable to the study medication in the preceding 4 weeks.

Listwise deletion was used to handle missing data and this resulted in the deletion of 13 of the 5281 events for study 1, in the deletion of 1 of 4604 events in study 2, and in the deletion of 10 of the 1074 events in study 3. At subject level, this resulted in deletion of 23 of 376 observations in study 1, 48 of 366 observations in study 2 and 17 of 100 observations in study 3, mostly due to no events being reported in the BLE or ATP period.

Quantitative assessment

Factor analysis

Exploratory factor analysis (EFA) was used to assess underlying dimensions of the Likert scale items (items 5 through 15 for study 1 and 2 and items 5 through 10 for study 3) at event and at subject level. At event level, EFA was conducted on the polychoric correlation matrices of the Likert scale items, because these data are ordinal. At subject level, Pearson correlation matrices of the average Likert scale items were used. Factor analysis was estimated using the maximum likelihood (ML) method. The number of factors to be retained was determined by inspecting the eigenvalues and scree-plots. Furthermore, Parallel Analyses (PA) based on a minimum rank factor analysis [11] were conducted to find the number of factors under possible violations of the multivariate normality assumption.

Reliability

Internal consistency was assessed using Cronbach's alpha coefficient, that provides a lowerbound for reliability. Inter-item and item-rest correlations were assessed using Pearson correlation coefficients. Inter-item correlations were used to assess the relationship between individual items within the SED. Item-rest correlations were used to assess the relationship between individual items and the total item sum score of the remaining items. Two different sum scores were used: an unweighted sum score of all SED Likert scale items, referred to as the SED total score, and an unweighted SED sexual function sum score, consisting of all relevant SED Likert scale items based on the results of the factor analyses and/or theoretical arguments. For the calculation of the SED sum score(s) and all statistical analyses, the answer categories of 16-item SED items 11 ('afraid of pain'), 12 ('disgust'), and 13 ('distracting thoughts') and 11-item SED item 8 ('distracting thoughts') were reversed, so that the answer categories of all items had the same direction.

Validity

At event level, construct validity was assessed by comparing the means of the SED item scores that were included in the scale with the dichotomous items assessing satisfaction, gratification and orgasm (answer options: "yes" and "no").

At subject level, construct validity was assessed using the Pearson correlation coefficients of the total score and domain scores of the Sexual Function Questionnaire [9] (SFQ; studies 1 and 2) and the Female Sexual Function Index [12] (FSFI; study 3) with the sum scores and related

domain items of the SED. The SFQ is a 34-item self report questionnaire that assesses seven domains of sexual function over the preceding 4 weeks: desire, arousal–sensation, arousal–lubrication, enjoyment, orgasm, pain, and partner relationship. The FSFI is a 19-item self report questionnaire that assesses six domains of sexual function over the preceding 4 weeks: desire, arousal, lubrication, orgasm, satisfaction and pain.

Also at subject level, known groups validity was assessed by comparing the mean SED scores over the ATP period between responders and non-responders, using independent sample T-test statistics. Responders were defined as those subjects who indicated experiencing improvement in the last 4 of the 8 weeks treatment using the SEG. Responders were subjects who answered "yes" to the question asking if they had experienced benefit from the medication during the last 4 weeks, and non-responders were those who answered "no". The responder classification was independent of study medication used.

Responsiveness

Responsiveness is the ability of the instrument to detect change when there is a known change in the measurement of interest. Responsiveness was assessed by comparing the means of the change from BLE to ATP in SED scores between responders and non-responders, by calculating independent sample T-test statistics.

Responsiveness was also assessed by determining the effect size statistics of the ability of the SED to measure change in sexual functioning using Guyatt's responsiveness index [13, 14].

$$Guyatt's\ index = \frac{(Change\ in\ SED\ scores\ for\ Responders) - (Change\ in\ SED\ scores\ for\ Non-Responders)}{Standard\ Deviation\ of\ Change\ in\ SED\ scores\ for\ Non-Responders}$$

Effect sizes of about 0.20 represent small effects, those of about 0.50 represent moderate effects, and those ≥ 0.80 represent large effects [13].

A two-sided 5% significance level was adopted for all statistical tests.

Qualitative assessment

Two iterative sets of cognitive debriefing interviews were held with women with HSDD in the United States by RTI Health Solutions, Research Triangle Park, NC, USA. Each set of cognitive interviews was held with 8 women. The purpose of the first set of interviews was to assess the content validity of the 16-item SED. The second set of interviews was carried out following SED adaptation, to confirm content validity and finalize the 11-item SED

Population cognitive debriefing interviews

All women had clinician-diagnosed HSDD. Comorbidity with female sexual arousal disorder and/or female orgasmic disorder (only as secondary diagnosis) was allowed. The participants were between the age of 21 and 70 year and were sexually active (i.e., vaginal penetration, sexual intercourse) in the past 3 months. They were able to read and speak English and all provided written consent prior to study entry. Women were excluded if they were pregnant or lactating, had other unexplained gynecological complaints such as clinically relevant abnormal uterine bleeding patterns, and/or had current sexual disorder of vaginismus or dyspareunia according to the DSM-IV or DSM-IV-TR.

Patients were recruited from a clinical site in West Palm Beach, Florida for the first round of interviews and from a clinical site in Houston, Texas, for the second round. All patients provided written consent and received $100 reimbursement. All study materials were reviewed and approved by RTI International's institutional review board (IRB) before any participants were recruited for the study.

Debriefing interview methods

The one-hour interviews in both rounds were led by the same two female interviewers (RTI Health Solutions, Research Triangle Park, NC, USA), both with experience in women's sexual medicine, using a semi structured interview guide. At the start of each interview in round 1, participants were asked to engage in a brief open-ended discussion describing their definitions of sexual desire and of satisfaction with sexual activity. Participants were then asked to provide feedback on the 16-item SED items while describing their thought processes out loud. The interviewers also asked targeted questions to get further information about the way in which the participants interpreted the items and thought about the response options. At the close of the debriefing interview, the interviewers asked whether the 16-item SED missed any important concept that participants deemed critical to measuring satisfaction with a sexual event and whether any items included in the SED seemed irrelevant to participants for inclusion in a sexual event diary. In round 2, at the start of each interview, participants were asked to engage in a brief open-ended discussion again describing their definitions of satisfaction with sexual activity. Cognitive debriefing of the 11-item SED was conducted using a similar "think-aloud" procedure and directed probes to delve further into the question-answering process. At the end of each patient interview, participants were asked whether the 11-item SED missed any important concept that participants deemed critical to measuring satisfaction with a sexual event and whether any items included in the 11-item SED seemed irrelevant to participants for inclusion in

a sexual event diary. Finally, from the concepts contained in the 11-item SED, participants were asked to select the three concepts that were most important to them in determining satisfaction with sexual activity.

Results

Participant characteristics

Baseline characteristics and demographics of women with HSDD who were included in psychometric assessments and who participated in the debriefing interviews are shown in Table 1.

Table 1. Baseline characteristic and demographics of study participants

	Psychometric assessment						**Qualitative assessment**			
	Clinical study						**Debriefing interviews**			
	1		**2**		**3**		**Round 1**		**Round 2**	
	n= 188		**n= 183**		**n=50**		**n= 8**		**n= 8**	
Menopausal status[§], n (%)										
Premenopausal	134	(71.3)	132	(72.1)	30	(60.0)	2	(25.0)	0	(0.0)
Postmenopausal	54	(28.7)	51	(27.9)	20	(40.0)	6	(75.0)	8	(100)
Age, mean (range)	44.2	(22-65)	43.7	(24-67)	46.1	(23-66)	50.0	(29-62)	61.8	(54-69)
Race, n (%)										
Caucasian	152	(80.9)	112	(61.2)	31	(63.0)	3	(37.5)	5	(62.5)
Black	18	(9.6)	56	(30.6)	14	(28.0)	3	(37.5)	2	(25.0)
Asian	4	(2.1)	1	(0.5)	0	(0.0)	0	(0.0)	0	(0.0)
Other	14	(7.4)	14	(7.7)	5	(10.0)	2	(25.0)	1	(12.5)
Clinician secondary diagnosis, n (%)										
FSAD	25	(13.3)	36	(19.7)	13	(26.0)	2	(25.0)	1	(12.5)
FOD	18	(9.6)	19	(10.4)	4	(8.0)	5	(62.5)	0	(0.0)
FSAD and FOD	27	(14.4)	30	(16.4)	12	(24.0)	0	(0.0)	0	(0.0)
none	118	(62.8)	98	(53.6)	21	(42.0)	1	(12.5)	7	(87.5)

§ Perimenopausal women were not included in the three clinical trials because the variable nature of their hormonal status could impact the studies' results in an unpredictable manner. Abbreviations: FSAD= Female Sexual Arousal Disorder, FOD= Female Orgasmic Disorder.

16-item SED (study 1 and 2)

Event level

Factor analysis

Inspection of the eigenvalues and scree-plots of the ML factor analyses and results of the PA revealed that one factor should be retained. The first factor had an eigenvalue of 5.48, explaining 49.8% of the variance for study 1 and an eigenvalue of 5.56, explaining 50.5% of the variance for

study 2. Items 5 through 10 and,15 had moderate (>.50) to strong (>.80) loadings, with less strong contributions of items 12, 13 and 14 (Table 2). The SED sexual function sum score was derived using items 5 through 10 and items 13 through 15. Item 11 and 12 were not included because overall they had the lowest factor loadings.

Reliability

Items 5 through 10 and 13 through 15 gave a high Cronbach's alpha for both studies (see Table 3). Most inter-item Pearson correlation coefficients were over 0.30, see ***Supplementary, Tables A and B.*** Only items 11 ('afraid pain'), 12 ('disgust') in both studies and 13 ('distracting thoughts') in study 1 had correlations below 0.30. Item-rest Pearson correlation coefficients between 16-item SED items 4 ('satisfied'), 3 ('gratified') and 16 ('orgasms'), with answer options "yes"/ "no", and the sum scores were all larger than 0.30 in both studies 1 and 2. Here the item-rest correlation coefficients between items 4 ('satisfied') and 3 ('gratified') and the sum scores were somewhat larger compared to 16 ('orgasm') for both studies. The item-rest Pearson correlation coefficients between the 16-item SED Likert items 5 to 15 and the sum scores were mostly larger than 0.30, except for item 11 ('afraid of pain') for both studies.

Table 2. Maximum Likelihood Factor Analyses: Factor loadings of the 1-factor solution for the SED items

	1-Factor solution						
	Study 1		Study 2			Study 3	
16-item SED	event	subject	event	subject	11-item SED	event	subject
5. 'Pleasurable'	0.918	0.915	0.924	0.933	5. 'Sexual desire'	0.953	0.934
6. 'Letting go'	0.848	0.850	0.856	0.841	6. 'Mentally aroused'	0.966	0.981
7. 'In the mood for sex'	0.851	0.856	0.855	0.830	7. 'Physically aroused'	0.960	0.962
8. 'Vaginal arousal'	0.893	0.920	0.939	0.928	8. 'Distracting thoughts'§	0.668	0.594
9. ' Sexually aroused'	0.951	0.960	0.964	0.956	9. 'Letting go'	0.884	0.803
10. 'Body image'	0.566	0.484	0.546	0.526	10. 'Pleasurable'	0.947	0.952
11. 'Afraid pain'§	0.125	0.087	0.176	0.120			
12. 'Disgust'§	0.420	0.338	0.397	0.283			
13. 'Distracting thoughts'§	0.410	0.370	0.490	0.442			
14. 'Partner attractiveness'	0.500	0.466	0.451	0.420			
15. 'Partner's actions'	0.749	0.796	0.709	0.703			

§Reversed variable. Remarks: Dichotomous items 1-4, and item 11/16 were not included in the Factor Analyses. Item 11/16 are the same item ('Orgasm'), but numbered 11 in the 11-item SED and 16 in the 16-item SED.

Table 3. Cronbach's alpha coefficients of all items on event level and subject level

		Internal consistency#		
Level of analysis	**Items**	**study 1**	**study 2**	**study 3**
Event level	5 – 15	0.89	0.89	
	5, 6, 7, 8, 9, 10, 13§, 14, 15*	0.91	0.91	
	5 – 10*			0.95
Subject level	5 – 15	0.90	0.90	
	5, 6, 7, 8, 9, 10, 13§ ,14 ,15*	0.92	0.92	
	5 – 10*			0.95

*items that loaded on factor 1 and/or which were theoretically important for measuring sexual function
#Cronbach's alpha (range -1.00 to +1.00). §Reversed variable

Validity

All items that were included in the 16-item SED of study 1 and 2 showed strong known groups validity, since the mean differences in SED item and sum scores between "yes" and "no" responders on SED items measuring 'gratified', 'satisfied', and 'orgasm' were highly significant (all but one were $P<0.0001$), all in the expected direction (Table 4).

Table 4. Known groups validity: comparison between 16-item SED item and sum score means (SE) and "Yes" and "No" answers on 16-item SED items 3, 4 and 11

	3. Did you find this sexual activity gratifying?						4. Were you satisfied with the sexual activity?						16. Did you have an orgasm?					
	Yes		No		Test statistics		Yes		No		Test statistics		Yes		No		Test statistics	
SED item	mean	SE	Mean	SE	t-value	p-value*	mean	SE	Mean	SE	t-value	p-value*	mean	SE	Mean	SE	t-value	p-value*
Study 1																		
5. 'Pleasurable'	3.68	0.02	1.86	0.02	78.18	<0.0001	3.69	0.02	1.99	0.02	69.12	<0.0001	3.80	0.02	2.47	0.02	49.12	<0.0001
6. 'Letting go'	3.62	0.02	2.08	0.03	49.80	<0.0001	3.62	0.02	2.22	0.03	44.95	<0.0001	3.74	0.02	2.56	0.02	39.58	<0.0001
7. 'In the mood for sex'	3.27	0.02	1.81	0.02	49.67	<0.0001	3.26	0.02	1.95	0.03	42.17	<0.0001	3.35	0.02	2.30	0.02	33.86	<0.0001
8. 'Vaginal arousal'	3.31	0.02	1.63	0.02	61.82	<0.0001	3.32	0.02	1.74	0.02	55.87	<0.0001	3.49	0.02	2.10	0.02	46.98	<0.0001
9. ' Sexually aroused'	3.41	0.02	1.71	0.02	61.01	<0.0001	3.42	0.02	1.84	0.02	53.67	<0.0001	3.55	0.02	2.25	0.02	42.84	<0.0001
10. 'Body image'	3.67	0.02	2.97	0.03	19.38	<0.0001	3.67	0.02	3.04	0.03	17.80	<0.0001	3.73	0.02	3.19	0.02	16.97	<0.0001
11. 'Afraid pain'§	4.81	0.01	4.71	0.02	4.43	<0.0001	4.80	0.01	4.72	0.02	3.72	0.0002	4.84	0.01	4.72	0.02	6.48	<0.0001
12. 'Disgust'§	3.98	0.02	3.36	0.03	11.65	<0.0001	4.83	0.01	4.55	0.02	11.83	<0.0001	4.84	0.01	4.64	0.02	10.31	<0.0001
13. 'Distracting thoughts'§	3.98	0.02	3.36	0.03	18.01	<0.0001	4.00	0.02	3.36	0.03	19.14	<0.0001	4.08	0.02	3.5	0.02	19.91	<0.0001
14. 'Partner attractiveness'	3.89	0.02	3.33	0.03	15.83	<0.0001	3.88	0.02	3.41	0.03	13.24	<0.0001	3.83	0.02	3.63	0.02	5.99	<0.0001
15. 'Partner's actions'	3.66	0.02	2.37	0.03	38.48	<0.0001	3.65	0.02	2.50	0.03	33.59	<0.0001	3.66	0.02	2.89	0.02	23.05	<0.0001
SED Total	42.83	0.12	30.45	0.15	63.90	<0.0001	42.86	0.12	31.43	0.16	56.17	<0.0001	43.89	0.14	34.24	0.15	47.26	<0.0001
SED Sexual function	32.49	0.11	21.13	0.14	62.96	<0.0001	32.51	0.11	22.05	0.15	55.05	<0.0001	33.22	0.13	24.89	0.14	42.48	<0.0001
Study 2																		
5. 'Pleasurable'	3.77	0.02	1.84	0.02	70.50	<0.0001	3.76	0.02	1.87	0.02	70.07	<0.0001	3.89	0.02	2.43	0.02	48.18	<0.0001
6. 'Letting go'	3.63	0.02	1.94	0.03	49.12	<0.0001	3.64	0.02	1.95	0.03	50.35	<0.0001	3.79	0.02	2.40	0.03	40.95	<0.0001
7. 'In the mood for sex'	3.34	0.02	1.76	0.03	47.99	<0.0001	3.31	0.02	1.83	0.03	44.00	<0.0001	3.45	0.02	2.23	0.03	35.18	<0.0001
8. 'Vaginal arousal'	3.46	0.02	1.66	0.02	62.10	<0.0001	3.44	0.02	1.74	0.02	56.88	<0.0001	3.64	0.02	2.14	0.02	46.85	<0.0001
9. ' Sexually aroused'	3.53	0.02	1.68	0.02	64.78	<0.0001	3.51	0.02	1.75	0.02	59.50	<0.0001	3.67	0.02	2.22	0.02	44.83	<0.0001
10. 'Body image'	3.83	0.02	3.08	0.04	18.62	<0.0001	3.84	0.02	3.05	0.04	19.82	<0.0001	3.94	0.02	3.24	0.03	19.91	<0.0001
11. 'Afraid of pain'§	4.76	0.01	4.61	0.03	5.74	<0.0001	4.77	0.01	4.60	0.02	6.48	<0.0001	4.78	0.01	4.65	0.02	5.85	<0.0001
12. 'Disgust'§	4.75	0.01	4.30	0.03	14.54	<0.0001	4.76	0.01	4.28	0.03	15.37	<0.0001	4.75	0.01	4.47	0.02	11.02	<0.0001

13. 'Distracting thoughts'§	4.09	0.02	3.16	0.04	22.92	<0.0001	4.11	0.02	3.12	0.04	24.85	<0.0001	4.14	0.02	3.46	0.03	19.55	<0.0001
14. 'Partner attractiveness'	4.04	0.02	3.47	0.04	13.87	<0.0001	4.06	0.02	3.41	0.04	16.40	<0.0001	4.00	0.02	3.72	0.03	8.01	<0.0001
15. 'Partner's actions'	3.77	0.02	2.43	0.03	34.84	<0.0001	3.79	0.02	2.41	0.03	37.46	<0.0001	3.82	0.02	2.89	0.03	25.39	<0.0001
SED Total	43.69	0.13	30.01	0.17	63.76	<0.0001	43.72	0.13	30.12	0.16	64.54	<0.0001	44.86	0.15	33.86	0.17	48.12	<0.0001
SED Sexual function	33.47	0.12	21.03	0.15	63.54	<0.0001	33.48	0.13	21.15	0.15	64.19	<0.0001	34.34	0.15	24.74	0.16	44.58	<0.0001

*2-sided P-values. §Reversed variable. Abbreviation: SE= standard error of the mean, SED= Sexual Event Diary. Remarks: Study 1: Gratified= "Yes": n=3788 and "No": n=1479. Satisfied= "Yes": n=3649 and "No": n=1618. Orgasm= "Yes": n= 2790 (items 5 and 6 n=2789) and "No": n= 2479. Study 2: Gratified= "Yes": n=3237 and "No": n=1366. Satisfied= "Yes": n=3220 and "No": n=1383. Orgasm= "Yes": n= 1383 and "No": n=2415.

Subject level

Factor analysis

Inspection of the eigenvalues and scree-plots of the ML factor analyses and results of the PA revealed that one factor should be retained. The first factor had an eigenvalue of 5.41, explaining 49.1% of the variance for study 1 and an eigenvalue 5.28, explaining 48.0% of the variance for study 2. Items 5 to 9 and 15 had strong factor loadings, with less strong contributions of items 10, 13 and 14 (Table 2). Sexual function sum score was derived using the same items as for the event level analyses.

Reliability

The items gave a high Cronbach's alpha for both studies (see Table 2). Most inter-item Pearson correlation coefficients were over 0.30 (see *Supplementary, Tables C and D*). Only items 11 ('afraid pain'), 12 ('disgust') for both studies, and 13 ('distracting thoughts') for study 1 had correlations lower than 0.30. Item-rest Pearson correlation coefficients between 16-item SED items 4 ('satisfied'), 3 ('gratified') and 16 ('orgasms'), with answer options "yes"/ "no", and the sum scores were all larger than 0.30 in both studies 1 and 2. The item-rest Pearson correlation coefficients between the 16-item SED Likert items 5 to 15 and the sum scores were mostly larger than 0.30, except for item 11 ('afraid pain') for both studies.

Validity

Convergent validity was deemed adequate, with correlation coefficients between the SFQ domains with their related 16-item SED items ranging from 0.33 to 0.77 for study 1 and from 0.22 to 0.80 for study 2 during both periods (see *Supplementary, Tables E and F*). These results provided support for adequate to strong convergence of the 16-item SED. The lower, but adequate convergence of SFQ 'Orgasm' and SFQ34 'how satisfied have you been' was expected, since the SFQ items assessed intensity on these concepts over a 4-week period, while their related SED variables represented the different concepts frequency of satisfactory events and orgasms. Low correlations between SFQ 'Partner relation' and SED 'partner' items were also expected, since the 'partner' items were dissimilar. The SFQ assesses the fear of negative impact of sexual dysfunction on the relationship, while the SED assesses the partner's proficiency and attractiveness during the sexual event.

Responders, as defined by the SEG, had significantly higher 16-item SED item scores in ATP compared to non-responders on almost all items (see *Supplementary, Table G*). SED item

scores for 10 (‘body image’) for study 1 and 11 (‘afraid of pain’) and 12 (‘disgust’) for both studies did not differ significantly between responders and non-responders.

Responsiveness

Responders had a significantly higher increase in change from BLE to ATP in 16-item SED scores compared to non-responders (Table 4). Only item 12 (‘disgust’) for both studies, and 11 (‘afraid of pain’) for study 1, did not differ significantly between responders and non-responders. Guyatt's effect size ranges were 0.80-1.23 for study 1 and 0.59-1.29 for study 2 indicating moderate to strong ability to detect changes in 16-item SED item scores (Table 5**)**. Exceptions were items 11 (‘afraid of pain’), 12 (‘disgust’), and 13 (‘distracting thoughts’) for both studies and 14 (‘partner attractive’) for study 1, since their effect sizes were small (0.07-0.44 and 0.05-0.41 for studies 1 and 2, respectively).

Table 5. Known Groups Responsiveness – Mean (SD) Change in 16-item SED scores from Baseline establishment to Active Treatment Period in responders and non-responders as defined by the SEG

	Study 1							Study 2						
	Responder [n=72^]		Non-responder [n=76]		T-value	p-value*	Guyatt's responsiveness	Responder [n=76]		Non-responder [n=48]		t-value	p-value*	Guyatt's responsiveness
	Mean change [SE]		Mean change [SE]					Mean change [SE]		Mean change [SE]				
SED3 'Gratified'†	2.48	[0.40]	0.42	[0.27]	4.25	<0.0001	0.87	2.14	[0.32]	0.73	[0.34]	2.89	0.005	0.59
SED4: 'Satisfied'†	2.56	[0.40]	0.32	[0.27]	4.65	<0.0001	0.95	2.08	[0.33]	0.75	[0.32]	2.92	0.004	0.60
SED5: 'Pleasurable'	1.12	[0.12]	0.19	[0.11]	5.83	<0.0001	0.97	0.97	[0.11]	0.24	[0.16]	3.82	0.0002	0.65
SED6: 'Letting go'	1.23	[0.11]	0.23	[0.11]	6.44	<0.0001	1.08	1.16	[0.12]	0.13	[0.11]	6.10	<0.0001	1.29
SED7: 'In the mood for sex'	1.33	[0.12]	0.36	[0.11]	6.10	<0.0001	1.04	1.10	[0.13]	0.25	[0.15]	4.22	<0.0001	0.84
SED8: 'Vaginal arousal'	1.17	[0.12]	0.15	[0.11]	6.22	<0.0001	1.02	1.11	[0.13]	0.33	[0.15]	3.91	0.0002	0.75
SED9: ' Sexually aroused'	1.25	[0.13]	0.34	[0.11]	5.28	<0.0001	0.96	1.17	[0.12]	0.23	[0.15]	4.90	<0.0001	0.89
SED10: 'Body image'	0.75	[0.12]	0.08	[0.09]	4.62	<0.0001	0.88	0.88	[0.14]	0.25	[0.13]	3.38	0.001	0.72
SED11: 'Afraid pain'§	0.19	[0.06]	0.13	[0.08]	0.50	0.616	0.07	0.29	[0.09]	-0.03	[0.12]	2.20	0.030	0.38
SED12: 'Disgust'§	0.22	[0.06]	0.14	[0.06]	0.94	0.350	0.15	0.32	[0.08]	0.29	[0.10]	0.26	0.793	0.05
SED13: 'Distracting thoughts'§	0.46	[0.12]	0.06	[0.11]	2.47	0.015	0.44	0.82	[0.11]	0.39	[0.15]	2.32	0.022	0.41
SED14: 'Partner attractiveness'	0.24	[0.12]	-0.11	[0.11]	2.23	0.027	0.38	0.53	[0.11]	0.02	[0.12]	3.01	0.003	0.60
SED15: 'Partner's actions'	0.86	[0.13]	0.10	[0.11]	4.51	<0.0001	1.23	0.82	[0.13]	0.20	[0.14]	3.16	0.002	0.62
SED16: 'Orgasms'†	2.04	[0.28]	-0.11	[0.20]	6.28	<0.0001	0.80	1.59	[0.27]	0.40	[0.26]	3.17	0.002	0.66
SED Total	8.83	[0.86]	1.69	[0.70]	6.45	<0.0001	1.17	9.17	[0.84]	2.28	[1.00]	5.22	<0.0001	0.99
SED Sexual function	8.43	[0.82]	1.41	[0.67]	6.61	<0.0001	1.20	8.56	[0.82]	2.03	[0.91]	5.17	<0.0001	1.03

*2-sided tests were used. ^Study 1 items 3 & 4: n=73. §Reversed variable. †Counts of "yes" answers on these items were used according to efficacy analyses. Abbreviations: SE= standard error of the mean, SED= Sexual Event Diary, SEG = Subjective Evaluation of Gain.

11-item SED (study 3)

Cognitive debriefing interviews and item reduction

Content validity of the SED was assessed by conducting two iterative sets of cognitive debriefing interviews. The 16-item SED was tested in round 1, and following adaptation, the 11-item SED was tested in round 2. Each round included eight women (Table 1) who met the in- and exclusion criteria. Six of the interview participants were classified as having low sensitivity to sexual cues and two as demonstrating dysfunctional over-activity of sexual inhibitory mechanisms. This subdivision is based on the dual control model of sexual response and is substantiated by cognitive [15, 16], psychophysiological [6,7,15-17], subjective [6,7,16], neuroanatomical [18,19] and pharmacological [6,7,15,16,18] evidence. This information was collected to ascertain if these groups differed in opinions/perceptions regarding what is important for satisfactory sex. No differences were observed. Based on the results of these interviews in round 1, revisions were made to the item set, including the removal of items addressing concepts less important to patients and the development of new items to capture concepts of greater importance to patients. SED item wording of the two versions, including a brief description of the type of change made is summarized in *Supplementary, Table H*. In round 1, all patients found the items 4 (‘satisfied’) and 16 (‘orgasm’) and the majority of patients found the items 5 (‘pleasure’), 6 (‘letting go’) and 13 (‘distracting thoughts’) clear, easy to understand and answer, and relevant for a sexual event diary. Therefore, these items were retained. Items 4 (‘satisfied’), 6 (‘letting go’), and 13 (‘distracting thoughts’) were not modified, and for the items 5 (‘pleasure’) and 16 (‘orgasm’) minor modifications to wording were made, e.g., refinements to US English. The majority of the patients indicated that the ‘gratified’ item was measuring the same as the ‘satisfaction’ item; half of the patients preferred the item assessing ‘satisfaction’, and several said the term ‘gratification’ was not clear (38%). As a consequence, the ‘gratified’ item was deleted without replacement. Item 9 (‘sexually aroused’) was deleted and replaced since participants not always interpreted the question as asking for mental arousal. The item 7 (‘in the mood for sex’) was deleted and replaced, since participants found this item was not fully measuring desire. Several patients indicated that overall physical arousal was also important to assess instead of only vaginal arousal, as a result the item 8 (‘vaginal arousal’) was deleted and replaced. Items 10 (‘body image’), 11 (‘afraid of pain’), 12 (‘disgust’) and 14 (‘partner attractiveness’) and 15 (‘partner’s actions’) were deemed irrelevant to a sexual event diary by interview participants. These items were therefore deleted without replacement. The deletion of item 11 (‘afraid of pain’) and item 12 (‘disgust’) was also justified by evaluating the quantitative assessment results. Both items had

overall the lowest factor loadings. The 11-item SED was then tested in round 2, which resulted in strong content validity. Nearly all participants found each item included in the SED clear, easy to answer and important to capture in a sexual event diary, so no changes were made to the final SED. An exception was the deletion of the "How would you rate" stem of SED items 6-9, since this item formulation generally did not perform well. The wording was changed to a more direct form, e.g., "How physically aroused or excited did you become during the sexual activity?". The psychometric data gathered in studies 1 and 2 supported the item selection for the SED.

Event level

Factor analysis

Inspection of the eigenvalues and scree-plots of the ML factor analyses and results of the PA revealed that one factor should be retained. The factor had an eigenvalue of 4.89, explaining 81.4% of the variance. The one-factor structure showed high loadings for all SED items (Table 2). Because of this finding, the SED total sum score and the SED sexual function sum score were equal and consisted of all Likert scale items.

Reliability

Cronbach's alpha coefficient was high (Table 3). The Pearson correlation coefficients, which were calculated for assessing the SED inter-item and item-rest correlations, were all larger than 0.30 (P<0.0001, see *Supplementary, Table I*).

Validity

All SED items showed strong construct validity. The mean differences in SED scores between "yes" and "no" responders on SED items measuring 'satisfied' and 'orgasm' were highly significant (P<0.0001) and the results were in the expected direction (Table 6).

Table 6. Mean (SE) in 11-item SED scores separately for answering SED item 4 "Were you satisfied with the sexual activity?" and 11 "Did you have an orgasm?" with "Yes" or "No" – study 3

	4. Were you satisfied with the sexual activity?						11. Did you have an orgasm?					
	Yes		No		Test statistics		Yes		No		Test statistics	
SED item	mean	SE	Mean	SE	t-value	p-value*	mean	SE	Mean	SE	t-value	p-value*
5. How would you rate your level of sexual desire during the sexual activity?	2.61	0.04	0.83	0.05	30.74	<0.0001	2.63	0.04	1.22	0.05	21.74	<0.0001
6. How mentally aroused or excited did you become during the sexual activity?	2.56	0.04	0.78	0.05	29.73	<0.0001	2.60	0.05	1.14	0.05	22.37	<0.0001
7. How physically aroused or excited did you become during the sexual activity?	2.72	0.04	0.92	0.05	31.63	<0.0001	2.80	0.04	1.25	0.05	25.56	<0.0001
8. To what extent did you have distracting thoughts? §	2.62	0.04	1.47	0.07	14.81	<0.0001	2.65	0.05	1.70	0.06	12.96	<0.0001
9. To what extent were you able to let yourself go?	2.66	0.04	0.83	0.05	27.50	<0.0001	2.79	0.05	1.13	0.05	24.49	<0.0001
10. How pleasurable was the sexual activity to you?	2.82	0.04	0.87	0.05	33.81	<0.0001	2.89	0.04	1.25	0.05	26.01	<0.0001
SED Sexual function	15.99	4.71	5.70	5.15	32.85	<0.0001	16.36	0.22	7.69	0.26	25.79	<0.0001

*2-sided P-values. §Reversed variable.
Abbreviation: SE= standard error of the mean.
Remarks: Satisfied ="Yes": n= 650 and "No": n=414 . Orgasm= "Yes": n=527 and "No": n=537.

Subject level

Factor analysis

Inspection of the eigenvalues and scree-plots of the ML factor analyses and results of the PA revealed that one factor should be retained. The factor had an eigenvalue of 4.66, explaining 77.7% of the variance. The one-factor structure showed high loadings for all items (Table 2). Also here, the SED total sum score was equal to the SED sexual function sum score and consisted of all Likert scale items.

Reliability

Cronbach's alpha coefficient was high (Table 3). The majority of the Pearson correlation coefficients, which were calculated for assessing the SED inter-item and item-rest correlations, were larger than 0.30 (P<0.0001), see *Supplementary, Table J*.

Validity

Convergent validity was deemed adequate, with correlation coefficients between the FSFI domains with their related SED items ranged from 0.36 to 0.79 during both periods (see *Supplementary, Table K*). These results provided support for adequate to strong convergence of the 11-item SED. The adequate convergence between FSFI item 16 ('how satisfied have you been') and SED item 4 ('satisfied') was expected, since FSFI item 16 measured intensity of satisfaction over a 4-week period, while the related SED item 4 ('satisfied') measured the different concept frequency of satisfactory events.

Known groups validity was good. Responders scored significantly higher compared to non-responders (P<0.05, see *Supplementary, Table L*) on all items during the ATP, except on items 8 ('distracting thoughts'), 9 ('letting go') and 11 ('orgasm').

Responsiveness

For responders increase in SED item scores from BLE to ATP was significantly higher than for non-responders (P<0.05, see Table 7) with exception of items 8 ('distracting thoughts') and 11 ('orgasm') showing strong known groups responsiveness.

The Guyatt's effect sizes ranged from 0.73-1.58 indicating a very good ability to detect changes in SED item scores, see Table 7. An exception was item 8 ('distracting thoughts') which had a small effect size (0.14).

Table 7. Known Groups Responsiveness – Mean (SD) Change in SED scores from Baseline establishment to Active Treatment Period in responders and non-responders as defined by the SEG – study 3

	Responder [n=18]		Non-responder [n=10]		t-value	p-value*	Guyatt's responsiveness
	Mean change	SE	Mean change	SE			
4. Were you satisfied with the sexual activity?†	1.17	0.52	-1.40	0.83	2.76	0.011	0.97
5. How would you rate your level of sexual desire during the sexual activity?	1.06	0.21	0.06	0.26	2.92	0.007	1.22
6. How mentally aroused or excited did you become during the sexual activity?	1.05	0.22	-0.37	0.28	3.89	0.0006	1.58
7. How physically aroused or excited did you become during the sexual activity?	0.94	0.20	-0.21	0.27	3.41	0.002	1.34
8. To what extent did you have distracting thoughts?§	0.56	0.26	0.43	0.31	0.32	0.751	0.14
9. To what extent were you able to let yourself go?	0.73	0.18	-0.07	0.35	2.29	0.031	0.73
10. How pleasurable was the sexual activity to you?	0.91	0.22	-0.44	0.34	3.50	0.002	1.25
11. Did you have an orgasm?†	0.83	0.48	-0.60	0.56	1.86	0.074	0.81
SED Sexual function	5.25	0.98	-0.60	1.30	3.59	0.001	1.42

*2-sided P-values. §Reversed variable. †Counts of "yes" answers on these items were used according to efficacy analyses. Abbreviations: SE= standard error of the mean, SED= Sexual Event Diary, SEG = Subjective Evaluation of Gain.

Discussion

A standardized event log, the SED, was developed for assessment of sexual satisfaction and sexual functioning of a single sexual event. The questions in the SED are directed at a discrete sexual event instead of being directed at sexual functioning over a longer period of time, e.g., four weeks. Measuring discrete sexual events gives a more valid assessment of efficacy of on-demand investigational drugs on sexual functioning of women with HSDD/FSIAD compared to questionnaires that are directed at assessment of sexual function over longer periods of time, e.g., SFQ, FSFI. This is because the influence of on-demand medication is predominantly present during an event and the data from event logs are therefore more proximate estimations of an on-demand drug's efficacy than data from monthly questionnaires, giving the advantage of minimized recall bias and increased precision. The results of the present study show that the SED is a valid and reliable measure for assessing female sexual function during discrete sexual events.

Reliability, validity and responsiveness were confirmed for the majority of the 11-item SED items, based on evidence from cognitive debriefing interviews and psychometrical assessments in patients with HSDD. These findings indicated that most items measured the same concept of interest, the construct we intended to measure and changes in sexual functioning when change was reported. Our clinical trials' primary endpoint measure 'change in the number of satisfying sexual events from baseline', proved to be an excellent measure. This measure was comprehensive and correlated strongly with, and has excellent discriminating ability in, all aspects of sexual functioning. Furthermore, the SED showed to have a clear one-factor structure, indicating that the resulting scale (sum score) measured the same concept, and this scale showed excellent reliability, validity and responsiveness.

Discrepancies were found in our results regarding 11-item SED items 8 ('distracting thoughts'), 9 ('letting go '), and 11 ('orgasm'). These appeared to be less valid and/or responsive according to psychometric assessments, whereas debriefing interviews revealed excellent content validity of all items included in the SED, indicating that they are appropriate and comprehensive relative to its intended measurement concept, population, and use. For item 11 ('orgasm') a lower validity than the other SED items was expected, because sexual satisfaction in women is less dependent on reaching orgasm than in men [20]. Nevertheless, for a substantial number

of women, orgasm is an important aspect of sexual functioning [20], and should therefore be assessed in clinical trials that investigate the efficacy of drugs for HSDD/FSIAD. The main reason 8 (‘distracting thoughts’) and 9 (‘letting go’) were included in the 11-item SED was to capture information on inhibition for the subtype of women in which HSDD/FSIAD is caused by dysfunctional over-activity of sexual inhibitory mechanisms. These women’s sexual excitation is hampered by overactivity of normal inhibitory processes in the brain [18], which may be expressed behaviorally as excessive distraction or an inability to let oneself go during sexual activity.

A limitation of the present study is that the final focus groups/cognitive interviews were performed in predominantly postmenopausal women, and predominantly women with low sensitivity to sexual cues. It is not expected that additional research in premenopausal and high inhibitory subjects would necessitate additional items or adjustment of items to reach concept saturation. There is no literature that suggest that there is a difference between women of these subgroups (pre- vs. postmenopausal, and low sensitive vs. high inhibitory) with respect to which aspects of a sexual event are important. This will need to be confirmed in future research.

The SED was developed and validated as a part of a drug development program for HSDD/FSIAD. Development of and modifications to the SED were based on the premise that the instrument had to be a valid and reliable tool for use in such a program. The data that were used for the here described validation were also gathered in this program. Despite of this focus, however, the SED may also show merit in the assessment of sexual functioning of discrete sexual events in other areas of research, e.g., recreational drug use and sexual risky behaviors.

In conclusion, the 11-item SED has proven to be an excellent tool for measuring sexual satisfaction and sexual functioning over a single sexual event, and is therefore suitable for use in clinical trials assessing the efficacy of on-demand drugs in women with HSDD/FSIAD.

Acknowledgments

XXX

References

1 Davison SL, Bell RJ, La China M, Holden SL, Davis SR. Assessing sexual function in well women: Validity and reliability of the Monash women's health program female sexual satisfaction questionnaire. J Sex Med 2008;5:2575-86.

2 American Psychiatric Association. DSM-IV: Diagnostic and statistical manual of mental disorders. 4th edition. Text revision. Washington, DC: American Psychiatric Press; 2000.

3 American Psychiatric Association. DSM-5: Diagnostic and Statistical Manual of Mental Disorders. 5th edition. Arlington, VA: American Psychiatric Publishing; 2013.

4 Joffe HV, Chang C, Sewell C, Easley O, Nguyen C, Dunn S, Lehrfeld K, Lee L, Kim M, Slagle AF, Beitz J. FDA Approval of Flibanserin-Treating Hypoactive Sexual Desire Disorder. N Engl J Med 2016;374:101-4.

5 Tuiten A, van Rooij K, Bloemers J, Eisenegger C, van Honk J, Kessels R, Kingsberg S, Derogatis LR, de Leede L, Gerritsen J, Koppeschaar HPF, Olivier B, Everaerd WTAM, Frijlink HW, Höhle D, de Lange RPJ, Böcker KBE, Pfaus JG. Personalized medicines for female sexual interest/arousal disorder. In submission.

6 Poels S, Bloemers J, van Rooij K, Goldstein I, Gerritsen J, van Ham D, van Mameren F, Chivers M, Everaerd W, Koppeschaar H, Olivier B, Tuiten A. Toward personalized sexual medicine (part 2): testosterone combined with a PDE5 inhibitor increases sexual satisfaction in women with HSDD and FSAD, and a low sensitive system for sexual cues. J Sex Med 2013;10:810–23.

7 van Rooij K, Poels S, Bloemers J, Goldstein I, Gerritsen J, van Ham D, van Mameren F, Chivers M, Everaerd W, Koppeschaar H, Olivier B, Tuiten A. Toward personalized sexual medicine (part 3): testosterone combined with a Serotonin1A receptor agonist increases sexual satisfaction in women with HSDD and FSAD, and dysfunctional activation of sexual inhibitory mechanisms. J Sex Med 2013;10:824–37.

8 Clayton AH, Althof SE, Kingsberg S, DeRogatis LR, Kroll R, Goldstein I, Kaminetsky J, Spana C, Lucas J, Jordan R, Portman DJ. Bremelanotide for female sexual dysfunctions in premenopausal women: a randomized, placebo-controlled dose-finding trial. Women's Health 2016;12:325–37.

9 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.

10 U.S. Department of Health and Human Services. Low Sexual Interest, Desire, and/or Arousal in Women: Developing Drugs for Treatment. Guidance for Industry (Draft Guidance); 2016.

11 Timmerman ME, and Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychol Methods 2011;16: 209-20.

12 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.

13 Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

14 Guyatt GH, Walter SD, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171-8.

15 van der Made F, Bloemers J, van Ham D, Yassem WE, Kleiverda G, Everaerd W, Olivier B, Tuiten A. Childhood sexual abuse, selective attention for sexual cues and the effects of testosterone with or without vardenafil on physiological sexual arousal in women with sexual dysfunction: a pilot study. J Sex Med 2009a;6:429–39.

16 van der Made F, Bloemers J, Yassem WE, Kleiverda G, Everaerd W, van Ham D, Olivier B, Koppeschaar H, Tuiten A. The influence of testosterone combined with a PDE5-inhibitor on cognitive, affective, and physiological sexual functioning in women suffering from sexual dysfunction. J Sex Med 2009b;6:777–90.

17 Gerritsen G, van der Made F, Bloemers J, van Ham D, Kleiverda G, Everaerd W, Olivier R, Levin R, Tuiten A. The clitoral photoplethysmograph: A new way of assessing genital arousal in women. J Sex Med 2009;6:1678–87.

18 Bloemers J, van Rooij K, Poels P, Goldstein I, Everaerd W, Koppeschaar H, Chivers M, Gerritsen J, van Ham D, Olivier B, Tuiten A. Toward personalized sexual medicine (Part 1): Integrating the "dual control model" into differential drug treatments for hypoactive sexual desire disorder and female sexual arousal disorder. J Sex Med 2013;10:791–809.

19 Bloemers J, Scholte HS, van Rooij K, Goldstein I, Gerritsen J, Olivier B, Tuiten A. Reduced gray matter volume and increased white matter fractional anisotropy in women with hypoactive sexual desire disorder. J Sex Med 2014;11:753–67.

20 Salisbury CMA, Fisher WA. ‘‘Did You Come?’’ A Qualitative Exploration of Gender Differences in Beliefs, Experiences, and Concerns Regarding Female Orgasm Occurrence During Heterosexual Sexual Interactions. J Sex Res 2014;51:616–31.

Supporting Information

Additional Supporting information may be found in the online version of this article at the publisher's web-site:

- Supplementary Tables:
  - Table A. Inter-item correlations of the 16-item SED items, at event level – study 1
  - Table B. Inter-item correlations of the 16-item SED items, at event level – study 2
  - Table C. Inter-item correlations of the 16-item SED items, at subject level – study 1
  - Table D. Inter-item correlations of the 16-item SED items, at subject level – study 2
  - Table E. Convergent validity. Pearson correlations: 16-item SED items with the SFQ domains and item 34 – study 1
  - Table F. Convergent validity. Pearson correlations: 16-item SED items with the SFQ domains and item 34 – study 2
  - Table G. Known Groups validity – Mean (SD) in 16-item SED scores in the Active Treatment Period in responders and non-responders as defined by the SEG – studies 1 and 2
  - Table H. Summary item tracking matrix SED and SED: Content validation based on patient debriefing interviews
  - Table I. Inter-item correlations of the 11-item SED items, at event level – study 3
  - Table J. Inter-item correlations of the 11-item SED items, at subject level – study 3
  - Table K. Convergent Validity. Pearson correlations: 11-item SED items with the FSFI domains, and FSFI item 16 – study 3
  - Table L. Known Groups validity – Mean (SD) in 11-item SED scores in the Active Treatment Period in responders and non-responders as defined by the SEG – study 3

Appendix

Sexual Event Diary

The questions in this diary are designed to gather information about your sexual activity during the study. Please complete this diary as soon as possible every time you engage in sexual activity.

For the purposes of this diary, "sexual activity" includes any activity which may result in sexual stimulation or sexual pleasure. Such activities include but are not necessarily limited to sexual intercourse (vaginal/anal), oral sex, genital stimulation with a partner, and masturbation (self-stimulation).

For each item, please select the answer that best describes your last sexual experience (the one you are now reporting).

Q1. When did this sexual activity occur?
[ ] Less than 8 hours ago
[ ] At least 8 but less than 24 hours ago
[ ] At least 24 but less than 48 hours ago
[ ] More than 48 hours ago (2 days)

Q2. In what kind(s) of sexual activity did you engage? Select all that apply.
[ ] Masturbation (self-stimulation)
[ ] Sexual intercourse
[ ] Oral sex performed on your partner
[ ] Oral sex performed on you
[ ] Stimulation of your partner's genitals (for example, using your hands or sexual aids)
[ ] Stimulation of your genitals (for example, using your or your partner's hands or sexual aids)
[ ] Other, please specify _____

Q3. Did you use the study medication prior to the sexual activity?
[ ] Yes
[ ] No

Q4. Were you satisfied with the sexual activity?
[ ] Yes
[ ] No

Q5. How would you rate your level of sexual desire during the sexual activity?

0	1	2	3	4
No desire				Extreme desire

Q6. How mentally aroused or excited did you become during the sexual activity?

0	1	2	3	4
Not aroused at all				Extremely aroused

Q7. How physically aroused or excited did you become during the sexual activity?

0	1	2	3	4
Not aroused at all				Extremely aroused

Q8. To what extent did you have distracting thoughts?

0	1	2	3	4
Not at all				Totally

Q9. To what extent were you able to let yourself go?

0	1	2	3	4
Not at all				Totally

Q10. How pleasurable was the sexual activity to you?

0	1	2	3	4
Not pleasurable at all				Extremely pleasurable

Q11. Did you have an orgasm?
[ ] Yes
[ ] No