



## Irrelevancies and Inaccuracies Miss the Target: A Response to Kingsberg et Al.

Glen I. Spielmans

Psychology, Metropolitan State University

I was disappointed by the response of Kingsberg et al. (2021), who failed to seriously engage with the most pressing concerns raised in my re-analysis (Spielmans, 2021). Kingsberg et al. (2021)'s commentary included irrelevant statements, inadequate or inaccurate arguments, unsubstantiated concerns regarding my level of knowledge, and mischaracterizations of my re-analysis. In this response, I closely examine the few instances in which they responded to concerns raised in my re-analysis, then refute various other statements in their commentary.

### Kingsberg et al.'s Concerns and Rebuttals to These Concerns

Table 1 shows the main concerns originally listed in table 1 of my re-analysis (Spielmans, 2021) alongside Kingsberg et al.'s (2021) response (if any), followed by summaries of my rebuttals to Kingsberg et al.'s responses. Kingsberg et al. responded, more or less directly, to four of my 10 main concerns. I thus assume they have no counterargument to six of my ten concerns.

One of my main concerns was regarding Kingsberg et al. (2019) not reporting 8 of the 11 protocol-specified efficacy outcomes. In response, Kingsberg et al. claimed that the journal in which they published (*Obstetrics & Gynecology*) lacked space to report all protocol-specified outcomes. *Obstetrics & Gynecology* allows up to 5500 words for original research articles, along with supplemental content that does not count toward the word count (Obstetrics & Gynecology, 2020). Further, Kingsberg et al. (2019) somehow found space to report 15 non-protocol specified efficacy outcomes, so there should also have been room to report the protocol-specified *a priori* analyses. Good Publication Practice (GPP3) standards (Battisti et al., 2015), to which Kingsberg et al. (2019) claimed to adhere, require following CONSORT data reporting standards – and CONSORT requires reporting of all protocol-specified measures (Schulz et al., 2010).

I stated that many efficacy outcomes reported by Kingsberg et al. (2019) were apparently post-hoc, due to their exclusion from the clinicaltrials.gov study protocols' lists of prespecified outcomes. Kingsberg et al. (2021) stated that I was incorrect about this because their statistical analysis plan was reviewed by the FDA prior to database lock. I believe Kingsberg et al. are conflating a) data analyses done for the sake of FDA approval with b) data analyses done for the sake of their *Obstetrics & Gynecology* manuscript. Outside of a brief mention that the

General Assessment Questionnaire has not been validated, I see no description of the categorical efficacy measures reported by Kingsberg et al. (2019) in the bremelanotide New Drug Application (United States Food and Drug Administration, 2019). Also, if these were in fact *a priori* measures, they should have been listed as such in the clinicaltrials.gov study protocol entries. In the absence of any evidence to the contrary, I stand by my claim that the vast majority of efficacy outcomes reported by Kingsberg et al. (2019) seem to have been derived post-hoc.

Kingsberg et al. devoted a few paragraphs to justifying changing the number of satisfactory sexual events to a secondary outcome and stated that I demonstrated a lack of expertise by criticizing this change. However, my re-analysis did not criticize the change of outcomes. Rather, I critiqued the lack of transparency surrounding this change. Switching satisfactory sexual events to a secondary outcome was not reported in the journal article, which again violates CONSORT standards.

In their commentary, Kingsberg et al. stated that I am “unaware of the validation that was conducted, at the direction of the FDA, to establish clinically meaningful cut-offs of the various patient reported outcomes to define clinical benefit.” Kingsberg et al. cited Revicki et al. (2018) to support this statement. It is unclear whether Kingsberg et al. are claiming that *all* of the patient-reported outcomes were validated in such a manner. In any case, readers can rest assured that I am quite aware of Revicki et al. (2018). In fact, I cited and discussed it in my re-analysis (Spielmans, 2021). Revicki et al. (2018) derived cutoff scores for only the FSFI-D and FSDS-DAO #13, not the bevy of other categorical outcomes reported by Kingsberg et al. (2019). I noted that Revicki et al.'s (2018) treatment response criteria categorized many participants as “responders” on the FSFI-D or FSDS-DAO #13 although their scores on an exit survey reflected non-meaningful change. Further, the FDA disagreed with the sponsor's definition of response on the FSFI-D, finding it to indicate too little change to be meaningful.

It is odd that Kingsberg et al. (2021) now refer to “clinically meaningful cutoffs” as defined in Revicki et al., yet they did not even report these categorical outcomes (FSFI-D and FSDS-DAO #13) in their 2019 paper! According to FDA definitions of clinically meaningful change, bremelanotide had a very modest benefit on the FSFI-D (NNT of 13) and no benefit on the FSDS-DAO #13. The validation of cutoff points selected for any other categorical efficacy outcome remains a mystery, not

**Table 1.** Debate surrounding main points from Spielmans (2021).

Main concern raised in Spielmans et al. (2021)	Response by Kingsberg et al. (2021)	My response to Kingsberg et al. (2021)
Most protocol-specified outcomes are unreported	“... [Spielmans] lists as a ‘problem’ that not everything in the study protocol was included in our Kingsberg et al. (2019) paper. In contrast to the Journal of Sex Research that provided 20 pages for Spielmans’ article, many, high impact, scientifically rigorous journals have a page limit	-Not reporting protocol-specified outcomes decreases transparency and violates CONSORT standards -Between <i>Obstetrics &amp; Gynecology’s</i> 5500-word limit and online appendices which do not count toward the word limit, all outcomes could have easily been reported.
Reporting of 15 non-protocol specified efficacy outcomes, which were apparently post-hoc	States that “the extensive statistical analysis plan that was prepared and reviewed by FDA defined each analysis in detail prior to database lock”. Thus, all analyses were purportedly determined prior to database lock.	The clinicaltrials.gov study protocols did not list 15 outcomes reported by Kingsberg et al. (2019). Most of these outcomes are not even mentioned in the FDA New Drug Application. I see no evidence that the outcomes, with the cutoff points used by Kingsberg et al. (2019), were defined <i>a priori</i> .
Several variables reported as showing favorable “trends” or as favoring treatment lack any numerical data	None	
Creating dichotomous outcomes from continuous outcomes without justification	None	
Lack of empirical justification for post-hoc measures	Cites Revicki et al. (2018) as providing validation for “patient reported outcomes”	Revicki et al. (2018) examined only cutoff points for treatment response on the FSFI-D and FSFS-DAO #13, outcomes which were unreported in Kingsberg et al. (2019). Kingsberg et al. provided no references to support validation of any post-hoc measures reported in their analyses.
Absolute benefit is incalculable for nearly all categorical analyses	None	
Number of dropouts due to adverse events is not reported by group	None	
Data reporting does not match CONSORT or GPP3 guidelines	None	
Change in coprimary outcome is unreported	The FDA allowed the change of primary outcomes and such a change was in line with evolving measurement standards for HSDD clinical trials.	I don’t dispute the change of primary outcome. Rather, I stated that the change of endpoints was unreported by Kingsberg et al. (2019), which Kingsberg et al. (2021) do not address.
Author and nonauthor contributions are unclear	None	

referenced by Kingsberg et al. in either their original paper or their current commentary. Thus, my critique about their seemingly arbitrary cutoff points remains unchallenged.

## Additional Concerns: Irrelevancies, Inaccuracies, and Distractions

### Meritless Attacks

Kingsberg et al. stated that “The attacks levied in [Spielmans’s] article and published by this journal, against virtually all parties involved in the development of [bremelanotide], the FDA and *Obstetrics & Gynecology* are without merit.” The best defense against truly meritless attacks is to describe *why* the attacks lack merit. Kingsberg et al., despite their bluster, failed to provide clear logic or evidence regarding how my criticisms of their paper (or “attacks”, if you like) lacked merit. Science improves through pointing out flaws. I hope my re-analysis of bremelanotide leads to both a) more complete data reporting and clearer justification for measures used in clinical trials and b) better peer reviews of clinical trial manuscripts.

### Too Late to Criticize and Insufficient Focus on Drug Benefits

Kingsberg et al. wrote that “With pre-specified efficacy requirements met and safety data reviewed, [bremelanotide] was

approved by the FDA. This surpasses a retrospective meta-analysis where the interpretation was clearly focused on the shortcomings of the treatment rather than the benefit it can provide.” They seem to imply that once a drug is approved by the FDA, any subsequent critiques of the drug’s clinical trials are no longer valid because they are retrospective. This is puzzling and will be addressed later in my response. While Kingsberg et al. maintained a strong, consistent focus on the “benefit it can provide,” the job of reporting information about dropout due to adverse events, poor treatment persistence, inadequate data reporting, and questionable outcome measures fell on me.

### Distractions: Irrelevancies

Kingsberg et al. noted that the FDA did not convene an advisory committee to discuss bremelanotide’s approval. This is entirely irrelevant. They accurately stated that my paper did not acknowledge a “decade of diligent research” on bremelanotide, which again is entirely irrelevant to my criticism of their paper regarding two specific phase III bremelanotide trials. They also stated that study sites received IRB approval and the trials were conducted according to good clinical practice; I said nothing to the contrary in my re-analysis. They stated that I may not understand the pernicious effects of untreated HSDD on individuals and relationships, citing no sources to support these claims. The phase III trials provided

no evidence that bremelanotide improved relationships and individual benefit is questionable, as I described in my re-analysis. They also stated that “The true measure of the clinical effectiveness of BMT is whether it can help patients with HSDD,” which “can only be addressed by a patient and her healthcare provider.” This is a reasonable point, and it actually underlines the need for my re-analysis, which adds much-needed clinically relevant information about bremelanotide’s effects as well as calling for a fuller reporting of the data on bremelanotide. If a healthcare provider is truly to practice evidence-based medicine, surely there is a need for transparently-reported data on treatment effects in clinical trials.

### ***Distractions: Rhetorical Devices and Ad Hominem Comments***

Kingsberg et al. referred to my analysis as “retrospective” five times. Perhaps this rhetorical device was an attempt to denigrate the validity of my analyses without directly addressing any of my concerns. “Retrospective” analyses have demonstrated that the published antidepressant clinical trial literature overstates efficacy relative to data from the same trials on file with the FDA (Jureidini et al., 2016; Kirsch et al., 2008; Turner et al., 2008). Following Kingsberg et al.’s apparent logic, one might wonder if such findings are invalid, if only clinical trial results as originally published in medical journals should be trusted, regardless of their actual accuracy or transparency.

The title of Kingsberg et al.’s commentary begins with “Failure of a meta-analysis”, which they contrast in their text with the “successful” bremelanotide trials. Perhaps this is another attempt to disparage my re-analysis without refuting any of its empirical findings. They also implicitly attack the *Journal of Sex Research*, stating “In contrast to the *Journal of Sex Research* that provided 20 pages for Spielman’s article, many high impact, scientifically rigorous journals have a page limit”. Impact factor does not necessarily imply greater rigor, but it seems that Kingsberg et al. (2021) have overlooked that the *Journal of Sex Research*’s impact factor rates near the top of sex- and gender-related journals (Zucker, 2021). In any case, no actual evidence is provided by Kingsberg et al. to support their apparent concern that the *Journal of Sex Research* lacks rigor. Also, as described earlier, Kingsberg et al. had plenty of room to report data completely and transparently in *Obstetrics & Gynecology*, but opted not to do so.

Kingsberg et al. (2021) stated that my interpretation of the data is based on “limited personal perspective” and is “uninformed”. Yet my conclusion that bremelanotide is “generally not useful” is based upon the empirical findings of the same clinical trials that Kingsberg et al. (2019) reported upon. More specifically, the rate of persistence (both completing a clinical trial and agreeing to participate in the follow-up open label trial) was much higher on placebo than bremelanotide: OR = .30, 95% CI = .24–.38, NNH: 4. Kingsberg et al. (2019) did not report this key finding and their commentary also ignores it. Would I be “uninformed” to ask: If bremelanotide is effective and well-tolerated, why were participants taking bremelanotide much less likely to complete a clinical trial?

Rather than specifically address my main points, Kingsberg et al. sought to diminish my credibility as a researcher via ad hominem arguments. For instance, they wrote “A researcher with expertise in sexual medicine would likely know the rigor and oversight of these trials” and note on multiple occasions that I am not a sexual medicine clinician. But it doesn’t take a sexual medicine clinician to document specific problems with inadequate data reporting, quite modest drug efficacy, and problematic tolerability of bremelanotide. Each of my claims was backed by evidence. It is very telling that Kingsberg et al. were unable to logically refute the accuracy of any of my analyses. They also stated that I am “biased”, while not describing what my alleged bias is, and providing zero evidence to support this claim. For those interested in biases, I point to the directly relevant commercial conflicts of interest which affect their entire authorship list. Indeed, four of the authors of Kingsberg et al. work either for the company that conducted the phase III trials and currently markets bremelanotide (Palatin Technologies) or a company that previously marketed bremelanotide (AMAG Pharmaceuticals).

If one takes the language of Kingsberg et al. at face value, one would likely conclude that Glen Spielman, an inexperienced, uninformed, biased researcher conducted a failed retrospective meta-analysis of the successful, rigorous clinical trials that led to the FDA’s approval of bremelanotide for hypoactive sexual desire disorder based on the comprehensive benefit it provided for participants. However, one might elect to ignore the ad hominem comments or various derogatory adjectives used to describe my re-analysis. Instead, one might actually read the empirical findings of my re-analysis of the bremelanotide trials, then compare them to Kingsberg et al.’s commentary. One might notice a striking lack of relevant arguments from Kingsberg et al. One might ask why six of my 10 main concerns were entirely unaddressed by Kingsberg et al. One might also note the lack of logical arguments or relevant evidence provided by Kingsberg et al. regarding the four of my 10 main points which they challenged.

Despite many clear points of disagreement, there is one area on which I concur entirely with Kingsberg et al. (2021). They recommend, and I strongly agree, that all readers “avail themselves of the published results from Kingsberg et al. (2019), where figure 3 presents the results of [bremelanotide] over placebo for eight individual patient reported outcomes within each of the two Phase III studies.” In figure 3, Kingsberg et al. see strong evidence for bremelanotide’s effectiveness, whereas I see a mirage of efficacy, consisting of incompletely reported measures of questionable validity. As described in my re-analysis, for each of the figure 3 outcomes: a) Kingsberg et al. provided no evidence that the cutoff scores for treatment response are valid, b) none were listed on the clinicaltrials.gov study protocols, so they appear to be post-hoc, c) none are reported according to CONSORT standards. Kingsberg et al. seem to view the measures and data reporting included in figure 3 as examples of good measurement practice and transparent data reporting. In actuality, figure 3 does not conform to the Good Publication Practice (GPP3) standards they inaccurately claimed to follow in their 2019 paper.

## Conclusion

When serious concerns with their data reporting and measurement practices were uncovered, Kingsberg et al. had several choices, including any combination of: a) demonstrating that some or all of the raised concerns lacked merit, b) admitting some faults in their manuscript and agreeing to set the record straight, c) denying fault without providing valid arguments or evidence, d) seeking to discredit the researcher who found the aforementioned problems rather than addressing the issues at hand. Unfortunately, they chose c) and d). I was looking forward to meaningful debate with Kingsberg et al., or perhaps for Kingsberg et al. to share the results on all protocol-listed outcomes. In review, they raised the following points to combat four of 10 claims in my analysis: claiming they did not need to report all protocol-listed outcomes(!), stating without evidence that all of their efficacy outcomes were *a priori*, inaccurately claiming that they validated cutoffs used for categorical efficacy measures used in their 2019 paper, and misstating that I critiqued their change of satisfactory sexual events to a secondary outcome. Their failure to even debate six of the ten main points raised in my re-analysis further validates my concerns with their article. I hope that future clinical trials of bremelanotide or any other treatment implement good measurement and data reporting practices, to avoid the many still-unresolved problems which occurred in Kingsberg et al. (2019)'s article.

## References

- Battisti, W. P., Wager, E., Baltzer, L., Bridges, D., Cairns, A., Carswell, C. I., Citrome, L., Gurr, J. A., Mooney, L. A., Moore, B. J., Peña, T., Sanes-Miller, C. H., Veitch, K., Woolley, K. L., & Yarker, Y. E. (2015). Good publication practice for communicating company-sponsored medical research: GPP3. *Annals of Internal Medicine*, 163(6), 461–464. <https://doi.org/10.7326/M15-0288>
- Jureidini, J. N., Amsterdam, J. D., & McHenry, L. B. (2016). The citalopram CIT-MD-18 pediatric depression trial: Deconstruction of medical ghostwriting, data mischaracterisation and academic malfeasance. *The International Journal of Risk & Safety in Medicine*, 28(1), 33–43. <https://doi.org/10.3233/JRS-160671>
- Kingsberg, S. A., Clayton, A. H., Portman, D., Williams, L. A., Krop, J., Jordan, R., Lucas, J., & Simon, J. A. (2019). Bremelanotide for the treatment of hypoactive sexual desire disorder: Two randomized phase 3 trials. *Obstetrics and Gynecology*, 134(5), 899–908. <https://doi.org/10.1097/AOG.0000000000003500>
- Kingsberg, S. A., Clayton, A. H., Portman, D., Williams, L. A., Krop, J., Jordan, R., Lucas, J., & Simon, J. A. (2021). Failure of a “meta-analysis”: A commentary on Glen Spielmans’ “Re-analyzing phase III bremelanotide trials for ‘hypoactive sexual desire disorder in women’”. *Journal of Sex Research*. <https://doi.org/10.1080/00224499.2021.1902926>
- Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine*, 5(2), e45. <https://doi.org/10.1371/journal.pmed.0050045>
- Obstetrics & Gynecology. (2020). *Instructions for authors: Obstetrics & gynecology*. Lippincott Williams & Wilkins. <https://journals.lww.com/greenjournal/Pages/instructionsforauthors.aspx>
- Revicki, D. A., Clayton, A. H., Stouch, B. C., Portman, D. J., Kingsberg, S. A., DeRogatis, L. R., & Jordan, R. (2018, February 20). *Developing clinically meaningful responder thresholds for primary endpoints for clinical trials in premenopausal women with hypoactive sexual desire disorder*. International Society for CNS Trials and Methodology.
- Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *PLoS Medicine*, 7(3), e1000251. <https://doi.org/10.1371/journal.pmed.1000251>
- Spielmans, G. I. (2021). Re-analyzing phase III bremelanotide trials for “hypoactive sexual desire disorder” in women. *Journal of Sex Research*. <https://doi.org/10.1080/00224499.2021.1885601>
- Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, 358(3), 252–260. <https://doi.org/10.1056/NEJMs065779>
- United States Food and Drug Administration. (2019). *New drug application multi-disciplinary review and evaluation standard 210557: Vyleesi/Bremelanotide*. United States Food and Drug Administration. [https://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2019/210557Orig1s000TOC.cfm](https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/210557Orig1s000TOC.cfm)
- Zucker, K. J. (2021). Archives at 50 years. *Archives of Sexual Behavior*, 50(1), 1–6. <https://doi.org/10.1007/s10508-021-01913-0>