Menopause: The Journal of The North American Menopause Society
Vol. 15, No. 5, pp. 885-889
DOI: 10.1097/gme.0b013e318182f84b

Measuring symptom relief in studies of vaginal and vulvar atrophy: the most bothersome symptom approach

Bruce Ettinger, MD,[1] Howard Hait, MS,[2] Kathleen Z. Reape, MD,[2] and Haibin Shu, MS[2]

Abstract

Objective: To assess the importance and usefulness of self-reported symptom data, especially the most bothersome symptom, in the evaluation of treatment for vulvovaginal atrophy.

Design: This was a double-blind, placebo-controlled multicenter study. Women rated symptoms associated with vaginal atrophy (vaginal dryness, vaginal/vulvar irritation/itching, vaginal/vulvar soreness, and dyspareunia) before and during treatment and selected one moderate to severe symptom as the most bothersome.

Results: Among 310 women (n = 156 placebo), vaginal dryness and dyspareunia were most commonly classified as moderate to severe and as most bothersome (44.4% and 30.2%, respectively). For both symptoms, the effect size favoring active treatment consistently increased as the cohort was more narrowly defined (all treated women, women who classified the symptom as moderate or severe, and those who classified the symptom as most bothersome). Compared with the standardized effect sizes for all women, those calculated from the most bothersome symptom were 49% and 62% greater for dyspareunia and dryness, respectively.

Conclusions: The most bothersome symptom approach represents a meaningful new standard for measurement of self-assessed vulvovaginal atrophy symptom change, but evaluation of change in individual symptoms remains an important, unbiased primary analysis of efficacy in vulvovaginal atrophy studies.

Key Words: Estrogen therapy – Hormone therapy – Vaginal atrophy – Vulvovaginal atrophy – Most bothersome symptom.

In January 2003, the US Food and Drug Administration (FDA) published a draft guidance for the conduct of clinical studies for the treatment of vulvovaginal atrophy (VVA).[1] Three years later, the FDA published a separate, more comprehensive, guidance outlining the manner in which self-reported scalar outcome measures should be developed and used to document symptomatic improvement among study participants.[2] For the treatment of VVA, a new product is now required to demonstrate efficacy at three coprimary endpoints: change in maturation index, change in vaginal pH, and change in severity of the most bothersome symptom (MBS). The first two objective measures have been widely used in clinical trials of VVA for many years, but the newly defined patient-reported MBS measure had not been used before the publication of the 2003 FDA guidance.

The MBS is derived from a selected list of symptoms, the most commonly reported being vaginal dryness, vaginal (or vulvar) irritation/itching, vaginal (or vulvar) soreness, and dyspareunia. At baseline, participants are instructed to rate each of these symptoms as not present, mild, moderate, or severe and then must select a single symptom among those classified as moderate or severe as the MBS. The MBS is then followed through to the end of treatment, and the change in its severity is used to evaluate symptomatic improvement.

The FDA proposed this new symptom measure because studies of VVA have lacked uniformity in symptom evaluation.[3] As the indicator of treatment efficacy, some studies have reported the change of symptom severity for individual symptoms,[4,5] whereas others have employed a composite score of several symptoms, weighted for severity.[6-10] Few required that vaginal symptoms at inclusion meet a level of moderate to severe intensity,[4,6,11,12] and only two published studies have reported change in MBS.[11,12]

The MBS construct is appealing because it examines each symptom individually and also requires those symptoms under consideration to be at least moderate in severity. Thus, each subject treated in the study contributes to the analyses, but only MBS symptoms are assessed when treatment efficacy is judged. Conversely, the MBS metric poses a number of potential statistical problems: it is not known whether this approach will provide variables with a smaller or larger treatment effect than other metrics, the precision error of this new construct is not known, and the numbers of women with each MBS symptom to be enrolled in clinical trials of VVA is not known. Thus, it is difficult to determine in advance whether a study under consideration will be adequately powered. Furthermore, not knowing in advance

Received March 27, 2008; revised and accepted June 11, 2008.

From the [1]University of California, San Francisco, CA; and [2]Duramed Research, Inc., Bala Cynwyd, PA.

Funding/Support: This work was supported by Duramed Research, Inc., a subsidiary of Barr Pharmaceuticals, Inc.

Financial disclosure: None reported.

Address correspondence to: Kathleen Reape, MD, Duramed Research, Inc., One Belmont Avenue, Suite 1100, Bala Cynwyd, PA 19004. E-mail: kreape@barrlabs.com

which symptom will be rated as the MBS, one cannot predict what the a priori distribution of women needs to be with respect to each symptom. Conceivably, all participants in one study may report vaginal dryness as their MBS, whereas all those in another study may report dypareunia as their MBS.

A recently published study of synthetic conjugated estrogens B 0.3-mg tablets (SCE-B) (Enjuvia; Duramed Pharmaceuticals, Inc., Pomona, NY) versus placebo for the treatment of VVA was designed to measure differences in symptom outcomes using the MBS endpoint.[12] Using the data obtained from that study, we sought to evaluate the effect size and statistical test results associated with various analyses of each symptom.

METHODS

Study design

The primary objective of the previously published study was to evaluate the safety and efficacy of oral SCE-B 0.3 mg once daily for the treatment of signs and symptoms of vaginal atrophy.[12] This was a double-blind, placebo-controlled multicenter investigation conducted at 42 sites in the United States. A total of 310 women with at least one moderate to severe symptom among six symptoms (vaginal dryness, vaginal irritation/itching, vaginal soreness, dysuria, dyspareunia, and bleeding after intercourse) were randomized and treated with either SCE-B (n = 154) or placebo (n = 156) for up to 12 weeks. The efficacy analysis included 125 women taking SCE-B and 123 women taking placebo who provided sufficient data for all three coprimary endpoints (vaginal pH, maturation index, and change in MBS) from baseline to week 12/end of treatment. The rate of study discontinuation was low; 82% of women taking SCE-B and 72% of women taking placebo completed 12 weeks of study. Last observed outcomes from those discontinuing prematurely were carried forward in the analyses. In the current report, we describe further analyses of these women to determine whether the magnitude of treatment effect was influenced by the MBS definition and whether other expressions of the symptom data lead to different conclusions about treatment efficacy.

Statistical analyses

For each symptom, the severity change from baseline was calculated as the postbaseline value (0 = no symptom present, 1 = mild, 2 = moderate, 3 = severe) minus the baseline value. Analysis of dyspareunia was limited to women reporting sexual intercourse at both baseline and at end of treatment. An analysis of covariance was used to compare mean changes in symptom severity between placebo and active treatment groups and to estimate the variance of those changes. Standardized treatment effect was calculated as the mean difference in change between the two groups divided by the S (Standard deviation) of that difference. A standardized effect size of 0.2 was considered small, 0.5 was medium, and 0.8 was large.[13]

RESULTS

Table 1 gives the baseline frequency distributions of the women who reported each self-assessed symptom as moderate to severe in intensity and as the MBS. Vaginal dryness and dyspareunia were the most frequently reported moderate to severe symptoms and were also most frequently classified as the MBS (44.4% and 30.2% of study participants, respectively). Because these two symptoms were most commonly reported as moderate to severe and also most commonly chosen as MBS, we focused specifically on them. Note that for dryness, the 110 women included in the MBS cohort represented only 49% of the total moderate to severe participant base with dryness; for dyspareunia, the MBS cohort of 75 women constituted 61% of the total moderate to severe subject base with dyspareunia.

Figures 1 and 2 show, respectively for vaginal dryness and dyspareunia, the mean changes from baseline in symptom severity in three ways: for all treated women, for those who classified the symptom as moderate or severe, and for those who classified the symptom as the MBS. For both symptoms, the treatment effect (difference in means between active and placebo) consistently increased as the subject cohort was more narrowly defined. In all cases for both symptoms, a strongly statistically significant ($P \leq 0.001$, adjusted for baseline severity) effect favoring SCE-B over placebo was observed, despite the substantial decrease in available sample size when the moderate to severe and MBS symptom definitions were used.

Table 2 compares changes in the standardized effect size by type of subject population used. Similar to the findings for

TABLE 1. *Baseline frequency distribution of moderate/severe vulvovaginal atrophy symptoms and the most bothersome symptom (MBS) by treatment group*

	No. of women					
	SCE-B (n = 125)		Placebo (n = 123)		Total (N = 248)	
Symptom	Moderate/severe	MBS	Moderate/severe	MBS	Moderate/severe	MBS
Dryness	114	56	110	54	224	110
Dyspareunia	55	35	67	40	122	75
Itching/irritation	47	19	49	15	96	34
Soreness	45	4	44	3	89	7
Difficulty during urination	36	11	30	11	66	22
Bleeding after intercourse	6	0	4	0	10	0

SCE-B, synthetic conjugated estrogens B.

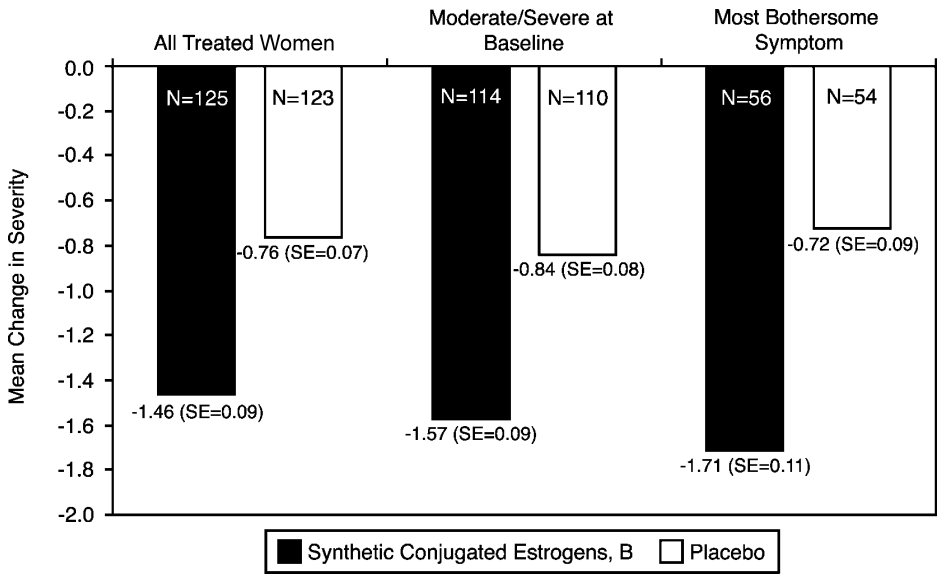

FIG. 1. Mean changes in vaginal dryness score from baseline to 12 weeks: by all treated women, by those reporting moderate to severe dryness, and by those reporting dryness as the most bothersome symptom.

treatment effect, the standardized effect size is greatest for the MBS, next greatest for those reporting baseline severity as moderate to severe, and smallest using data from all treated women. Thus, compared with the standardized effect sizes for all women, those calculated from the MBS were 49% and 62% greater for dyspareunia and dryness, respectively. Standardized effect sizes were large (≥1.0) for both dryness and dyspareunia when reported as the MBS. No consistent trends were seen in the magnitude of variances across the three analytic variations; differences among different metrics were relatively small (eg, compared with the variance calculated from changes of all women, the variances for the MBS were 24% higher for dyspareunia and 22% lower for dryness).

DISCUSSION

Our further analysis of a clinical study of treatment for VVA indicates that using the MBS increased the effect size and allowed statistically significant treatment effects to be shown in relatively small subgroups. Whereas demonstration of treatment effect was shown for the entire cohort, the apparent treatment effects are greater when mild symptoms are excluded from the analyses and when analyses focus on symptoms that are most bothersome. Our analysis of this

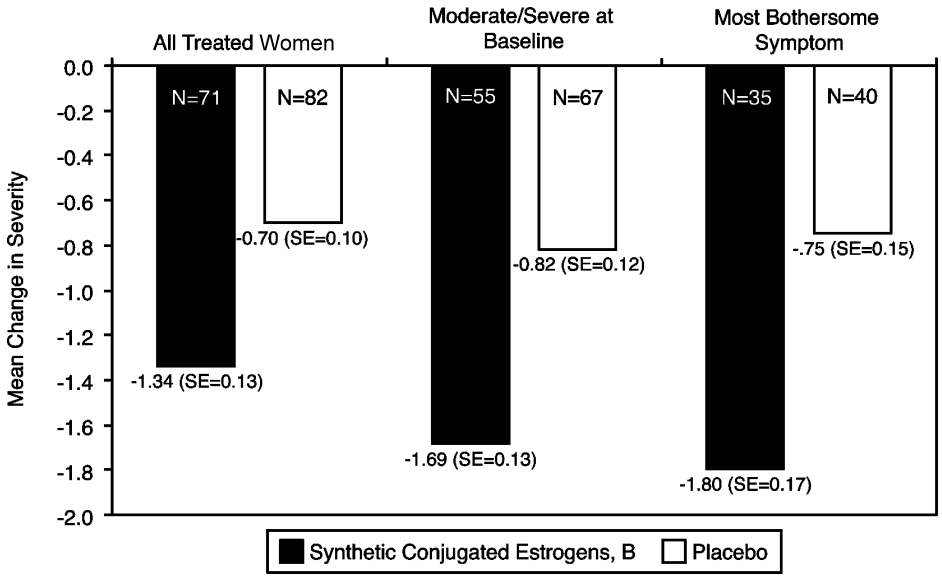

FIG. 2. Mean changes in dyspareunia score from baseline to 12 weeks: by all treated women, by those reporting moderate to severe dyspareunia, and by those reporting dyspareunia as the most bothersome symptom.

TABLE 2. *Standardized effect size comparisons for the reduction in vulvovaginal atrophy symptom severity between SCE-B and placebo*

	Dryness			Dyspareunia		
	All treated women	Moderate/severe at baseline	MBS	All treated women	Moderate/severe at baseline	MBS
No. of women	248	224	110	153	122	75
Mean baseline severity	2.33	2.48	2.53	2.17	2.61	2.72
SCE-B reduction in severity vs placebo						
Mean	0.70	0.73	0.99	0.64	0.87	1.05
Variance	0.67	0.70	0.52	0.85	0.93	1.05
Standardized effect size	0.86	0.88	1.39	0.70	0.91	1.04
P	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

MBS, most bothersome symptom; SCE-B, synthetic conjugated estrogens B.

single clinical trial of VVA shows that vaginal dryness and dyspareunia make up three fourths of the MBS cohort and the other symptoms associated with VVA are rarely most bothersome.

Although measurement and analytic techniques for vaginal cytology and vaginal pH are well understood, there appears to be little consensus on a generally accepted instrument for measuring symptoms, both to qualify women for study participation and to measure change from baseline to end of treatment. Because published studies have used different symptom endpoints, it has been difficult to differentiate among competing treatments' effects on symptoms of VVA.[3]

Clinical trials could be planned for women with moderate to severe symptoms, but one would have to restrict the number of symptoms under evaluation to just two or three to garner enough enrollees with a rating of moderate or severe for each of those symptoms. Otherwise, women would have to be enrolled in the study until reaching the requisite number of moderately to severely affected participants for each symptom. This would be wasteful of resources because there could be many noninformative women who qualified for enrollment based on one symptom but failed to qualify based on the others.

Questions remain about the clinical relevance of the MBS. The more severe or more troubled subgroup is an appropriate target for treatment, ie, consistent with the FDA's basis of approval of such products. Given that benefits of treatment apply to a broad spectrum of VVA symptoms, is it appropriate for product information material to state "relieves most bothersome symptoms of VVA" without proving this symptom by symptom? Because there is limited statistical power to show improvements in symptoms of low prevalence, eg, vaginal irritation or bleeding after intercourse, should these symptoms be included in the efficacy "umbrella"?

By eliminating women with mild severity who may still show a response and by eliminating women without a particular symptom at baseline who may develop the symptom during the course of study, investigators create an artificial population. Although we suspect that the response of the MBS may be the best reflection of treatment benefit for women who have very troubling symptoms, the MBS metric created by the FDA has no reliable distributional characteristics, has an overstatement of the expectations of treatment, and also represents a selected patient population.

What lessons can be learned from results reported from a wide variety of VVA study designs and what can be recommended for the optimal study design? First and foremost, self-reported improvement in VVA symptoms must be included as a coprimary outcome, along with objective measures such as maturation index and vaginal pH. This has not been the case in previous studies.[14,15] Second, inclusion should be limited to women with one and preferably two moderate to severe VVA symptoms; otherwise, treatment effects will be too small to document.[8–10] Third, changes in both individual symptoms and a composite score of all symptoms allow clinicians to better evaluate the effects of treatment. Fourth, change in symptom score relative to placebo should be at least one unit on the four-unit severity scale; this corresponds to a one-category improvement and to relief (defined as having none or mild symptoms) for those with moderate symptoms. Finally, alleviation of the most bothersome VVA symptom can be practically demonstrated in a properly designed study; MBS data should be shown symptom by symptom, whenever possible.

CONCLUSIONS

The FDA-recommended metric of the MBS is a major step forward in standardizing measurement of self-assessed VVA symptom changes. Evidence that the severity of individual moderate to severe symptoms is being reduced goes hand-in-hand with the MBS findings. Using data from our analyses, others should be able to better plan future studies of VVA treatment.

REFERENCES

1. US Department of Health and Human Services. Food and Drug Administration. Center for Drug Evaluation and Research (CDER). Guidance for Industry. estrogen and estrogen/progestin drug products to treat vasomotor symptoms and vulvar and vaginal atrophy symptoms-recommendations for clinical evaluation. Available at: http://www.fda.gov/cder/guidance/5412dft.pdf Accessed September 23, 2007.
2. US Department of Health and Human Services. Food and Drug Administration. Center for Drug Evaluation and Research (CDER). Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. Available at: http://www.fda.gov/cder/guidance/5460dft.htm Accessed September 23, 2007.

3. Suckling J, Lethaby A, Kennedy R. Local oestrogen for vaginal atrophy in postmenopausal women. Cochrane Database Syst Rev 2006;4: CD001500.
4. Barnabei VM, Cochrane BB, Aragaki AK, et al. Menopausal symptoms and treatment-related effects of estrogen and progestin in the Women's Health Initiative. *Obstet Gynecol* 2005;105:1063-1073.
5. Notelovitz M, Mattox JH. Suppression of vasomotor and vulvovaginal symptoms with continuous oral 17β-estradiol. *Menopause* 2000;7: 310-317.
6. Rioux JE, Devlin C, Gelfand MM, Steinberg WM, Hepburn DS. 17β-Estradiol vaginal tablet versus conjugated equine estrogen vaginal cream to relieve menopausal atrophic vaginitis. *Menopause* 2000;7:156-161.
7. Parsons A, Merritt D, Rosen A, et al. Study Groups on the Effects of Raloxifene HCI with Low-Dose Premarin Vaginal Cream. Effect of raloxifene on the response to conjugated estrogen vaginal cream or nonhormonal moisturizers in postmenopausal vaginal atrophy. *Obstet Gynecol* 2003;101:346-352.
8. Pinkerton JV, Shifren JL, La Valleur J, Rosen A, Roesinger M, Siddanti S. Influence of raloxifene on the efficacy of an estradiol-releasing ring for treating vaginal atrophy in postmenopausal women. *Menopause* 2003;10:45-52.
9. Speroff L, Haney AF, Gilbert RD, Ellman H. Efficacy of a new, oral estradiol acetate formulation for relief of menopausal symptoms. *Menopause* 2006;13:442-450.
10. Speroff L. Efficacy and tolerability of a novel estradiol vaginal ring for relief of menopausal symptoms. *Obstet Gynecol* 2003;102: 823-834.
11. Simon JA, Bouchard C, Waldbaum A, Utian W, Zborowoski J, Snabes MC. Low doses of transdermal estradiol gel for treatment of symptomatic postmenopausal women. *Obstet Gynecol* 2007;109: 588-596.
12. Simon JA, Reape KZ, Wininger S, Hait H. A randomized, double-blind, placebo-controlled trial to evaluate the efficacy and safety of synthetic conjugated estrogens B for the treatment of vulvovaginal atrophy in healthy postmenopausal women. *Fert Steril* Epub ahead of print November 28, 2007.
13. Cohen J. A power primer. *Psychol Bull* 1992;112:155-159.
14. Marx P, Schade G, Wilbourn S, Blank S, Moyer DL, Nett R. Low-dose (0.3 mg) synthetic conjugated estrogens A is effective for managing atrophic vaginitis. *Maturitas* 2004;47:47-54.
15. Utian WH, Shoupe D, Bachmann G, Pinkerton JV, Pickar JH. Relief of vasomotor symptoms and vaginal atrophy with lower doses of conjugated equine estrogens and medroxyprogesterone acetate. *Fert Steril* 2001;75:1065-1079.