FEMALE SEXUAL FUNCTION

Pooled Analysis Confirms Flibanserin's Unimpressive Efficacy, Raises Measurement Questions: A Commentary on Simon et al

Flibanserin's efficacy for hypoactive sexual desire disorder (HSDD) has been debated for over a decade. In the latest contribution to this topic, Simon et al. reported statistically significant benefits for flibanserin over placebo on Female Sexual Function Index (FSFI) domains including desire, arousal, lubrication, orgasm, pain (premenopausal women only), and satisfaction.[1] Results were reported separately for premenopausal women (3 trials, n = 2,368) and postmenopausal women (one trial, n= 895). The large sample size is a strength, allowing some precision in estimating treatment effects.

Flibanserin's efficacy is likely not clinically meaningful. For continuous measures (instruments reporting mean change, rather than categories such as responder and/or non-responder), effect sizes of less than d = .20 are typically considered less than small.[2] For FSFI Pain, effect sizes were nil for postmenopausal women and very small for premenopausal women. For postmenopausal women, on other FSFI domains, 4 of 5 statistically significant effect sizes were less than small, with only the effect on the desire domain (FSFI-D) entering the small range. If one includes data from an additional trial on postmenopausal women which was discontinued early, effects are reduced further. On a related note, flibanserin does not have regulatory approval for treating HSDD in postmenopausal women. For premenopausal women, effects (adjusted for covariates) were small on desire, arousal, lubrication and satisfaction (d = .20 to .28) and very small for orgasm and pain.

In examining whether flibanserin generates clinically meaningful benefits, the authors isolated the FSFI domain with the highest effect size (FSFI-D), stating that effects on the FSFI-D are comparable to the effects of antidepressant drugs for depression (d = .32) and medications for anxiety and obsessive-compulsive disorder. This comparison discards the five other FSFI domains. Further, many have questioned the clinical significance of antidepressant treatment for depression.[3] The effect sizes of flibanserin on the FSFI-D (d = .28 and .26 for premenopausal women and postmenopausal women, respectively) are lower than the effect sizes the authors provided for antianxiety medication (d = .40-.41) or medication for obsessive-compulsive disorder (d = .44). Even if effect sizes were equivalent across conditions, cautions abound. Given that different participants with different levels of impairment completed measures of quite different symptoms, an effect size of, say, .30 for a HSDD treatment may not mean the same thing as an effect size of .30 for a depression treatment.

Simon et al. noted that another recent paper found higher rates of "clinically meaningful improvement" on the Patient Global Impression of Improvement (PGI) scale for those taking flibanserin vs. placebo.[4] However, their definition of "clinically meaningful" improvement was at least a "minimally improved" PGI rating. Nobody seeks treatment aiming for minimal improvement. A more stringent approach might link "clinically meaningful" improvement to PGI ratings of "much improved" or "very much improved"; however, no statistical analysis was provided using this more rigorous definition of treatment response. It is clear that "much improved" or "very much improved" scores were substantially less common than "minimally improved".[4]

Simon et al. also provided the number needed to treat (NNT) to generate a FSFI-D treatment response that would not have been obtained if all participants had received placebo. NNT's for flibanserin (4.7 for premenopausal women, 9.4 for postmenopausal women) were compared favorably to NNT values for antidepressant therapies (NNTs of 7-9). This compares apples to oranges: Meeting an arbitrary cutoff for treatment response for one condition does not logically equate to meeting an arbitrary cutoff for treatment response for another condition. Further, they provide no evidence that their definition of treatment response is meaningful.

Additionally, their NNT calculation for premenopausal women shows how changing the participants included for analysis changes results. For the full efficacy dataset, using response rates on the FSFI-D[4] (47.1% vs 35% for premenopausal and 38.9% vs 26.3% for postmenopausal women), NNT values are 8.26 for premenopausal women (rounds to 9; NNT rounds to next highest integer) and 7.94 (rounds to 8) for postmenopausal women. The current authors generated a NNT of 4.7 by including only women who completed the study rather than the full analysis set; such results do not generalize to those who did not complete a study.

There is no gold standard measure for assessing HSDD treatment outcomes. Thus, Simon et al.'s use of outcome measures across differing domains of sexual functioning has potential merit. However, FSFI items were originally developed for women with female sexual arousal disorder, not HSDD. The only study examining the FSFI-D's content validity for HSDD found that 33 of 75 (44%) women with HSDD said it did not entirely capture their sexual desire and/or interest problems[5]; this is not impressive. The validity of the other FSFI domains has not been well-studied in women with HSDD. Thus, it is unclear what numerical treatment benefits on FSFI domains actually mean. Unfortunately, measures of questionable validity plague the HSDD treatment literature.[6]

Scores on all FSFI domains except desire are invalid for women who are sexually inactive.[7] All women in the flibanserin studies agreed to be sexually active monthly. Yet given their

HSDD diagnosis, some participants almost certainly refrained from sexual activity. The inclusion of their invalid scores renders study results inaccurate to whatever extent participants were sexually inactive during one or more of the five 28-day FSFI measurement periods during a study.

The authors aptly note that their pooled analysis did not include measures of relational satisfaction or overall well-being. Across clinical trials in general, too much focus is placed on symptomatic measures and not enough placed on how treatments impact a broader scope of outcomes that matter to patients. If treatment improves HSDD symptoms by a small amount compared to placebo, but relational satisfaction and overall well-being show no change, is this a positive outcome? If a patient has "minimally improved", is that a marker for success? Assessing a broader scope of outcomes is crucial in better understanding how treatments truly impact people's lives.

Careful assessment is needed to avoid pathologizing women who simply want to have less sex than their partner. Among women who are appropriate treatment candidates, it is difficult to quantify either the extent of their problems or their improvement without valid, broad-based assessment. Better measurement tools are needed. Treatment benefits must be considered in the context of treatment risks. Women are strongly advised to not take flibanserin within two hours of drinking alcohol (or skip a dose if they've had 3 or more drinks) due to increased risk of severe hypotension and syncope.[8] Given the substantial percentage of women who consume alcohol at least occasionally, this is problematic. Further, the difference in reported rates of somnolence and/or fatigue and/or sedation of 21% vs 8% for flibanserin vs placebo is notable.[8] Relative to placebo, about the same percentage of women (12% or 13%) experience a FSFI-D treatment "response" (see caveats above) or report somnolence and/or fatigue and/or sedation.

Reporting data from trials completed over a decade ago adds to our understanding but the delay in reporting such results is disappointing. Simon et al.'s insights into flibanserin's effects across FSFI domains confirm the drug's underwhelming efficacy. For women who personally struggle with low desire, better treatments are needed, and a firmer measurement foundation is necessary to better understand both low desire and the effects of treatments.

Glen I. Spielmans, PhD
Metropolitan State University, Department of Psychology, Saint Paul, MN, USA

Corresponding Author: Glen I. Spielmans, PhD, Metropolitan State University, Department of Psychology, Saint Paul, MN, USA. Tel: 651-999-5826; Fax: 651.999.5822; E-mail: glen.spielmans@metrostate.edu

Conflict of Interest: Dr. Spielmans holds shares in Vanguard Healthcare, a mutual fund investing heavily in pharmaceutical firms.

Funding: None.

https://doi.org/10.1016/j.esxm.2022.100579

REFERENCES

1. Simon JA, Clayton AH, Goldstein I, et al. Effects of flibanserin on subdomain scores of the Female Sexual Function Index in women with hypoactive sexual desire disorder. **Sex Med 2002;10:100570.**
2. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition Erlbaum; 1988.
3. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. **PLoS Med 2008;5:e45.**
4. Simon JA, Clayton AH, Kim NN, et al. Clinically meaningful benefit in women with hypoactive sexual desire disorder treated with flibanserin. **Sex Med 2022;10:100476.**
5. Revicki DA, Margolis MK, Bush EN, et al. Content validity of the Female Sexual Function Index (FSFI) in pre-and postmenopausal women with hypoactive sexual desire disorder. **J Sex Med 2011;8:2237–2245. doi: 10.1111/j.1743-6109.2011.02312.x.**
6. Lim-Watson MZ, Hays RD, Kingsberg S, et al. A systematic literature review of health-related quality of life measures for women with hypoactive sexual desire disorder and female sexual interest/arousal disorder. **Sex Med Rev 2022;10:23–41. doi: 10.1016/j.sxmr.2021.07.003.**
7. Meyer-Bahlburg HF, Dolezal C. The female sexual function index: A methodological critique and suggestions for improvement. **J Sex Marital Ther 2007;33:217–224. doi: 10.1080/00926230701267852.**
8. Food and Drug Administration. Highlights of prescribing information, flibanserin (Addyi).. Published online; 2019. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/022526s009lbl.pdf.