

ORIGINAL STUDY

TX-001HR is associated with a clinically meaningful effect on severity of moderate to severe vasomotor symptoms in the REPLENISH trial

Ginger D. Constantine, MD,¹ James A. Simon, MD,² Andrew M. Kaunitz, MD,³ James H. Pickar, MD,^{4,5} Dennis A. Revicki, PhD,⁶ Shelli Graham, PhD,⁷ Brian Bernick, MD,⁷ and Sebastian Mirkin, MD⁷

Abstract

Objective: The aim of the study was to evaluate the clinically meaningful effect of oral TX-001HR (17 β -estradiol [E2]/progesterone [P4]) capsules on hot flushes severity (vasomotor symptoms [VMS] severity scale) using the patient-reported Clinical Global Impression (CGI).

Methods: REPLENISH (NCT01942668) was a phase 3, randomized, double-blind, placebo-controlled, multicenter trial that evaluated TX-001HR in postmenopausal women (40-65 y) with a uterus. Those with frequent moderate to severe hot flushes ($\geq 7/d$ or $\geq 50/wk$) were randomized in a VMS substudy to daily E2/P4 (1/100, 0.5/100, 0.5/50, or 0.25/50 mg/mg), or placebo. Patients rated VMS severity from 1 (mild) to 3 (severe) and symptom improvements with the CGI. CGI results were an anchor in a nonparametric discriminant analysis to define clinically important differences (CIDs) and minimal CID in VMS severity at weeks 4 and 12.

Results: In the VMS substudy ($n = 726$), determined CID and minimal CID severity thresholds were reductions of 0.525 and 0.350 points at week 4, respectively, and 0.775 and 0.225 points at week 12. Significantly more women taking the two highest E2/P4 doses (1/100 and 0.5/100) versus placebo met CID severity thresholds at weeks 4 (40% and 44% vs 17%; $P < 0.05$) and 12 (56% and 48% vs 29%; $P < 0.05$).

Conclusion: REPLENISH trial data demonstrated that E2/P4 1/100 and 0.5/100 provided clinically meaningful improvements in hot flushes severity in postmenopausal women. In conjunction with previously demonstrated clinically meaningful VMS frequency improvements, these data support oral E2/P4 1/100 and 0.5/100 for postmenopausal women with a uterus seeking treatment for moderate to severe VMS.

Key Words: Estradiol – Hot flushes – Menopause – Progesterone – Severity – Vasomotor.

Vasomotor symptoms (VMS) represent one of the most troublesome symptoms of menopause and are experienced to some degree by most postmenopausal women.¹ Moderate to severe hot flushes increase sharply in the first 2 years before the final menstrual period

(FMP), reaching a peak in the first few years after the FMP, slowly decreasing to premenopausal levels until 8 to 9 years after FMP.^{2,3} In a meta-analysis (10 studies; 35,445 participants), bothersome symptoms peaked about 1 year earlier and decline more rapidly than other nonbothersome symptoms.³

Received March 18, 2020; revised and accepted April 30, 2020.

From the ¹EndoRheum Consultants, LLC, Malvern, PA; ²IntimMedicine Specialists, Washington, DC; ³University of Florida College of Medicine-Jacksonville, Jacksonville, FL; ⁴Columbia University Medical Center, New York, NY; ⁵KMITL Faculty of Medicine, Bangkok, Thailand; ⁶Patient-Centered Research, Evidera, Bethesda, MD; and ⁷TherapeuticsMD, Boca Raton, FL.

Clinical Trial Registration: <https://clinicaltrials.gov/ct2/show/NCT01942668>.

Funding/support: TherapeuticsMD sponsored the study, and provided support for the medical writing assistance of Dominique Verlaan, PhD, CMPP (Precise Publications, LLC).

Financial disclosure/conflicts of interest: Dr. Constantine consults to multiple pharmaceutical companies including but not limited to TherapeuticsMD, and has stock options from TherapeuticsMD. Dr. Simon has served (within the past year, or current) as a consultant/advisor to AbbVie, AMAG, Bayer Healthcare, CEEK Enterprises, Covance, Dare Bioscience, Duchesnay, Hologic, KaNDy/NeRRe Therapeutics, Mitsubishi Tanabe, Shionogi, Sprout2, and TherapeuticsMD; has received (within the past year, or current) grant/research support from AbbVie,

Bayer Healthcare, Endoceutics, GTx, Ipsen, Myovant Sciences, ObsEva SA, TherapeuticsMD, and Viveve Medical; has also served (within the past year, or current) on the speaker's bureaus of AbbVie, AMAG, Duchesnay, and TherapeuticsMD; and is a stockholder (direct purchase) in Sermonix Pharmaceuticals. Dr. Pickar has consulted for Pfizer, Shionogi, and TherapeuticsMD and has stock options with TherapeuticsMD. Dr. Kaunitz has served as a consultant to or on the advisory board of AMAG, Mithra, and Pfizer; and has received research support (to University of FL) from Mithra and TherapeuticsMD. Dr. Revicki is a consultant for AMAG, Palatin Technologies, and TherapeuticsMD. Dr. Graham, Dr. Bernick, and Dr. Mirkin are employees of TherapeuticsMD with stock/stock options.

Address correspondence to: Ginger D. Constantine, MD, EndoRheum Consultants, LLC, 212 Mine Rd, Malvern, PA 19355.

E-mail: endorheum@gmail.com

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Almost half of women reported VMS 4 years after the FMP, and 10% reported symptoms 12 years after the FMP.³ A 16-year US study of women with natural menopause reported that moderate to severe VMS lasted almost 15 years (mean 9 years), and VMS of any severity also lasted up to 15 years (mean 10 years).⁴

Bothersome VMS negatively affect health-related quality of life, including women's social and emotional functioning, mental health, energy levels, bodily pain, sleep, and sexual satisfaction.⁵⁻⁷ The severity of VMS specifically adversely affects quality of life, and may be more bothersome than frequency of symptoms.⁸ In the 2010 US National Health and Wellness Survey, women experiencing moderate to severe VMS had lower mean health status scores, used more health-care resources, and had higher impairment while at work and in activities of daily living than those with mild symptoms.⁹

Although hormone therapy (HT) can significantly reduce the frequency and severity of VMS, levels of reduction in VMS frequency and severity that are *clinically* significant for women have not been systematically evaluated. In general, the benefits of treatment to patients in clinical trials are measured using patient-reported outcomes (PROs). These instruments are now becoming a standard assessment tool used by the US Food and Drug Administration (FDA) when evaluating a product for approval.¹⁰⁻¹² PROs are important assessments that can help determine effects that are clinically meaningful to patients by gathering data on a patient's health condition reported directly from the patient, and not interpreted by a clinician or anyone else.^{11,12} For each study, PROs are specifically tailored to the population and drug being evaluated. Although PROs are not normally used to determine clinically meaningful reductions in specific endpoints such as VMS frequency and severity in patients, they can be used as anchors in analyses to determine clinically important differences (CIDs) in these specified outcomes. For VMS frequency, several methods have been proposed that include a difference in the given number of hot flushes of the active versus placebo groups, in responders, and in PRO anchor-based analyses, but these have never been performed for VMS severity. In this analysis, the Clinical Global Impression (CGI) scale, a subjective PRO, is used as an anchor to determine CID in VMS severity that is meaningful to patients.

In the 12-week VMS substudy of the REPLENISH trial, most doses of TX-001HR (TherapeuticsMD, Boca Raton, FL), a once-daily, oral capsule containing 17 β -estradiol (E2) and progesterone (P4), significantly reduced the frequency and severity of moderate to severe VMS in postmenopausal women with a uterus.¹³ At week 12, VMS severity significantly improved by 0.76 to 1.12 points from baseline and 0.24 to 0.57 points compared with placebo (Fig. 1).¹³ The 1 mg E2/100 mg P4 dose was approved by the US FDA as Bijuva ([estradiol and progesterone] capsules; TherapeuticsMD) in October 2018 for the treatment of moderate to severe VMS in postmenopausal women with a uterus. Overall, TX-001HR was shown to have a clinically meaningful effect to participants by reducing VMS frequency.¹⁴

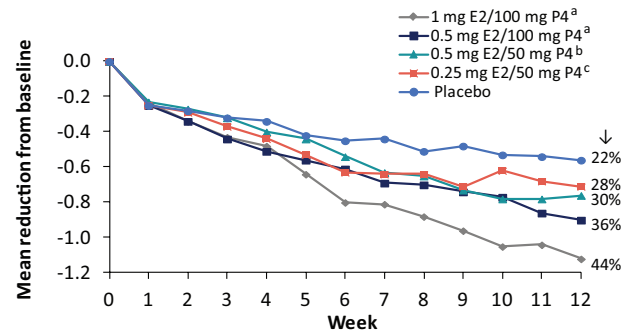


FIG. 1. Reductions in moderate to severe vasomotor symptoms severity in the REPLENISH trial. $P < 0.05$ from ^aweeks 3 to 12, ^bweeks 7, 9 to 12, ^cweeks 6, 7, 9 versus placebo. E2, estradiol; P4, progesterone. Figure adapted from Lobo et al 2018.¹³

The objective of this analysis was to determine the clinically meaningful effects of oral E2/P4 when reducing the severity of hot flushes in postmenopausal women.

METHODS

Study design

The REPLENISH (NCT01942668) study was a phase 3, randomized, double-blind, placebo-controlled, multicenter trial that evaluated TX-001HR for the treatment of moderate to severe VMS in postmenopausal women with a uterus.¹³ Healthy menopausal women, 40 to 65 years of age with body mass index of 34 kg/m² or below, seeking treatment for VMS could participate in this study. Exclusion criteria were typical of other HT trials and are described elsewhere.¹³ To evaluate the efficacy of TX-001HR in a subgroup of participants with the most severe VMS in the study, women with moderate to severe hot flushes ($\geq 7/d$ or $\geq 50/wk$) were enrolled in a 12-week VMS substudy and randomized 1:1:1:1 to daily oral E2/P4 (mg/mg) 1/100, 0.5/100, 0.5/50, 0.25/50, or placebo. Women with fewer VMS were randomized 1:1:1:1 to E2/P4 doses to assess the endometrial and general safety of TX-001HR for up to 1 year; endometrial and general safety were also assessed in women enrolled in the VMS substudy. Details of the study, including results of the primary efficacy and safety endpoints, as well as clinical meaningful thresholds for VMS frequency have been published.^{13,14}

Clinical meaningfulness

The four coprimary efficacy endpoints of the REPLENISH trial were mean changes from baseline to weeks 4 and 12 in the frequency and severity of moderate to severe VMS with E2/P4 compared with placebo. Participants in the VMS substudy completed a daily VMS diary and recorded the number and severity of hot flushes up to week 12. Severity of VMS was defined as mild (sensation of heat without sweating), moderate (sensation of heat with sweating, able to continue activity), or severe (sensation of heat with sweating, causing cessation of activity). Baseline weekly severity score was calculated by: [(number of moderate hot flushes for

7 days) × 2 + (number of severe hot flushes for 7 days) × 3] / (total number of moderate to severe hot flushes over 7 days), and on treatment weekly severity score was calculated by: [(number of mild hot flushes for 7 days) × 1 + (number of moderate hot flushes for 7 days) × 2 + (number of severe hot flushes for 7 days) × 3] / (total number of mild, moderate, and severe hot flushes over 7 days).

The clinical meaningfulness of treatment to patients was assessed using the CGI scale, which was a secondary endpoint of the trial. Participants answered the following question at baseline, and weeks 4, 8, and 12: “Rate the total improvement, whether or not in your judgment it is due entirely to drug treatment. Compared to your condition at admission to the study, how much has it changed?” using a 7-point Likert scale corresponding to very much improved, much improved, minimally improved, no change, minimally worse, much worse, and very much worse. The proportion of participants for each CGI category was summarized at weeks 4 and 12, and each E2/P4 dose was compared with placebo using the Fisher’s exact test. Note that when a subset of participants from the VMS substudy were queried about what the term “condition” meant, all confirmed it referred to their VMS.

To relate the clinical meaningfulness of treatment to changes in VMS severity that are clinically significant to patients, the CGI responses were grouped into three categories: clinically meaningful (much or very much improved) corresponding to CIDs, minimally improved corresponding to minimal CIDs (MCIDs), and no change or worse (no change to very much worse). Based on these CGI response groups, a three categorical variable was constructed (regardless of treatment), and a nonparametric discriminate analysis was conducted utilizing bootstrapping methods,^{15,16} to measure CIDs and MCIDs reductions in moderate to severe VMS severity at weeks 4 and 12. The proportion of women who were responders based on the calculated CID and MCID thresholds with E2/P4 were compared with those taking placebo using the Fisher’s exact test.

All analyses were performed in the modified intent-to-treat (MITT)-VMS population (primary efficacy population),

which included women who had been randomized to the VMS substudy, took at least one dose of study medication, had ≥5 days of VMS diary data at baseline, and had ≥4 days of VMS diary data for one on-treatment week.

RESULTS

Disposition and demographics

Of the 1,835 women randomized in the REPLENISH trial, a total of 766 women were enrolled in the VMS substudy and 726 were included in the MITT-VMS population. Of these women, 89% (647/726) completed the 12-week VMS efficacy substudy. At baseline, women included in the MITT-VMS population were on average 54.6 years of age (range, 40-65) and had a mean BMI of 26.6 kg/m² (range, 14.0-34.5); most were white (66.9%), or African American (31.0%; Table 1). The weekly severity of moderate to severe VMS at baseline ranged from 2.50 to 2.54 points (possible range of 1 through 3).

Clinical meaningfulness

At week 4, the percentage of women who reported that they were very much or much improved on the CGI ranged from 50% to 63% for the E2/P4 groups compared with 33% for placebo (Fig. 2A). The greatest proportion of women who reported no change or reported worsening severity of symptoms were those taking placebo (28%). At week 12, the range of women who reported they were very much or much improved for the E2/P4 groups was 73% to 82% compared with 53% for the placebo group (Fig. 2B), with the highest percentage of improvement reported for the 1/100 dose. Statistically significant improvements were observed for all timepoints for the E2/P4 groups compared with the placebo group; all groups showed improvements over time.

Based on the best discrimination between women who reported much or very much improvement and those women who reported minimal improvement (Fig. 3A and B), the threshold for reporting a CID decrease in weekly severity of VMS was a decrease of 0.525 points at week 4 and a decrease of 0.755 points at week 12. The threshold for reporting an MCID was based on the best discrimination between women

TABLE 1. Participant demographics and baseline characteristics of the modified intent-to-treat -vasomotor symptoms population

Characteristic	Estradiol/progesterone, mg/mg				Placebo
	1/100	0.5/100	0.5/50	0.25/50	
N	141	149	147	154	135
Age, y	54.7 ± 4.8	54.9 ± 4.5	54.8 ± 4.6	54.5 ± 3.8	54.3 ± 4.3
Race, n (%)					
White	95 (67.4)	99 (66.4)	99 (67.3)	102 (66.2)	91 (67.4)
African American	45 (31.9)	48 (32.2)	43 (29.3)	48 (31.2)	41 (30.4)
Other ^a	1 (0.7)	2 (1.3)	5 (3.4)	4 (2.6)	3 (2.2)
BMI, kg/m ²	26.5 ± 3.9	27.1 ± 4.3	26.6 ± 3.9	26.4 ± 4.0	26.6 ± 3.8
Time since menopause, y	6.1 ± 5.5	6.5 ± 5.4	6.0 ± 4.8	5.2 ± 4.8	5.7 ± 4.9
Bilateral oophorectomy	3 (2.1)	3 (2.0)	1 (0.7)	1 (0.6)	0
Baseline VMS parameters					
Weekly frequency	74.4 ± 35.3	72.1 ± 27.8	75.9 ± 28.0	77.0 ± 30.4	72.4 ± 23.3
Weekly severity	2.54 ± 0.32	2.51 ± 0.25	2.50 ± 0.23	2.51 ± 0.26	2.52 ± 0.25

Data presented as mean ± SD, unless stated otherwise.

BMI, body mass index; SD, standard deviation; VMS, vasomotor symptoms.

^aOther includes other (n = 10), American Indian or Alaska Native (n = 2), Native Hawaiian or Pacific Islander (n = 2), and unknown (n = 1).

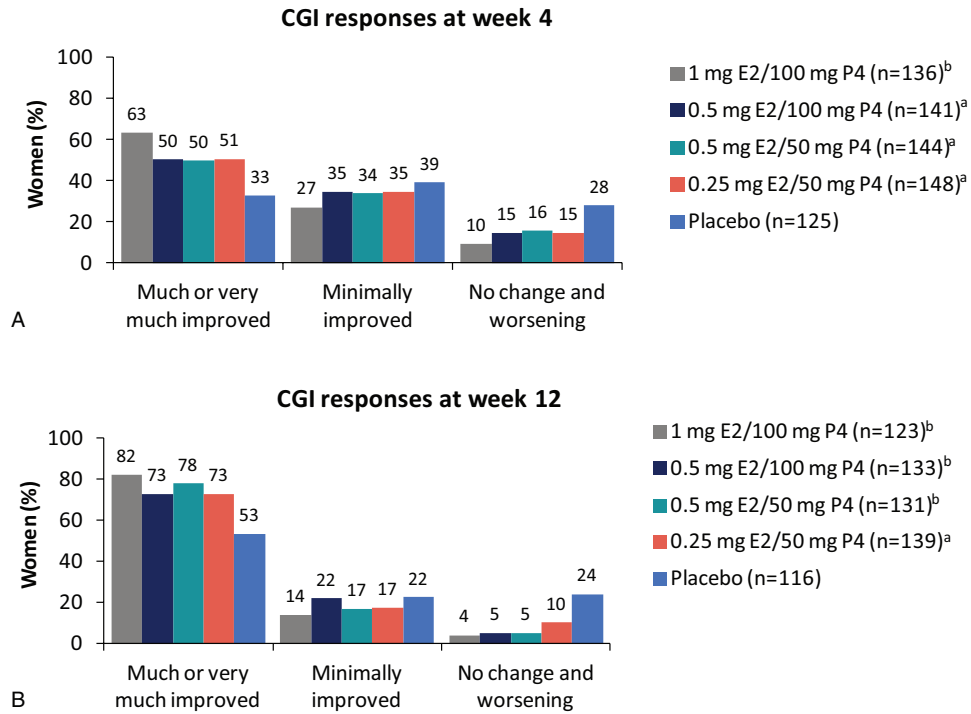


FIG. 2. Clinical Global Impression (CGI) responses at (A) week 4 and (B) week 12. ^a*P* < 0.01; ^b*P* < 0.001 versus placebo, calculated with Fisher's exact test.

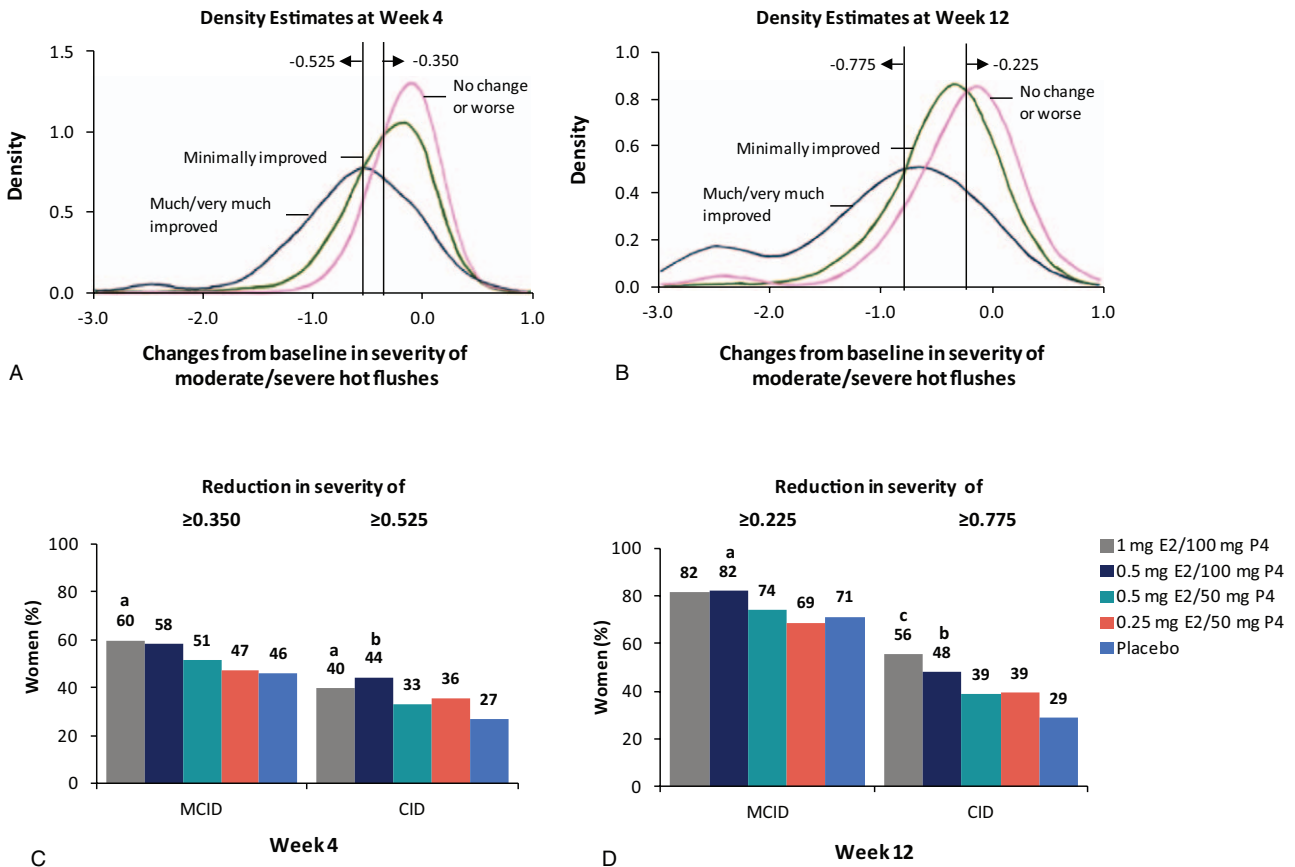


FIG. 3. Clinical meaningfulness threshold analysis at (A) week 4 and (B) week 12; CGI-based CID and MCID in VMS severity at (C) week 4 and (D) week 12. ^a*P* < 0.05; ^b*P* < 0.01; ^c*P* < 0.001 versus placebo, calculated with Fisher's exact test. CID, clinically important difference; CGI, Clinical Global Impression; E2, estradiol; MCID, minimal clinically important difference; P4, progesterone.

who reported minimal improvements and those who reported no change or worse, and was a decrease of 0.350 points at week 4 and a decrease of 0.225 points at week 12.

Based on these VMS severity thresholds, there were significantly more clinical responders for the E2/P4 groups than with placebo at week 4 (Fig. 3C) and week 12 (Fig. 3D) for both MCID and CID.

DISCUSSION

In this post hoc analysis from the REPLENISH trial, reducing VMS severity was considered clinically meaningful by participants. This is the first report of an HT product demonstrating that reductions in VMS severity observed with treatment were clinically meaningful to women.

Careful assessment of patients' views on benefits and risks are an important part of the regulatory decision-making process when evaluating the efficacy of a product. The FDA considers patients to be the experts on what it is like to live with their condition; however, patients' main complaints may not be factored explicitly into drug development plans, including measures of a particular drug's benefit in planned clinical trials. Clinically meaningful effects of an intervention can be measured with PROs, which assess how a patient feels or functions in daily life. An MCID can be defined as the smallest change in a PRO endpoint score that is considered beneficial and important to patients,^{11,12} whereas a CID can be defined as a threshold at which the treatment outcome is experienced as relevant and clinically meaningful by the patient.

In VMS clinical trials, the efficacy assessment of products for menopausal VMS is often based on statistically significant reductions of VMS severity and frequency with treatment compared with placebo as well as comparing the proportion of clinically meaningful responders by treatment groups using PROs.¹⁷⁻²⁰ A reduction of at least 2 hot flushes per day or 14 hot flushes per week has been suggested, by some, to determine whether a treatment is clinically beneficial to patients; however, this does not consider a possible clinically meaningful reduction in VMS severity. An absolute reduction of 2 hot flushes per day or 14 hot flushes per week between treatment and placebo,¹⁷⁻¹⁹ is dependent on placebo response rates and may not be sufficient to determine the lowest effective dose in patients who may have more severe, but less frequent hot flushes; treatment for these patients may be different.

While, the CGI anchor-based method used here has been previously used to determine VMS frequency thresholds in two studies evaluating treatment for moderate to severe VMS in menopausal women, it has not been used to calculate VMS severity thresholds.^{14,15} In fact, few studies have evaluated clinically meaningful effects of VMS severity. One study of desvenlafaxine found MCID results similar to those found in this study, where an MCID threshold for severity of 0.20 was reported between women who were satisfied with treatment and those who had a neutral experience at week 12; no CID threshold was calculated.¹⁸ The results from our investigation demonstrated CIDs in VMS severity to be reductions of 0.525 points or more at week 4 and 0.775 points or more at week 12,

and based on these thresholds, women had clinically meaningful benefit in their VMS severity when treated with 1/100 or 0.5/100 E2/P4 doses. These results are consistent with previously published REPLENISH data, which reported clinically meaningful benefits in their VMS frequency with the same doses,¹⁴ as well as significant reductions in the frequency and severity of moderate to severe VMS as early as week 4, which were maintained to week 12.¹³

REPLENISH trial's limitations include that women in the study were healthier than the general population, and resided in the United States, which may not be representative of worldwide women seeking treatment for VMS. The REPLENISH trial also had several strengths, including the fact that it was a well-designed, double-blind, randomized, placebo-controlled trial planned under FDA guidance, and was similar to other clinical trials that have evaluated treatments for menopausal VMS.

CONCLUSIONS

In the REPLENISH trial, postmenopausal women with a uterus receiving 1/100 and 0.5/100 E2/P4 doses had more clinically meaningful improvements in VMS severity score as measured by the CGI-based analysis. TX-001HR combines bioidentical E2 and P4 in a single softgel capsule and the two E2/P4 doses (1/100 and 0.5/100) demonstrated clinically significant relief from moderate to severe VMS in postmenopausal women with a uterus.

Acknowledgments: The authors would also like to acknowledge the medical writing assistance of Dominique Verlaan, PhD, CMPP of Precise Publications, LLC, which was supported by TherapeuticsMD.

REFERENCES

1. Williams RE, Kalilani L, DiBenedetti DB, et al. Frequency and severity of vasomotor symptoms among peri- and postmenopausal women in the United States. *Climacteric* 2008;11:32-43.
2. Freeman EW, Sammel MD, Sanders RJ. Risk of long-term hot flashes after natural menopause: evidence from the Penn Ovarian Aging Study cohort. *Menopause* 2014;21:924-932.
3. Politi MC, Schleinitz MD, Col NF. Revisiting the duration of vasomotor symptoms of menopause: a meta-analysis. *J Gen Intern Med* 2008;23:1507-1513.
4. Freeman EW, Sammel MD, Lin H, Liu Z, Gracia CR. Duration of menopausal hot flushes and associated risk factors. *Obstet Gynecol* 2011;117:1095-1104.
5. Ayers B, Hunter MS. Health-related quality of life of women with menopausal hot flushes and night sweats. *Climacteric* 2013;16:235-239.
6. Savolainen-Peltonen H, Hautamaki H, Tuomikoski P, Ylikorkala O, Mikkola TS. Health-related quality of life in women with or without hot flashes: a randomized placebo-controlled trial with hormone therapy. *Menopause* 2014;21:732-739.
7. Utian WH. Psychosocial and socioeconomic burden of vasomotor symptoms in menopause: a comprehensive review. *Health Qual Life Outcomes* 2005;3:47.
8. Pinkerton JW, Abraham L, Bushmakin AG, Cappelleri JC, Komm BS. Relationship between changes in vasomotor symptoms and changes in menopause-specific quality of life and sleep parameters. *Menopause* 2016;23:1060-1066.
9. Whiteley J, Wagner JS, Bushmakin A, Kopenhafer L, Dibonaventura M, Racketa J. Impact of the severity of vasomotor symptoms on health status, resource use, and productivity. *Menopause* 2013;20:518-524.
10. US Department of Health and Human Services (FDA). Guidance for industry - patient-reported outcome measures: use in medical product

- development to support labeling claims. December 2009. Available at: <https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf>. Accessed February 2018.
11. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171-178.
 12. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertain the minimal clinically important difference. *Control Clin Trials* 1989;10:407-415.
 13. Lobo RA, Archer DF, Kagan R, et al. A 17 β -estradiol-progesterone oral capsule for vasomotor symptoms in postmenopausal women: a randomized controlled trial. *Obstet Gynecol* 2018;132:161-170.
 14. Constantine GD, Revicki DA, Kagan R, et al. Evaluation of clinical meaningfulness of estrogen plus progesterone oral capsule (TX-001HR) on moderate to severe vasomotor symptoms. *Menopause* 2019;26:513-519.
 15. Gerlinger C, Gude K, Hiemeyer F, Schmelter T, Schafers M. An empirically validated responder definition for the reduction of moderate to severe hot flushes in postmenopausal women. *Menopause* 2012;19:799-803.
 16. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102-109.
 17. Archer DF, Freeman EW, Komm BS, et al. Pooled analysis of the effects of conjugated estrogens/bazedoxifene on vasomotor symptoms in the Selective Estrogens, Menopause, and Response to Therapy Trials. *J Womens Health (Larchmt)* 2016;25:1102-1111.
 18. Wyrwich KW, Spratt DI, Gass M, Yu H, Bobula JD. Identifying meaningful differences in vasomotor symptoms among menopausal women. *Menopause* 2008;15:698-705.
 19. US Food and Drug Administration. Center for Drug Evaluation and Research. Application number: 022247Orig1s000. Summary review. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/022247Orig1s000SumR.pdf. Accessed February 14, 2018.
 20. US Department of Health and Human Services (FDA). Guidance for industry: patient-reported outcome measures - use in medical product development to support labeling claims. December 2009. Available at: <http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf>. Accessed December 2016.