Research on the factors associated with placebo response in clinical trials for Vasomotor Symptoms

Kentaro Miyazaki

Department of Clinical Medicine (Pharmaceutical Medicine)

Graduate School of Pharmaceutical Sciences

Kitasato University

5-9-1 Shirokane, Minato-ku, Tokyo, 108-8641, Japan

Abstract

Background

Hot flashes and night sweats, also known as vasomotor symptoms (VMS), are the hallmark symptoms of menopause. Hormone therapy (HT) is recognized as the gold standard for the treatment of VMS; however, many women choose not to use HT due to the potential risks of side effects (e.g., heart disease and breast/endometrial cancer). The other treatment options include nonhormonal pharmacologic therapy (non-HT) and complementary and alternative medicine (CAM), but their efficacy and side-effect profile have limited their use for the treatment of VMS as a regulatory approved product. It was reported that high placebo response can be seen in the evaluation of active treatment in clinical studies for women with hot flashes and potentially undermine the evaluation of new treatments. Many studies have been conducted to seek the reasons of high placebo response within a study or among studies; however, there is no definite answer as of now. The aim of this study was to determine the factors associated with high placebo response in randomized, controlled, double-blind studies enrolling women with hot flashes.

Methods

In the research 1, the magnitude of placebo response was defined as the reduction in the mean number of hot flash frequency from baseline. In the research 2, coefficient of variation (CV) was calculated and used as an index of the variability of placebo response. To identify eligible studies, Embase, MEDLINE, and BIOSIS Previews were searched for English-language articles published between April 1975 and August 2020. Placebo-controlled, double-blind, randomized studies that assessed changes in hot flash frequency

were included if they satisfied the pre-defined criteria. We conducted univariate/ multivariate analyses for research 1, and simple linear regression/multiple regression analysis for research 2 using categorical and numerical data as explanatory variables. Categorical data included the following variables with levels in brackets: active treatment type (HT/non-HT/CAM), administration route (oral/non-oral), study region (in/excluded the US), breast cancer population (in/excluded), entry criteria of hot flash severity (moderate to severe only/all included), parallel or crossover study, placebo run-in period before treatment (yes/no), and menopausal status (postmenopausal only/include perimenopausal/include premenopausal). Numerical data included published year, pretreatment period duration, treatment period duration, number of sites, number of total participants, number of placebo participants, number of treatment arms, mean age, body mass index (BMI), and hot flash frequency at baseline.

Results

Forty-three of the 802 identified publications were included in the analysis. In the research 1, multivariate analysis identified five individual factors associated with placebo response: treatment period duration, number of treatment arms, and BMI for the higher placebo response; active treatment type and breast cancer population included for the lower placebo response. In the research 2, multiple regression analysis identified two individual factors associated with variability of placebo response: menopause status for higher variability; hot flash frequency at baseline for lower variability.

Conclusion

Several factors associated with placebo response in clinical studies of women with hot flashes were identified. Knowing these factors, especially relevant to the high placebo

response may enable proactive implementation of operational and analytic strategies that further aid in determining the true treatment effect of an intervention. The outcome of this research will directly contribute to the revitalization of the new drug development in the women's health area where the therapeutic options are limited against the high unmet needs.

Table of Contents

List of Tables

List of Figures

Abbreviations

BC	Breast cancer
BMI	Body mass index
CAM	Complementary and Alternative Medicine
CV	Coefficient of variation
E2	Estradiol
EMA	European Medicines Agency
FDA	The U.S. Food and Drug Administration
HT	Hormone therapy
P	Progesterone
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCTs	Randomized clinical trials
SD	Standard deviation
SE	Standard error
US	United States of America
VMS	Vasomotor symptoms

1. Introduction

Vasomotor symptoms (VMS), characterized by hot flashes and/or night sweats, are the most common menopausal symptoms experienced by women transitioning through menopause [1]. Hot flashes occur many of the menopausal women and often described as episodic sensations of heat, intense sweating, and flushing affecting the face and chest, which are often accompanied by palpitations and anxiety. Each particular episode lasts 3-10 minutes and episodes can recur with varying frequency. Some women experience hot flashes hourly or daily, whereas for others they may occur occasionally. The age at onset of hot flashes varies from woman to women. Night sweats are hot flashes that occur with heavy perspiration during sleep and cause sleep disruption.

VMS are experienced by majority of women during the menopausal transition. In the Study of Women's Health Across the Nation called SWAN, which is a multisite, longitudinal, cohort study of the menopausal transition being conducted in community-based group of women, 60-80% of women experience VMS at some point during the menopausal transition, with prevalence rates varying by racial/ethnic group [2]. Result of SWAN indicates that the occurrence and frequency of VMS peak in the late perimenopause and early postmenopausal years, or the several years surrounding the final menstrual period. Also, the recent epidemiological evidence indicates that hot flashes are experienced by 30-70% of premenopausal women [3], but they are likely to be mild in nature at these earlier stages of a women's reproductive life.

VMS may persist for a median of 7.4 years, and one third of women continue to experience moderate to severe VMS for 9.4 years after their final menstrual period [4]. While completely natural, VMS can significantly impact a women's life, affecting her

both physically, emotionally, and sometimes economically. VMS interferes with sleep, concentration/memory, cognitive function, productivity, and personal relationships, contributing to decreased quality of life [5, 6]. VMS increases healthcare resource utilization and direct/indirect costs [7], and compromises workplace productivity [8, 9].

In 1991, the National Heart, Lung, and Blood Institute, part of NIH, launched the Women's Health Initiative (WHI) to better understand how cardiovascular disease, breast cancer, colorectal cancer, and osteoporosis affect postmenopausal women and to reduce the number of women who develop and die from these diseases. More than 160,000 postmenopausal women ages 50 to 79 participated in the 15-year study, making it one of the largest prevention studies involving women in the United States. WHI results in 2002 found that postmenopausal women taking combination (estrogen and progestin) hormone therapy (HT) for menopause symptoms had an increased risk for breast cancer, heart disease, stroke, bold clots, and urinary incontinence. Although women using combined HT had a lower risk of fractures and colorectal cancer, these benefits did not outweigh the risks. As a result, many women stopped taking hormone therapy, reducing their risk for breast cancer [10].

Based on these historical facts, despite the high prevalence of menopause-related VMS, the current treatment options are limited for managing these symptoms. HT remains the gold standard for symptom relief with current treatment guidelines recommending the lowest effective dose and the shortest duration that is consistent with treatment goals and risks for the individual woman [11]. However, many women choose not to use HT [12, 13]. There is only one nonhormonal pharmacologic therapy, a selective serotonin reuptake inhibitor (SSRI), that is approved by the United States Food and Drug Administration (FDA) for the treatment of moderate to severe VMS. Other nonhormonal pharmacologic

therapies have been investigated; however, their efficacy and side-effect profile have limited their use for the treatment of VMS. In addition, some patients cannot take HT due to comorbid conditions (e.g., prior history of breast cancer) or by choice which varies in regions and countries. These limitations have led clinicians to search for other treatment options for hot flashes. Surprisingly, despite the extensive research, the pathophysiology of hot flashes is not entirely understood and likely to present the interplay between multiple central and peripheral physiologic systems [14]. Recently, through a series of pioneering studies conducted in humans and in animals, investigators suggested that the kisspeptin/neurokinin B/dynorphin (KNDy) neurons in the hypothalamus are the link between estrogen decline and hot flashes [15], and further investigation in research and clinical is expected. Given the impact and disruption to women's lives, there is a huge unmet need for improved management of VMS.

We strongly believe that development of the new therapies for VMS as one of the most bothersome symptoms in women's health will be more and more required by the entire society. The purpose of our research was to find new insights about the factors impacting the placebo response, which is critical for the success of a clinical trial. Through this research, we will be able to contribute to the women's health by enhancing the development of new therapies with effective and good safety profile.

2. Research 1

2.1 Background and Objectives

When developing a new drug, placebo-controlled studies are usually conducted, in which superiority of efficacy and sufficient safety for the new treatment is expected to be shown. At the same time, placebo response in clinical studies for treatment of hot flashes creates hurdles in understanding the true treatment effect of potential therapies. Previous studies have reported that treatment of menopausal women with placebo alone reduced hot flash frequency by up to 60% [16, 17]. Although several factors have been associated with this high placebo response [16], such as high anxiety at study entry, demographic and personal characteristics of participants, there has not been a systematic, thorough analysis of factors influencing the placebo response to inform the study designs and operational strategies.

As there are no good objective parameters, hot flash frequency is considered as one of the most important clinical outcomes for evaluating efficacy of treatments for VMS. Some potential new therapies are under development in accordance with regulatory guidance [18, 19]. These therapies must demonstrate adequate efficacy and safety in controlled clinical studies for regulatory approval. High placebo response due to intrinsic or extrinsic factors could negate the ability to understand the potential power of those therapies. Many studies have been conducted to seek the reasons of high placebo effect within a study or among studies; however, there is no definite answer as of now. Therefore, we conducted the research 1 with the aim of identifying factors associated with high magnitude of placebo response in randomized, controlled, double-blind studies enrolling women with hot flashes.

2.2. Methods

2.2.1. Search Strategy

We searched Medline, Embase, and BIOSIS Previews through ProQuest™, using the following search terms: “hot flash/hot flashes/hot flush/hot flushes”, “menopause/climacteric/konenki (Japanese term of climacteric status)”, “breast cancer”, “placebo”, “randomized/randomize/randomised/randomise” and “blind/blinded”. The search language was limited to English.

2.2.2. Selection Criteria

We used the following inclusion criteria: 1) randomised control studies (RCTs), 2) studies enrolling women with hot flashes, 3) studies evaluating changes in hot flash frequency, and 4) studies published in English. After the identification of target papers, we excluded reviews, abstracts, editorials, letters to the editor, preliminary reports, papers published year before 2000, and non-English articles. Studies reporting only changes in the percentage of hot flash frequency were also excluded.

The identification of relevant abstracts, the selection of clinical studies based on the criteria, and the subsequent data abstraction from full-text articles were confirmed at each step in duplicate. Any discrepancies were resolved by consensus.

2.2.3. Data Extraction

We evaluated the data of the placebo arm in each eligible study that clearly described the change in the mean number of hot flash frequency from baseline with standard deviation and/or standard error. Night sweats were also counted as hot flashes unless no

specific restriction was made in each study. Based on the study design, the extracted variables were classified as; A) categorical data: active treatment type, administration route, study region in/excluded the United States of America (US), breast cancer population in/excluded, entry criteria of hot flash severity, parallel/crossover study, placebo run-in period before treatment, and menopausal status, and B) numerical data: published year, pretreatment period duration, treatment period duration, number of sites, number of total participants, number of placebo participants, number of treatment arms, mean age, BMI, and hot flash frequency at baseline. The longest period evaluated in each study was selected as the treatment period.

2.2.4. Data Analysis

Placebo response was defined as a reduction from baseline in the mean number of hot flash frequency per week at the evaluated point. First, we conducted univariate analysis using the placebo response and the 18 explanatory variables. A significant association was defined as p-value < 0.2 in the univariate analysis to collect broad variables, and all the associated variables were incorporated into the multivariable model. Before conducting the multivariate analysis, correlation between the variables was examined using one of the following methods: Spearman's rank correlation (more than 0.7) for numeric/numeric factors, correlation ratio (more than 0.5) for numeric/categorical factors, and Cramér's V (more than 0.5) for categorical/categorical factors. The variables with a strong correlation were eliminated. In the multivariate analysis, a statistically significant association was defined as p-value < 0.05. We used a random-effects model to account for heterogeneity among studies. No interaction test was performed. All analyses were performed using R Ver 3.3.2 [20].

2.3. Results

2.3.1. Included studies for the analysis

A total of 802 articles from MEDLINE, Embase, and BIOSIS Preview were identified after the removal of 12 duplicates. After study selection, 268 articles were identified by applying the exclusion criteria. Out of the total 268 articles, 225 were excluded owing to insufficient data related to changes in the hot flash frequency in the placebo arm. Finally, 43 articles [21-63], including 45 RCTs, were eligible for the present meta-analysis (Figure 1). The validity of the included studies was evaluated using the criteria of the Cochrane risk-of-bias tool [64], which categorized the studies into low risk of bias, some concerns, and high risk of bias (Figure 2). The main characteristics of the selected RCTs are described in Table 1.

This meta-analysis included 5704 women with hot flashes from 45 RCTs. The data in this group were summarized separately for A) categorical and B) numerical data, in Table 2-1 and 2-2, respectively. Several types of HT (e.g., estrogen, combination of HTs) were defined as “HT” and used in 13 studies (28.9%). Supplements or plant-derived ingredients were defined as “CAM” and used in 13 studies (28.9%). Drugs expected to be used for treatment of hot flashes but not classified as HT nor CAM were defined as “non-HT” and used in 19 studies (42.2%). Most of the treatments were orally administered (88.9%). Over half of the studies had sites in the US (57.8%). More than 80% studies excluded breast cancer population (82.2%). Half of the studies set the entry criteria for moderate to severe hot flash severity according to the FDA guidance [18] (51.1%). Most of the studies were designed as parallel arm design (93.3%). Most of the studies did not utilize a placebo run-in period before treatment (91.1%). Regarding the target menopausal status, 32

studies (71.1%) were postmenopausal women only, whereas 9 studies (20.0%) included perimenopausal women and 4 studies (8.9%) included premenopausal women in addition to postmenopausal women. The median values of selected numerical data were shown in the bracket: publication year (2011), pretreatment period (2.0 weeks), treatment period (12 weeks), number of sites (9.0), number of total participants (205), number of placebo participants (61), number of treatment arms (2), mean age (53.7 years), BMI (26.5 kg/m^2), hot flash frequency at baseline (67.2 per week).

The forest plot including the placebo response and 95% CI in each study was shown in Figure 3.

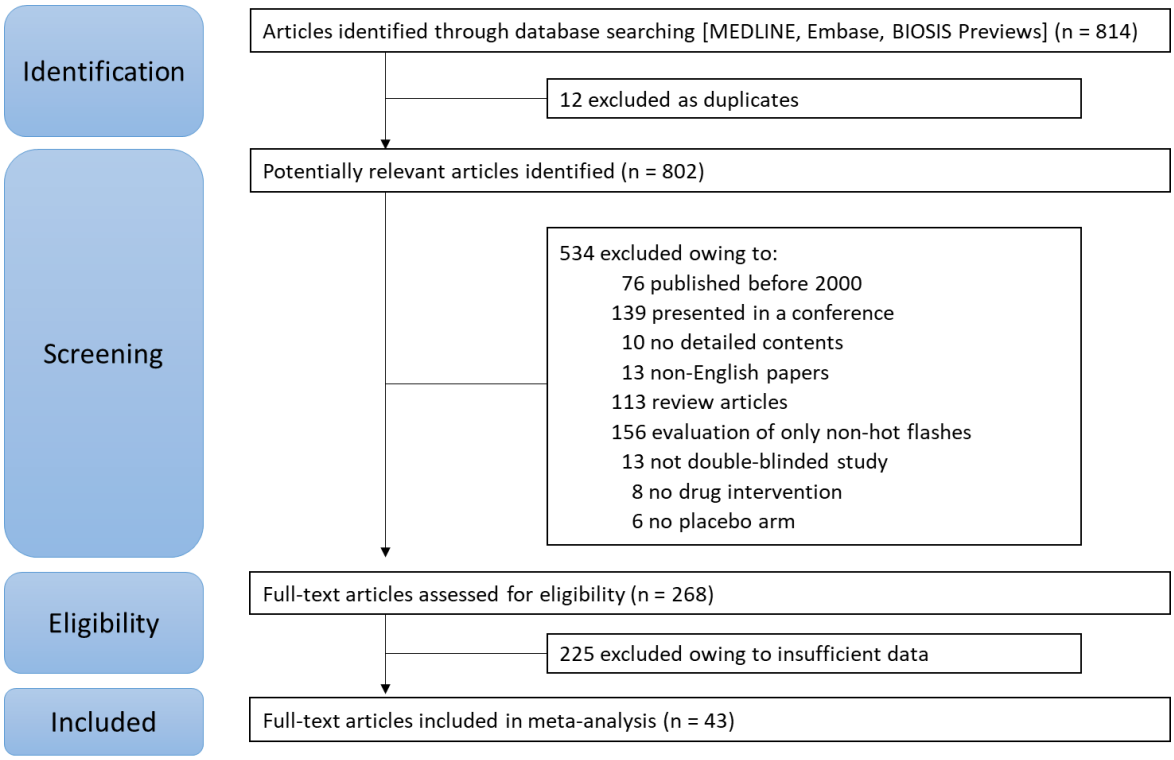


Figure 1. PRISMA flowchart for the selection of studies

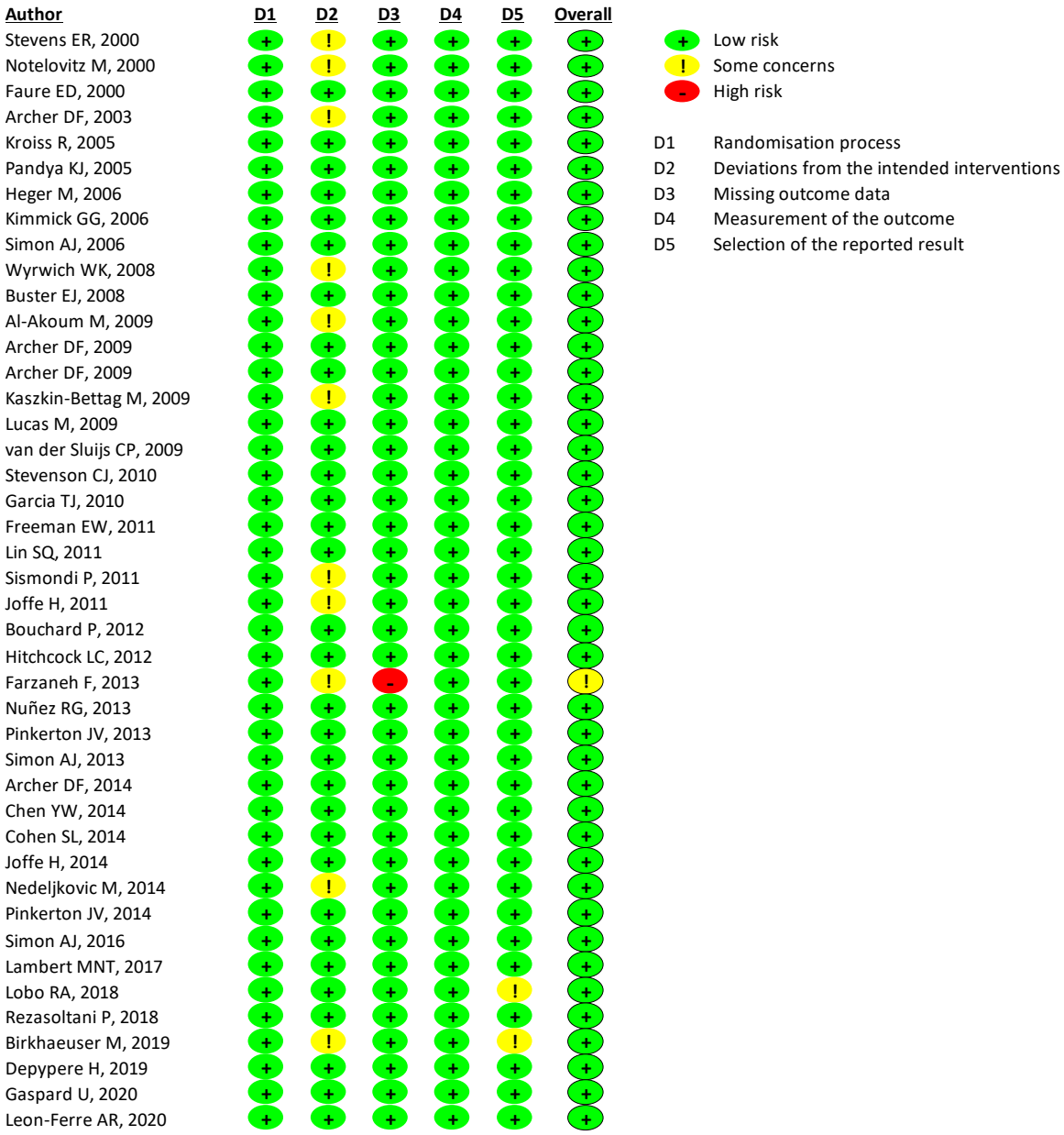


Figure 2. Summary of risk of bias in the included studies

Table 1. Characteristics of included randomized placebo-controlled trials

Published Year and Author	Active Treatment	Study Region	Characteristics of Participants in Placebo Arm					
			Number of Participants (Placebo/Total)	Mean Age (year)	Hot Flash Frequency at Baseline (per week)	Change from Baseline (per week)	BMI (kg/m^2)	Other Characteristics
2000 Stevens [21]	Synthetic conjugated Estrogens	US	47/120	48	94.1	-56.3	28.0	
2000 Notelovitz [22]	E2 (50 mg) plus norethindrone (140 mg) E2 (50 mg) plus norethindrone (250 mg) E2 (50 mg) plus norethindrone (400 mg)	US	53/220	53.6	81.1	-38.8	27.4	Placebo run-in
2002 Faure [23]	Phytosoya	France	21/75	53.9	65.8	-23.1	24.9	
2003 Archer [24]	Estradiol gel 1.25 g Estradiol gel 2.5 g	US	73/221	51	77.0	-39.9	26.4	
2005 Kroiss [25]	Tibolone 2.5 mg	Europe	26/70	59	15.4	6.3	25.4	BC included
2005 Pandya [26]	Gabapentin 300 mg Gabapentin 900 mg	US	113/420	54	61.6	-15.8	N/A	BC included
2006 Heger [27]	ERr731 (extract from the roots of Rheum rhaponticum)	Ukraine	42/110	48.6	103.6	0.0	25.7	
2006 Kimmick [28]	Sertraline 50 mg	US	22/62	52.3	38.5	-10.5	N/A	BC included

Published Year and Author	Active Treatment	Study Region	Characteristics of Participants in Placebo Arm					
			Number of Participants (Placebo/Total)	Mean Age (year)	Hot Flash Frequency at Baseline (per week)	Change from Baseline (per week)	BMI (kg/m^2)	Other Characteristics
2006 Simon [29]	Estradiol	US	100/200	51.8	95.2	-50.4	N/A	Placebo run-in
2008 Wyrwich [30]	Desvenlafaxine 50 mg Desvenlafaxine 100 mg Desvenlafaxine 150 mg Desvenlafaxine 200 mg	US	67/620	54.2	77.1	-42.3	26.7	
2008 Buster [31]	Estradiol spray	US	228/458	52.3	86.8	-37.2	26.9	
2009 Al-Akoum [32]	St. John's wort	Canada	22/47	54.0	53.9	-7.0	26.1	BC included
2009 Archer [33]	Desvenlafaxine 100 mg Desvenlafaxine 150 mg	US	150/458	53.4	76.3	-40.6	28.2	
2009 Archer [34]	Desvenlafaxine 100 mg Desvenlafaxine 150 mg	US	178/567	54.0	74.2	-34.3	26.3	
2009 Kaszkin-Bettag [35]	ERr731 (extract from the roots of Rheum rhaponticum)	Ukraine	49/112	49.6	84.7	-4.9	26.4	
2009 Lucas [36]	Omega-3 (ethyl-eicosapentaenoic acid)	Canada	46/120	50.2	16.3	-3.5	25.6	
2009 Slujis [37]	Chinese herbs	Australia	46/93	55.7	65.7	-19.7	26.1	

Published Year and Author	Active Treatment	Study Region	Characteristics of Participants in Placebo Arm					
			Number of Participants (Placebo/Total)	Mean Age (year)	Hot Flash Frequency at Baseline (per week)	Change from Baseline (per week)	BMI (kg/m^2)	Other Characteristics
2010 Stevenson [38]	E2 0.5mg/D2.5mg (dydrogesterone) E2 1mg/D5mg (dydrogesterone)	Europe	124/313	53.8	53.9	-34.3	26.6	
2010 Garcia [39]	Nutrafem	Singapore/ Philippines	28/159	54.7	22.6	-6.1	22.9	
2011 Freeman [40]	Escitalopram	US	97/205	54.4	67.6	-22.4	29.7	
2011 Lin [41]	Drospirenone/E2	China	61/249	51.9	50.3	-27.5	22.4	
2011 Sismondi [42]	Tibolone 2.5mg	Global	1290/3133	52.9	45.1	-12.4	27.1	BC included
2011 Joffe [43]	Estradiol	US	14/72	52.6	37.8	-15.4	N/A	
2012 Bouchard [44]	Desvenlafaxine 100mg	Europe/South Africa/Mexico	150/485	54.0	67.2	-40.7	26.0	
2012 Hitchcock [45]	Progesterone	Canada	46/133	54.4	44.1	-9.8	24.9	
2013 Farzaneh [46]	Primrose oil	Iran	28/56	51.9	37.8	-11.2	N/A	
2013 Nunez [47]	Bupropion	Brazil	47/55	49.0	47.0	-14.8	N/A	BC included
2013 Pinkerton [48]	Desvenlafaxine 100mg	US/Canada	181/396	54.0	83.3	-31.5	26.5	

Published Year and Author	Active Treatment	Study Region	Characteristics of Participants in Placebo Arm					
			Number of Participants (Placebo/Total)	Mean Age (year)	Hot Flash Frequency at Baseline (per week)	Change from Baseline (per week)	BMI (kg/m^2)	Other Characteristics
2013 Simon [49]	Paroxetine 7.5mg	US	305/614	53.0	81.6	-37.3	29.0	Placebo run-in
2013 Simon [49]	Paroxetine 7.5mg	US	284/570	54.0	76.3	-27.6	27.7	Placebo run-in
2014 Archer [50]	Drospirenone 0.25/E2 0.5mg	US	176/735	53.4	N/A	-31.9	27.8	
2014 Chen [51]	Melatonin 3 mg	US	19/95	59.0	18.9	-4.2	25.0	BC included
2014 Cohen [52]	Omega-3	US	169/355	55.0	53.2	-18.9	27.1	
2014 Joffe [53]	Estradiol 0.5mg Venlafaxine 75mg	US	137/339	54.3	53.9	-15.4	27.6	
2014 Nedeljkovic [54]	Chinese herbal medicine (Zhi Mu 14)	Switzerland	9/40	53.4	54.8	3.2	23.0	
2014 Pinkerton [55]	Gabapentin	US	294/600	54.0	82.6	-45.5	N/A	
2016 Simon [56]	Oxybutynin	US	73/148	54.1	75.9	-32.8	27.2	
2017 Lambert [57]	Red Clover Isoflavone	Denmark	29/62	52.3	111.5	-5.5	25.5	
2018 Lobo [58]	E2 0.25/P50mg E2 0.5/P50mg E2 0.5/P100mg E2 1/P100mg	US	115/1845	54.5	72.4	-40.2	26.6	

Published Year and Author	Active Treatment	Study Region	Characteristics of Participants in Placebo Arm					
			Number of Participants (Placebo/Total)	Mean Age (year)	Hot Flash Frequency at Baseline (per week)	Change from Baseline (per week)	BMI (kg/m^2)	Other Characteristics
2018 Rezasoltani [59]	Vitamin E	Iran	35/83	52.0	N/A	-16.9	26.9	
2019 Birkhaeuser [60]	Esmirtazapine 2.25 mg Esmirtazapine 4.5 mg Esmirtazapine 9.0 mg Esmirtazapine 18.0 mg	Global	294/942	54.0	79.8	-34.2	25.6	
2019 Birkhaeuser [60]	Esmirtazapine 4.5 mg Esmirtazapine 9.0 mg Esmirtazapine 18.0 mg	Global	283/946	53.5	84.7	-29.1	25.7	
2019 Depypere [61]	Fezolinetant	Belgium	40/87	53.7	72.0	-35.3	26.5	
2020 Gaspard [62]	Estetrol (E4) 2.5 mg Estetrol (E4) 5 mg Estetrol (E4) 10 mg Estetrol (E4) 15 mg	Europe	55/260	53.7	65.9	-43.9	26.6	
2020 Leon-Ferre [63]	Oxybutynin 2.5 mg Oxybutynin 5 mg	US	38/150	58.2	67.2	-18.2	N/A	BC included

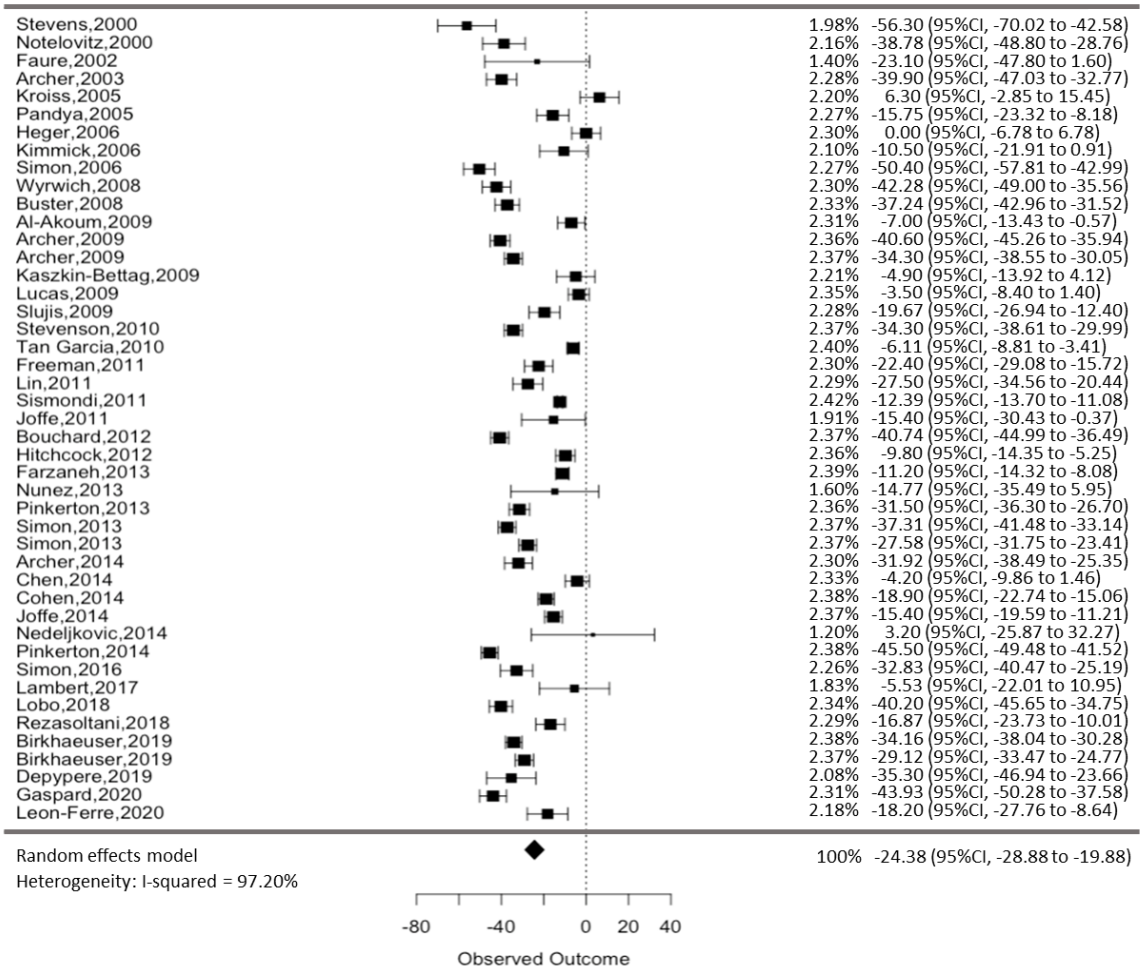


Figure 3. Forest plot of the effect size for placebo response; reduction of hot flash frequency

Table 2-1. Characteristics of the categorical data for studies included

Variables (A: Categorical Data)	Level	N (%)
Active Treatment Type	HT	13 (28.9)
	Non-HT	19 (42.2)
	CAM	13 (28.9)
Administration Route	Oral	40 (88.9)
	Non-Oral	5 (11.1)
Study Region In/Excluded the US	Included the US	26 (57.8)
	Excluded the US	19 (42.2)
Breast Cancer Population In/Excluded	Excluded	37 (82.2)
	Included	8 (17.8)
Entry criteria of hot flash severity	Moderate to Severe Only	23 (51.1)
	All Included	22 (48.9)
Parallel or Crossover Study	Parallel	42 (93.3)
	Crossover	3 (6.7)
Placebo Run-in Period before Treatment	No	41 (91.1)
	Yes	4 (8.9)
Menopausal Status	Postmenopausal only	32 (71.1)
	Include Perimenopausal	9 (20.0)
	Include Premenopausal	4 (8.9)

Abbreviations: CAM, complementary and alternative medicine; HT, hormone therapy; US, United States

Table 2-2. Characteristics of the numerical data for studies included

Variables (B: Numerical Data)	Median (Min. to Max.)	N
Published Year (year)	2011 (2000 to 2020)	45
Pretreatment Period Duration (week)	2.0 (1.0 to 8.5)	41
Treatment Period Duration (week)	12 (4 to 16)	45
Number of Sites	9.0 (1 to 245)	45
Number of Total Participants	205 (40 to 3133)	45
Number of Placebo Participants	61 (9 to 1290)	45
Number of Treatment Arms	2 (2 to 5)	45
Mean Age (year)	53.7 (48.0 to 59.0)	45
BMI (kg/m^2)	26.5 (22.4 to 29.7)	37
Hot Flash Frequency at Baseline (per week)	67.2 (12.4 to 111.5)	43

Abbreviations: BMI, body mass index; Max, maximum; Min, minimum; N, number of studies

2.3.2. Univariate analysis

Univariate regression analysis revealed that the following 13 variables (7 from categorical data and 6 from numerical data) were possibly related to placebo response ($p < 0.2$): active treatment type, administration route, study region in/excluded the US, breast cancer population in/excluded, entry criteria of hot flash severity, placebo run-in period before treatment, menopausal status, pretreatment period duration, treatment period duration, number of sites, number of treatment arms, BMI, and hot flash frequency at baseline (Table 3-1, 3-2).

Table 3-1. The result of univariate analysis in categorical data

Variables (A: Categorical Data)	Level	Estimate	SE	p-value
Active Treatment Type	HT	Reference		
	Non-HT	5.76	4.30	0.180*
	CAM	24.54	4.78	<0.001**
Administration Route	Oral	Reference		
	Non-Oral	-14.23	7.10	0.045**
Study Region In/Excluded the US	Included the US	Reference		
	Excluded the US	14.08	4.23	<0.001**
Breast Cancer Population In/Excluded	Excluded	Reference		
	Included	18.26	5.43	<0.001**
Entry criteria of hot flash severity	Moderate to Severe Only	Reference		
	All Included	26.52	2.08	<0.001**
Parallel or Crossover Study	Parallel	Reference		
	Crossover	10.95	9.64	0.256
Placebo Run-in Period before Treatment	No	Reference		
	Yes	-15.39	7.70	0.045**
Menopausal Status	Postmenopausal only	Reference		
	Include Perimenopausal	11.93	5.56	0.032*
	Include Premenopausal	12.91	8.09	0.111*

Abbreviations: CAM, complementary and alternative medicine; HT, hormone therapy; US, United States; SE, standard error

*p-value <0.2, **p-value <0.05

Table 3-2. The result of univariate analysis in numerical data

Variables (B: Numerical Data)	Estimate	SE	p-value
Published Year (year)	0.01	0.46	0.975
Pretreatment Period Duration (week)	-2.77	1.50	0.066*
Treatment Period Duration (week)	-1.50	0.76	0.047**
Number of Sites	-0.08	0.05	0.094*
Number of Total Participants	-0.01	0.00	0.207
Number of Placebo Participants	-0.01	0.01	0.525
Number of Treatment Arms	-6.18	2.01	0.002**
Mean Age (year)	0.91	1.06	0.387
BMI (kg/m^2)	-4.57	1.59	0.004**
Hot Flash Frequency at Baseline (per week)	-0.32	0.09	<0.001**

Abbreviations: BMI, body mass index; SE, standard error

*p-value <0.2, **p-value <0.05

2.3.3. Multivariate analysis

Based on the correlation analyses, the following nine variables were included in the multivariate regression analysis: active treatment type, breast cancer population in/excluded, placebo run-in period before treatment, pretreatment period duration, treatment period duration, number of sites, number of treatment arms, BMI, and hot flash frequency at baseline. Consequently, multivariate regression analysis revealed significant associations of placebo response with 1) active treatment type, 2) breast cancer population included, 3) treatment period duration, 4) number of treatment arms, and 5) BMI (Table 4). Three variables (treatment period duration, number of treatment arms, and BMI) were all associated with higher placebo response. Interestingly, both CAM as an active

treatment and inclusion of breast cancer patients demonstrated a lower placebo response compared to the other variables analyzed.

Table 4. The result of multivariate analysis

Variables	Level	Estimate	SE	p-value
Active Treatment Type	HT	Reference		
	Non-HT	4.27	3.86	0.270
	CAM	13.65	4.60	0.003**
Breast Cancer Population In/Excluded	Excluded	Reference		
	Included	17.10	5.82	0.003**
Placebo Run-in Period before Treatment	No	Reference		
	Yes	0.42	1.07	0.698
Pretreatment Period Duration (week)	N/A	0.42	1.07	0.698
Treatment Period Duration (week)	N/A	-2.81	0.94	0.003**
Number of Sites	N/A	-0.01	0.04	0.794
Number of Treatment Arms	N/A	-2.95	1.47	0.045**
BMI (kg/m^2)	N/A	-3.91	1.35	0.004**
Hot Flash Frequency at Baseline (per week)	N/A	-0.10	0.09	0.260

Abbreviations: BMI, body mass index; CAM, complementary and alternative medicine; HT, hormone therapy; SE, standard error

**p-value <0.05

2.4. Discussion

The results of our meta-analysis derived from 45 RCTs showed significant associations of placebo response with 1) active treatment type, 2) breast cancer population included, 3) treatment period duration, 4) number of treatment arms, and 5) BMI. The placebo response was lower in studies using CAM than in those using HT. In addition, the placebo response was lower in studies in which breast cancer population was included than in

those without. In contrast, a longer treatment period duration, greater number of treatment arms, and higher BMI were associated with higher placebo response.

With regard to the active treatment type, in addition to the multivariate regression analysis, the univariate regression analysis revealed that the placebo response tended to be low in studies using non-HT than those with HT. There are few non-HT/CAM that have been approved for VMS with clear efficacy and safety profiles. Based on these observations, we considered that the lower placebo response was possibly caused by the perception of weaker efficacy and/or lower expectation of the patients in the clinical studies with the target treatment.

Our research indicated a lower placebo response in studies including the breast cancer population than excluding this population. It has been suggested that hot flashes are artificially induced by the functional deterioration of the ovary or by the decrease in estrogen due to breast cancer treatment [65]. Most of the studies (37 out of 45) did not include patients who had history of breast cancer because the use of HT for cancer patients is prohibited in the prescribing information. We assumed that there was a large difference in patient profiles; thus, more studies on the association between hot flashes and breast cancer are needed. When investigating the association between hot flashes and breast cancer, researchers need to ensure that the active treatment for hot flashes should have little or no effect on the original cancer treatment.

The median of the treatment period duration in this dataset was 12 weeks (min. 4 weeks and max. 16 weeks). Our research showed that a longer treatment period duration led to a higher placebo response. This result is consistent with previous research [66]. Based on the result, we interpreted that checking the maintenance of efficacy by 12 weeks or longer period is important to confirm the true efficacy of the active treatment. Interestingly, the

pretreatment period duration or placebo run-in period before treatment (yes/no) did not affect the placebo response, but only the treatment period duration affected the placebo response as an individual factor. This meant that there is no need to set a pretreatment period and/or placebo run-in period, which has been considered to reduce the placebo response but would impose great burden on patients.

The median number of treatment arms in this dataset was two (min. two arms, max. five arms), including the placebo arm. Based on our research, the additional arms would lead to a higher placebo response. A higher number of treatment arms increases the possibility to receive active treatment. Therefore, this variable was considered to reflect the participants' expectation to obtain a higher chance of receiving active treatment.

The median BMI in this dataset was 26.5 kg/m^2 (min. 22.4 kg/m^2, max. 29.7 kg/m^2). In this range, the placebo response was becoming higher and higher. Prior research has shown that higher BMI is associated with hot flash frequency [67, 68], but recent evidence suggests that BMI has no effect on hot flash frequency [69]. Based on our findings, it would be prudent to consider excluding participants with extremely high or low BMI from clinical studies to minimize the placebo response. Minimally, researchers should consider a priori- defined analyses to understand how BMI variation within a population may impact overall trial results.

This meta-analysis identified factors associated with high placebo response in clinical studies of hot flashes. We focused on the hot flash frequency which can be measured as a certain number to see the effect of the treatment and most important endpoint for regulatory approvals.

3. Research 2

3.1. Background and Objectives

In the research 1, we identified the factors associated with placebo response, which was defined as the reduction in the mean number of hot flash frequency from baseline, in randomized, controlled, double-blind studies enrolling women with hot flashes. Even in the studies in which the same therapies and placebo were used, individual symptoms and treatment effect were different, and it would make the interpretation of the efficacy result more difficult.

There is no research focusing on the variability of placebo response as long as we confirmed, and we checked the tendency of the coefficient of variation (CV) in the two study groups; 1) studies in which the active treatment showed statistically significant efficacy compared to placebo, and 2) studies in which the active treatment did not show statistically significant efficacy compared to placebo. Interestingly, the CV in the group number 1 showed narrow CV degree in both active treatment and placebo arms, compared to the respective arms in the group number 2 (Appendix 1). Based on this pre-analysis, we hypothesized that the factors which could make less variability of CV would contribute to the success of the study. As we have focused on the placebo response, the objective of the research 2 was set to identify the factors associated with high variability of placebo response in randomized, controlled, double-blind studies enrolling women with hot flashes.

3.2. Methods

We used the same set of data with research 1. Given the participant selection criteria in each study were slightly different, we calculated the CV as the indicator for the relative variability of placebo response, instead of using standard deviation. The formula for calculating the CV is as follows: CV = (Standard Deviation / Mean) x 100. If the expected return in the denominator of the CV formula was negative or zero, we excluded those values from the analysis.

Once CV in each study was calculated, we conducted a simple linear regression analysis using the CV and the 18 explanatory variables. A significant association was defined as p-value < 0.2 in the simple linear regression analysis to collect broad variables. Before conducting the multiple regression analysis, correlation between the explanatory variables was examined using the same methods in the research 1, and the variables with a strong correlation were eliminated. However, we have included the item of "number of placebo participants" regardless of the result of the simple linear regression analysis, because the value of CV was anticipated to become smaller when the number of placebo participants would become larger. In the multiple regression analysis, a statistically significant association was defined as p-value < 0.05. All analyses were performed using R Ver 3.3.2 [20].

3.3. Results

3.3.1. Included studies for the analysis

Among the placebo data in 45 RCTs, there were 4 negative value of CV after the defined calculation. We eliminated those 4 negative values from the analysis to avoid any misleading.

3.3.2. Simple linear regression analysis

Simple linear regression analysis revealed that the following 10 variables (6 from categorical data and 4 from numerical data) were possibly related to the variability of placebo response ($p < 0.2$): active treatment type, study region in/excluded the US, breast cancer population in/excluded, entry criteria of hot flash severity, parallel or crossover study, menopausal status, pretreatment period duration, treatment period duration, number of treatment arms, and hot flash frequency at baseline (Table 5-1, 5-2).

Table 5-1. The result of simple linear regression analysis in categorical data

Variables (A: Categorical Data)	Level	Estimate	SE	p-value
Active Treatment Type	HT	Reference		
	Non-HT	0.39	0.42	0.364
	CAM	1.38	0.48	0.007**
Administration Route	Oral	Reference		
	Non-Oral	-0.52	0.59	0.386
Study Region In/Excluded the US	Included the US	Reference		
	Excluded the US	0.73	0.39	0.069*
Breast Cancer Population In/Excluded	Excluded	Reference		
	Included	1.33	0.48	0.008**
Entry criteria of hot flash severity	Moderate to Severe Only	Reference		
	All Included	1.36	0.33	<0.001**
Parallel or Crossover Study	Parallel	Reference		
	Crossover	1.41	0.72	0.058*
Placebo Run-in Period before Treatment	No	Reference		
	Yes	-0.63	0.65	0.338
Menopausal Status	Postmenopausal only	Reference		
	Include Perimenopausal	1.05	0.47	0.031**
	Include Premenopausal	1.70	0.59	0.007**

Abbreviations: CAM, complementary and alternative medicine; HT, hormone therapy; US, United States; SE, standard error

*p-value <0.2, **p-value <0.05

Table 5-2. The result of simple linear regression analysis in numerical data

Variables (B: Numerical Data)	Estimate	SE	p-value
Published Year (year)	-0.03	0.04	0.433
Pretreatment Period Duration (week)	-0.23	0.10	0.026**
Treatment Period Duration (week)	-0.09	0.07	0.164*
Number of Sites	-0.01	0.00	0.211
Number of Total Participants	0.00	0.00	0.277
Number of Placebo Participants	0.00	0.00	0.552
Number of Treatment Arms	-0.38	0.18	0.043**
Mean Age (year)	-0.11	0.09	0.230
BMI (kg/m^2)	-0.12	0.15	0.424
Hot Flash Frequency at Baseline (per week)	-0.02	0.01	0.037**

Abbreviations: BMI, body mass index; SE, standard error

*p-value <0.2, **p-value <0.05

3.3.3. Multiple regression analysis

Based on the correlation analyses, the following six variables were included in the multiple regression analysis: active treatment type, menopausal status, pretreatment period duration, number of placebo participants, number of treatment arms, and hot flash frequency at baseline. Consequently, multiple regression analysis revealed significant associations of the variability of placebo response with 1) menopausal status including premenopausal, and 2) hot flash frequency at baseline (Table 6). One variable (menopausal status include premenopausal) was associated with higher variability of placebo response, and another variable (hot flash frequency at baseline) was associated with lower variability of placebo response.

Table 6. The result of multiple regression analysis

Variables	Level	Estimate	SE	p-value
Active Treatment Type	HT	Reference		
	Non-HT	0.06	0.32	0.850
	CAM	0.70	0.43	0.113
Menopausal Status	Postmenopausal only	Reference		
	Include Perimenopausal	0.04	0.36	0.905
	Include Premenopausal	1.50	0.47	0.004**
Pretreatment Period Duration (week)	N/A	-0.03	0.09	0.755
Number of Placebo Participants	N/A	0.00*	0.00*	0.712
Number of Treatment Arms	N/A	-0.08	0.12	0.533
Hot Flash Frequency at Baseline (per week)	N/A	-0.02	0.01	0.046**

Abbreviations: CAM, complementary and alternative medicine; HT, hormone therapy; SE, standard error

*nearly equal to 0.00 after rounding up, **p-value <0.05

3.4. Discussion

The results of our multiple regression analysis targeted the placebo arm data showed significant associations of CV with 1) menopausal status include premenopausal, and 2) hot flash frequency at baseline, as an individual factor.

The CV was larger in studies in which premenopausal women were allowed to be included compared to those in which they were excluded. In general, VMS peak in the late perimenopause and early postmenopausal years, or the several years surrounding the final menstrual period. It would be high possible that the causes of VMS in the premenopausal women were derived from the background disease rather than natural aging (e.g., hysterectomy and/or oophorectomy). Although the number of studies with premenopausal women were very low, our result suggested that menopausal status would be very important to keep the homogeneity in a clinical study for VMS.

The CV was smaller in studies including higher number of hot flash frequency at baseline. The FDA strongly recommends enrolling only postmenopausal women who have a minimum of 7 to 8 moderate to severe hot flashes per day, or 50 to 60 per week at baseline. The median of hot flash frequency at baseline among the collected studies was 67.2 (min-max: 12.4-111.5) per week, but some studies without having such a minimum hot flash frequency nor severity requirement were also included. This suggested that a certain number of hot flash frequency at baseline would increase the success rate of the study by decreasing the CV.

The CV represents the variability of efficacy, meaning the variability of numeric change of hot flash frequency from baseline. Interestingly, both factors impacting the CV were those related to the subjects, so they are manageable if the researchers would carefully examine the enrollment criteria. Based on our findings, it would be advisable to

consider enrolling participants according to the recommendation by the FDA, at least for the criteria of menopausal status and baseline hot flash frequency.

4. Overall Discussion

HT is the gold standard for relieving hot flashes, and the efficacy of HT has been well recognized. At the same time, some treatment options (e.g., Premarin®) include a boxed warning for associated risks of endometrial cancer, cardiovascular disorders, breast cancer, and probable dementia. To overcome these limitations, several potential new therapies are currently under development; however, these alternative therapies are facing challenges in showing sufficient efficacy as well as good safety profiles.

Clinical studies of hot flashes generally follow the US Guidance for Industry [18] and/or the EMA Guideline [19]. RCTs with a placebo arm and four co-primary endpoints (hot flash frequency and severity at 4 and 12 weeks) are recommended in the US guidance [18]. Demonstrating significant differences in all these four co-primary endpoints is considered as a high hurdle for a newly developed therapy to obtain regulatory approval in the US. In addition to show a statistical significance between the active treatment and placebo, it is important to show a clinically meaningful difference in reduction of hot flash frequency compared to the placebo arm [70]. In order to reveal any benefits of existing therapies in the clinical setting, there have been many systematic reviews and meta-analyses comparing the efficacy of several active treatments as well as studies attempting to determine the factors affecting the high placebo response [17, 71]. A previous study by Li et al. targeted the percentage of change in hot flash frequency from baseline in the placebo arm and confirmed that the active treatment type was a covariate for the placebo response [66]. However, their results did not answer the question of what factors should be considered, except for the type of test drug in the actual clinical studies. Therefore, we investigated the factors associated with the placebo response exhaustively, which was defined as “reduction in the mean number of hot flash frequency per week at

the evaluation point from baseline". To the best of our knowledge, the research 1is the first meta-analysis assessing the variables associated with the placebo response on hot flash frequency in women using all available data from RCTs. In addition, we expanded our research into the variability of placebo response. As long as we confirmed, the research 2 is the first attempt assessing the variability of placebo response within the broad RCTs.

In the research 1, multivariate regression analysis confirmed two factors impacting a significant low placebo response: active treatment type and breast cancer population included, then three factors impacting a significant high placebo response: treatment period duration, number of treatment arms, and BMI. In the univariate regression analysis, some factors were also tended to be associated with the change of hot flashes from baseline. However, these factors were not included in the multivariate regression analysis due to the high correlations in Cramér's V. In the research 2, multiple regression analysis confirmed that clinical studies including premenopausal women compared to those with only postmenopausal women were associated with larger CV of placebo response, and hot flash frequency at baseline was associated with smaller CV. Both were factors related to the background information of subjects, not the study design. In the simple linear regression analysis, breast cancer population in/excluded and entry criteria of hot flash severity were also tended to be associated with the degree of CV. These factors were not included in the multiple regression analysis due to the high correlation in Cramér's V with the menopausal status.

When planning a clinical study targeting VMS with new/existing therapy, it would be critical to collect and analyze the information required in the targeting countries referring to the guidelines and the environment. Appropriate endpoints related to efficacy and

safety need to be set, and importantly, improvement of quality of life such as sleep is also critical for VMS treatment. Based on these premises, we would like to recommend implementing our research results to improve the success probability of the study by controlling placebo response: 1) no need for setting the pre-treatment period that would be the burden for participants, 2) minimize total treatment arms including placebo, 3) consider to evaluate efficacy not only for short period at 4 weeks but also for middle period at 12 weeks, 4) exclude participants with extreme background value of BMI, 5) plan a separate study for patients with extrinsic cause of VMS to keep the study homogeneous as much as possible.

In the real clinical settings, it is considered that improving both frequency and severity of hot flash are important to patients. In addition, improvement of quality of life such as sleep is also critical. Those important endpoints are all objectively evaluated by participants themselves, so any positive or negative information related to the active treatment would naturally impact their evaluation as a bias. We found out some important factors related to the placebo response, but still further research is needed because evaluation by study participants would be impacted by their real-time consideration. The COMMA (Core Outcomes in Menopause) group have recently clarified which aspects of vasomotor symptoms should be measured in clinical studies [72]. These consolidated research as well as our findings will enable improved standardization of study design to assess the treatment appropriately.

5. Limitation

There are several limitations in this research. First, the factors evaluated in this research were limited. Unknown or unidentified variables may have contributed to the placebo response, although every effort to eliminate placebo response were made in each clinical study by the specific in/exclusion criteria. We prioritized general factors available in the majority of RCTs. Second, our research included only RCTs that have been published. There may be studies which were not published due to the negative data in the active treatment arm. Third, the number of RCTs identified was not large because we focused on placebo-controlled studies meeting our criteria for examining hot flash frequency. Although frequency is one of the important aspects for patients, severity and quality of life should be considered in the clinical setting. Lastly, most of the RCTs identified did not include Asian women. Since the difference between races was not investigated, the results may not be generalizable across all populations. Therefore, the results of this research should be regarded as exploratory rather than conclusive.

6. Conclusion

The results of our study showed that three independent factors were associated with larger placebo response and two independent factors were associated with smaller placebo response in clinical studies of hot flashes. In addition, one factor was shown to be associated with larger CV of placebo response and another factor was associated with smaller CV.

Hot flash frequency is one of the most important endpoints for evaluating the efficacy of treatments of hot flashes, and the drugs under development need to show significant difference in efficacy compared with placebo. Therefore, it is important to implement the learning from this research when designing interventional studies for VMS.

References

1. 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(9): 924-932.
2. Gold E, Colvin A, Avis N, et al. Longitudinal analysis of vasomotor symptoms and race/ethnicity across the menopausal transition: Study of Women's Health Across the Nation (SWAN). *Am J Public Health* 2006;96(7):1226–1235.
3. Reed SD, Lampe JW, Qu C, et al. Premenopausal vasomotor symptoms in an ethnically diverse population. *Menopause* 2014;21(2):153–158.
4. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. *JAMA Intern Med.* 2015;175(4):531-539.
5. 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(7):732-739.
6. Utian WH. Psychosocial and socioeconomic burden of vasomotor symptoms in menopause: a comprehensive review. *Health Qual Life Outcomes* 2005;3:47.
7. Sarrel P, Portman D, Lefebvre P, et al. Incremental direct and indirect costs of untreated vasomotor symptoms. *Menopause* 2015;22(3):260-266.
8. Gartoulla P, Bell RJ, Worsley R, Davis SR. Moderate-severely bothersome vasomotor symptoms are associated with lowered psychological general wellbeing in women at midlife. *Maturitas* 2015;81(4):487-492.
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(5):518-524.

10. Website in Office on Women's Health: "Largest women's health prevention study ever – Women's Health Initiative". U.S. Department of Health & Human Services (HHS.gov).
11. The NAMS 2017 Hormone Therapy Position Statement Advisory Panel. The 2017 hormone therapy position statement of The North American Menopause Society. *Menopause* 2017;24(7):728-53.
12. Crawford SL, Crandall CJ, Derby CA, et al. Menopausal hormone therapy trends before versus after 2002: impact of the Women's Health Initiative Study results. *Menopause* 2019;26(6):588-97
13. Ameye L, Antoine C, Paesmans M, de Azambuja E, Rozenberg S. Menopausal hormone therapy use in 17 European countries during the last decade. *Maturitas* 2014;79(3):287-91. Erratum in: *Maturitas* 2015;81(1):237-8.
14. Thurston RC. Vasomotor symptoms. In: Crandall CJ, Bachman GA, Faubion SS, et al., eds. Menopause Practice: A Clinician's Guide. 6th ed. Pepper Pike, OH: *The North American Menopause Society*, 2019:43-55.
15. Rance NE, Dacks PA, Mittelman-Smith MA, Romanovsky AA, Krajewski-Hall SJ. Modulation of body temperature and LH secretion by hypothalamic KNDy (kisspeptin, neurokinin B and dynorphin) neurons: A novel hypothesis on the mechanism of hot flushes. *Front Neuroendocrinol* 2013;34(3):211-27.
16. Freeman EW, Ensrud KE, Larson JC, et al. Placebo improvement in pharmacologic treatment of menopausal hot flashes: time course, duration, and predictors. *Psychosom Med* 2015;77:167-175.

17. Margaret DD, Helena JT, Kerry MB, John ER, Henry GB. Predictors of placebo response in a randomized, controlled trial of phytotherapy in menopause. *Menopause* 2009;16:792-796.
18. Food and Drug Administration. 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: https://www.fda.gov/media/71359/download. Accessed Sep 1, 2020.
19. European Medicines Agency. Guideline on clinical investigation of medicinal products for hormone replacement therapy of oestrogen deficiency symptoms in postmenopausal women. Available at: https://www.ema.europa.eu/en/clinical-investigation-medicinal-products-hormone-replacement-therapy-oestrogen-deficiency-symptoms. Accessed Sep 1, 2020.
20. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
21. Stevens ER, Hanford K, Wason S, Cusack LS, Phelps VK. A 12-week clinical trial determining the efficacy of synthetic conjugated estrogens, A (SCE), in the treatment of vasomotor symptoms in menopausal women. *Int J Fertil Womens Med*. 2000;45:264-272.
22. Notelovitz M, Cassel D, Hille D, et al. Efficacy of continuous sequential transdermal estradiol and norethindrone acetate in relieving vasomotor symptoms associated with menopause. *Am J Obstet Gynecol* 2000;182:7-12.

23. Faure ED, Chantre P, Mares P. Effects of a standardized soy extract on hot flushes: a multicenter, double-blind, randomized, placebo-controlled study. *Menopause* 2002;9:329-334.
24. Archer DF, EstroGel Study Group. Percutaneous 17β-estradiol gel for the treatment of vasomotor symptoms in postmenopausal women. *Menopause* 2003;10:516-521.
25. Kroiss R, Fentiman IS, Helmond FA, et al. The effect of tibolone in postmenopausal women receiving tamoxifen after surgery for breast cancer: a randomised, double-blind, placebo-controlled trial. *BJOG* 2005;112:228-233.
26. Pandya KJ, Morrow GR, Roscoe JA, et al. Gabapentin for hot flashes in 420 women with breast cancer: a randomised double-blind placebo-controlled trial. *Lancet* 2005;366:818-824.
27. Heger M, Ventskovskiy MB, Borzenko I, et al. Efficacy and safety of a special extract of Rheum rhaponticum (ERr 731) in perimenopausal women with climacteric complaints: a 12-week randomized, double-blind, placebo-controlled trial. *Menopause* 2006:13:744-759.
28. Kimmick GG, Lovato J, McQuellon R, Robinson E, Muss BH. Randomized, double-blind, placebo-controlled, crossover study of sertraline (Zoloft) for the treatment of hot flashes in women with early stage breast cancer taking tamoxifen. *Breast J.* 2006;12:114-122.
29. Simon AJ, ESTRASORB Study Group. Estradiol in micellar nanoparticles: the efficacy and safety of a novel transdermal drug-delivery technology in the management of moderate to severe vasomotor symptoms. *Menopause* 2006;13:222-231.

30. Wyrwich WK, Spratt ID, Gass M, Yu H, Bobula DJ. Identifying meaningful differences in vasomotor symptoms among menopausal women. *Menopause* 2008;15:698-705.
31. Buster EJ, Koltun DW, Pascual GML, Day WW, Peterson C. Low-dose estradiol spray to treat vasomotor symptoms: a randomized controlled trial. *Obstetrics & Gynecology* 2008;111:1343-1351.
32. Al-Akoum M, Maunsell E, Verreault R, Provencher L, Otis H, Dodin S. Effects of Hypericum perforatum (St. John's wort) on hot flashes and quality of life in perimenopausal women: a randomized pilot trial. *Menopause* 2009;16:307-314.
33. Archer DF, Seidman L, Constantine DG, Pickar HJ, Olivier S. A double-blind, randomly assigned, placebo-controlled study of desvenlafaxine efficacy and safety for the treatment of vasomotor symptoms associated with menopause. *Am J Obstet Gynecol.* 2009;200:172.e1-10.
34. Archer DF, Dupont MC, Constantine DG, Pickar HJ, Olivier S. Desvenlafaxine for the treatment of vasomotor symptoms associated with menopause: a double-blind, randomized, placebo-controlled trial of efficacy and safety. *Am J Obstet Gynecol.* 2009;200:238.e1-10.
35. Kaszkin-Bettag M, Ventskovskiy MB, Solskyy S, et al. Confirmation of the efficacy of ERr 731 in perimenopausal women with menopausal symptoms. *Altern Ther Health Med.* 2009;15:24-34.
36. Lucas M, Asselin G, Mérette C, Poulin MJ, Dodin S. Effects of ethyl-eicosapentaenoic acid omega-3 fatty acid supplementation on hot flashes and quality of life among middle-aged women: a double-blind, placebo-controlled, randomized clinical trial. *Menopause* 2009;16:357-366.

37. van der Sluijs CP, Bensoussan A, Chang S, Baber R. A randomized placebo-controlled trial on the effectiveness of an herbal formula to alleviate menopausal vasomotor symptoms. *Menopause* 2009;16:336-344.
38. Stevenson CJ, Durand G, Kahler E, Pertynski T. Oral ultra-low dose continuous combined hormone replacement therapy with 0.5 mg 17β-oestradiol and 2.5 mg dydrogesterone for the treatment of vasomotor symptoms: Results from a double-blind, controlled study. *Maturitas* 2010;67:227-232.
39. Garcia TJ, Gonzaga F, Tan D, Ng YT, Oei LP, Chan WBC. Use of a multibotanical (Nutrafem) for the relief of menopausal vasomotor symptoms: a double-blind, placebo-controlled study. *Menopause* 2010;17:303-308.
40. Freeman EW, Guthrie KA, Caan B, et al. Efficacy of escitalopram for hot flashes in healthy menopausal women: a randomized controlled trial. *JAMA* 2011;305:267-274.
41. Lin SQ, Sun LZ, Lin JF, et al. Estradiol 1 mg and drospirenone 2 mg as hormone replacement therapy in postmenopausal Chinese women. *Climacteric* 2011;14:472-481.
42. Sismondi P, Kimmig R, Kubista E, et al. Effects of tibolone on climacteric symptoms and quality of life in breast cancer patients: data from LIBERATE trial. *Maturitas* 2011;70:365-372.
43. Joffe H, Petrillo LF, Koukopoulos A, et al. Increased estradiol and improved sleep, but not hot flashes, predict enhanced mood during the menopausal transition. *J Clin Endocrinol Metab.* 2011;96:E1044-1054.
44. Bouchard P, Panay N, de Villiers TJ, Vincendon P, Bao W, Cheng RJ, Constantine G. Randomized placebo- and active-controlled study of desvenlafaxine for menopausal vasomotor symptoms. *Climacteric* 2012;15:12-20.

45. Hitchcock LC, Prior CJ. Oral micronized progesterone for vasomotor symptoms: a placebo-controlled randomized trial in healthy postmenopausal women. *Menopause* 2012;19:886-893.
46. Farzaneh F, Fatehi S, Sohrabi RM, Alizadeh K. The effect of oral evening primrose oil on menopausal hot flashes: a randomized clinical trial. *Arch Gynecol Obstet.* 2013;288:1075-1079.
47. Nuñez RG, Pinczowski H, Zanellato R, et al. Bupropion for control of hot flashes in breast cancer survivors: a prospective, double-blind, randomized, crossover, pilot phase II trial. *J Pain Symptom Manage.* 2013;45:969-979.
48. Pinkerton JV, Constantine G, Hwang E, Cheng RJ. Desvenlafaxine compared with placebo for treatment of menopausal vasomotor symptoms: a 12-week, multicenter, parallel-group, randomized, double-blind, placebo-controlled efficacy trial. *Menopause* 2013;20:28-37.
49. Simon AJ, Portman JD, Kaunitz MA, et al. Low-dose paroxetine 7.5 mg for menopausal vasomotor symptoms: two randomized controlled trials. *Menopause* 2013;20:1027-1035.
50. Archer DF, Thomas S, Schaefers M, Gerlinger C, Gude K. A randomized, double-blind, placebo-controlled study of the lowest effective dose of drospirenone with 17β-estradiol for moderate to severe vasomotor symptoms in postmenopausal women. *Menopause* 2014;21:227-235.
51. Chen YW, Giobbie-Hurder A, Gantman K, et al. A randomized, placebo-controlled trial of melatonin on breast cancer survivors: impact on sleep, mood, and hot flashes. *Breast Cancer Res Treat.* 2014;145:381-388.

52. Cohen SL, Joffe H, Guthrie KA, et al. Efficacy of omega-3 for vasomotor symptoms treatment: a randomized controlled trial. *Menopause* 2014;21:347-354.
53. Joffe H, Guthrie KA, LaCroix ZA, et al. Low-dose estradiol and the serotonin-norepinephrine reuptake inhibitor venlafaxine for vasomotor symptoms: a randomized clinical trial. *JAMA Intern Med.* 2014;174:1058-1066.
54. Nedeljkovic M, Tian L, Ji P, et al. Effects of acupuncture and Chinese herbal medicine (Zhi Mu 14) on hot flushes and quality of life in postmenopausal women: results of a four-arm randomized controlled pilot trial. *Menopause* 2014;21:15-24.
55. Pinkerton JV, Kagan R, Portman D, Sathyanarayana R, Sweeney M. Phase 3 randomized controlled study of gastroretentive gabapentin for the treatment of moderate-to-severe hot flashes in menopause. *Menopause* 2014;21:567-573.
56. Simon AJ, Gaines T, LaGuardia DK. Extended-release oxybutynin therapy for vasomotor symptoms in women: a randomized clinical trial. *Menopause* 2016;23:1214-1221.
57. Lambert MNT, Thorup AC, Hansen ESS, Jeppesen PB. Combined Red Clover isoflavones and probiotics potently reduce menopausal vasomotor symptoms. *PLoS One* 2017;12:e0176590.
58. Lobo RA, Archer DF, Kagan R, et al. A 17β-estradiol-progesterone oral capsule for vasomotor symptoms in postmenopausal women: a randomized controlled trial. *Obstet Gynecol.* 2018;132:161-170.
59. Rezasoltani P, Elliyoun N, Ziaie T, Leyli KE, Aski SK, Sobhani A. Effect of vitamin E supplementation on plasma nitric oxide in menopausal women with hot flashes: a cross-over, randomized clinical trial. *Iran Red Crescent Med J.* 2018;20:e68800.

60. Birkhaeuser M, Bitzer J, Braat S, Ramos Y. Esmirtazapine treatment of postmenopausal vasomotor symptoms: two randomized controlled trials. *Climacteric* 2019;22:312-322.

61. Depypere H, Timmerman D, Donders G, et al. Treatment of menopausal vasomotor symptoms with fezolinetant, a neurokinin 3 receptor antagonist: a phase 2a trial. *J Clin Endocrinol Metab.* 2019;104:5893-5905.

62. Gaspard U, Taziaux M, Mawet M, et al. A multicenter, randomized study to select the minimum effective dose of estetrol (E4) in postmenopausal women (E4Relief): part 1. vasomotor symptoms and overall safety. *Menopause* 2020;27:848-857.

63. Leon-Ferre AR, Novotny JP, Wolfe GE, et al. Oxybutynin vs placebo for hot flashes in women with or without breast cancer: a randomized, double-blind clinical trial (ACCRU SC-1603). *JNCI Cancer Spectr.* 2019;4:pkz088.

64. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:I4898. Available at: https://www.bmj.com/content/366/bmj.l4898. Accessed Mar 1, 2021.

65. Avis NE. Breast Cancer Survivors and Hot Flashes: The Search for Nonhormonal Treatments. *J Clin Oncol.* 2008;26:5008-5010.

66. Li L, Xu L, Wu J, Dong L, Lv Y, Zheng Q. Quantitative analysis of placebo response and factors associated with menopausal hot flashes. *Menopause* 2017;24:932-937.

67. Shobeiri F, Jenabi E, Poorolajal J, Hazabehei SMM. The association between body mass index and hot flash in midlife women: a meta-analysis. *J Menopausal Med.* 2016;22:14-19.

68. Freedman RR. Hot flashes: behavioral treatments, mechanisms, and relation to sleep. *Am J Med.* 2005;118:124-130.

69. Karim R, Dang HM, Hodis HN, Stanczyk FZ, Brinton RD, Mack WJ. Association of hot flushes with ghrelin and adipokines in early late postmenopausal women versus late postmenopausal women. *Menopause* 2020;27:512-518.
70. Sarri G, Pedder H, Dias S, Guo Y, Lumsden MA. Vasomotor symptoms resulting from natural menopause: a systematic review and network meta-analysis of treatment effects from the National Institute for Health and Care Excellence guideline on menopause. *BJOG.* 2017;124:1514-1523.
71. Ziv-Gal A, Flaws JA. Factors that may influence the experience of hot flushes by healthy middle-aged women. *J Womens Health (Larchmt).* 2010;19:1905-1914.
72. Lensen S, Archer D, Bell R, et al. A core outcome set for vasomotor symptoms associated with menopause: the COMMA (Core Outcomes in Menopause) global initiative. *Menopause* 2021;28:852-858.

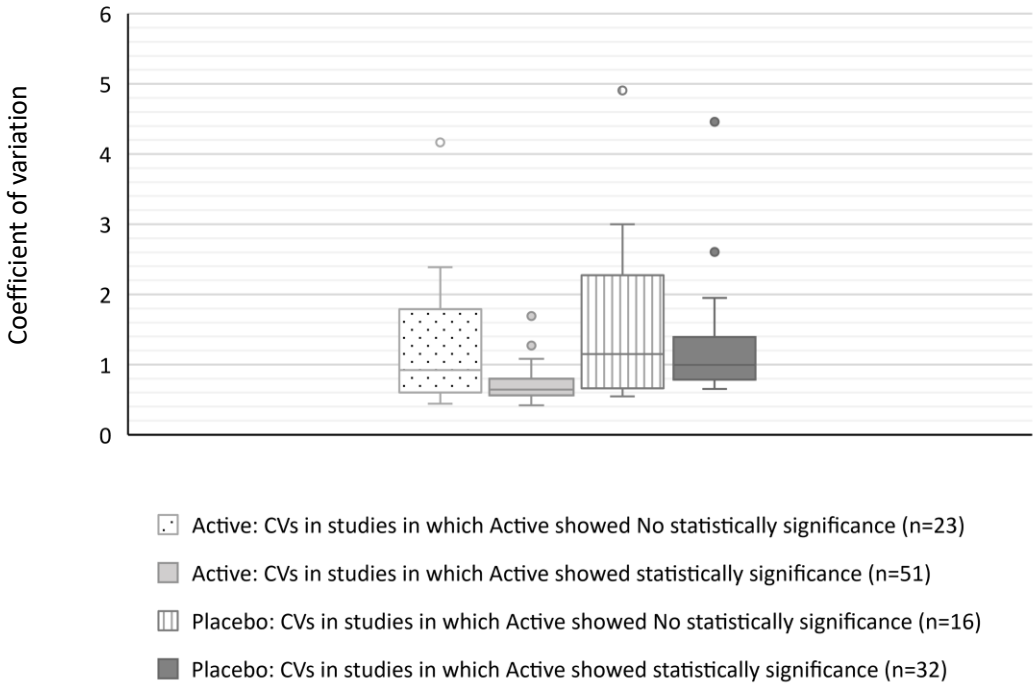
Acknowledgement

I would like to express my sincere gratitude to Professor Mamoru Narukawa for his dedicated and thoughtful guidance of my research. I would give my special thanks to Dr. Masayuki Kaneko for guiding me with insightful feedback and suggestions, especially from his professional statistical point of view. I also appreciate all members in our department for their warm encouragement. I would like to thank Professor Rie Kubota, Professor Nobuhiko Okada, and Professor Yusuke Suzuki for their reviews and helpful discussion on this study.

Not limited to Kitasato University, but I would like to thank many global colleagues in Astellas. Especially thanks to Mr. Kentaro Kuroishi who provided me with his statistical specialty, and Mrs. Marci English who supported me with her special and latest knowledge in women's health area including VMS.

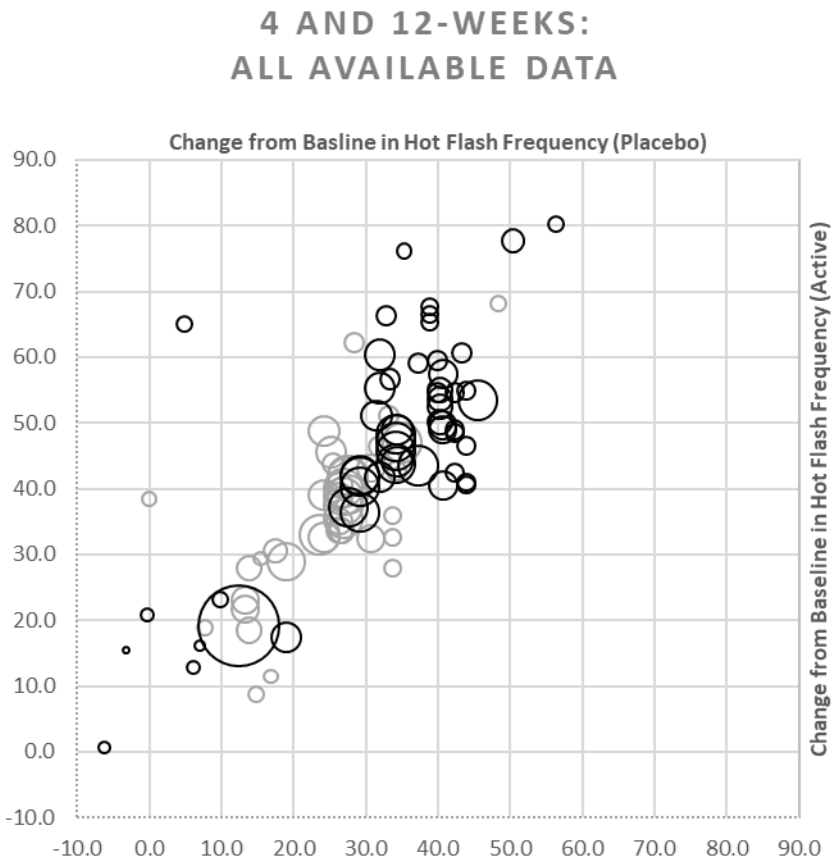
Last but certainly not least, I am extremely grateful to my family for continuous support and encouragement throughout my PhD course.

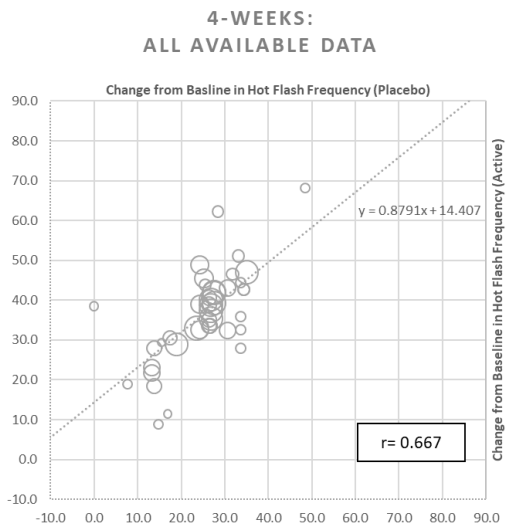
Appendices

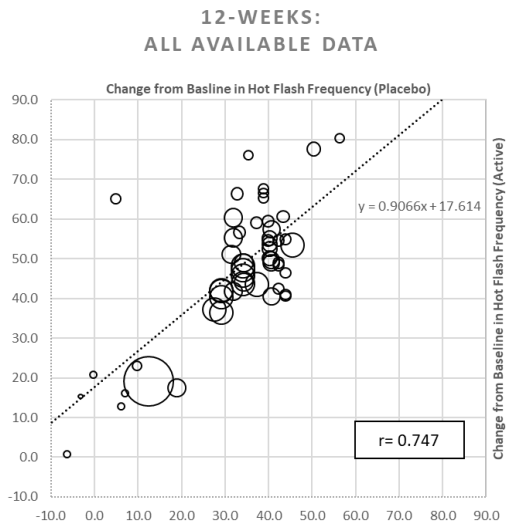


Appendix 1. CV seen in the studies considered success/failure in efficacy

The boxplots represent the 25th, 50th, and 75th percentiles, and the end of whiskers are the 5th and 95th percentile in each category.

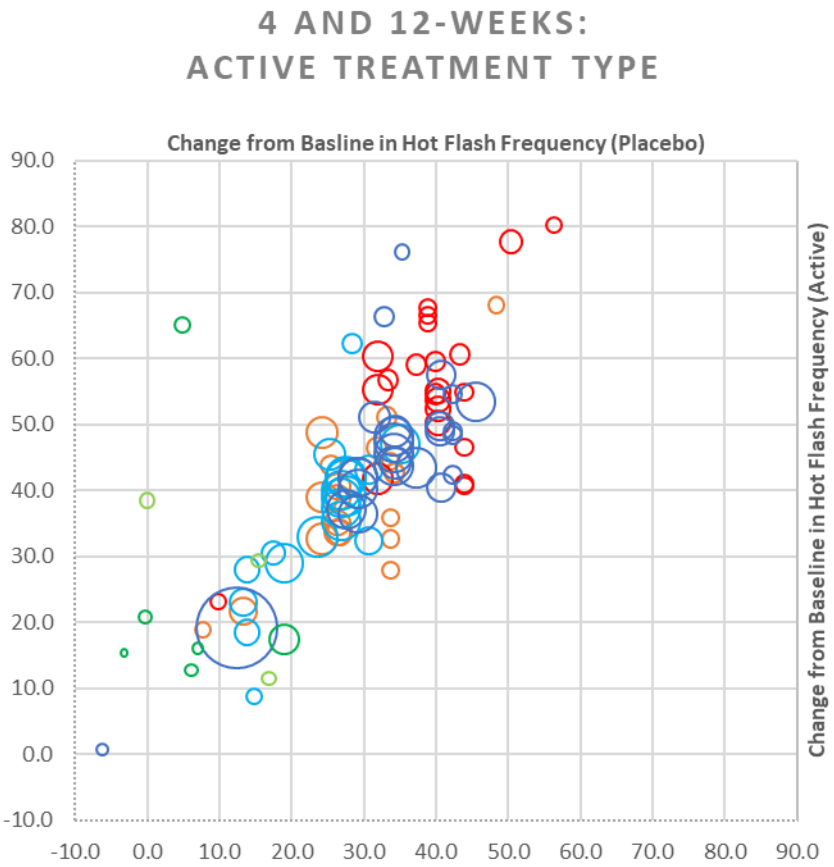


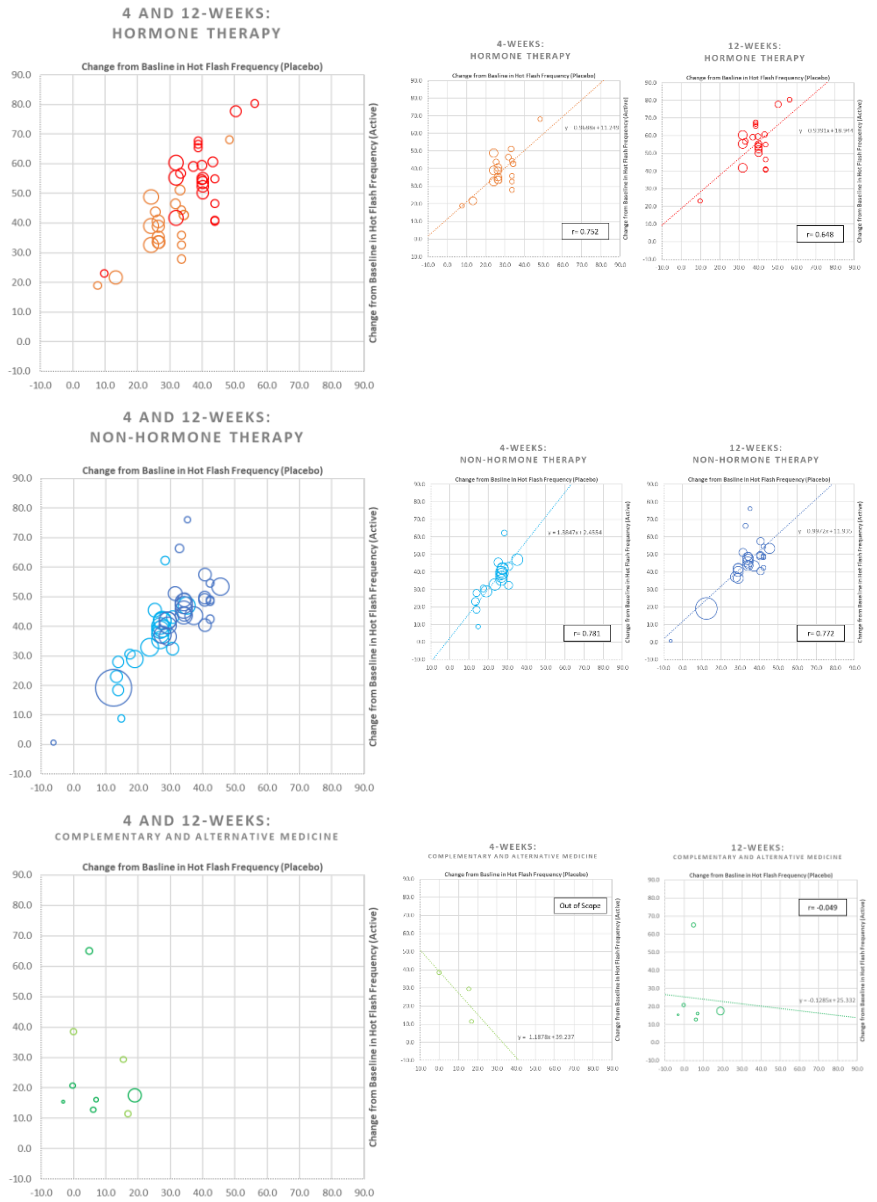




Appendix 2. Correlation of change from baseline in hot flash frequency between placebo and active treatment seen in the different treatment period duration

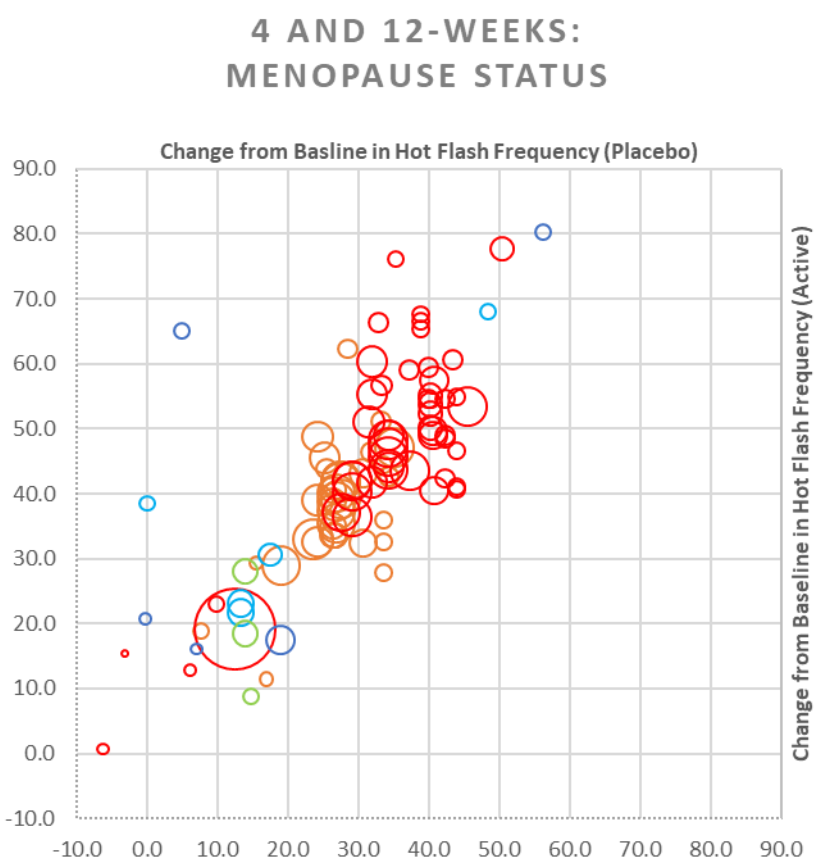
The size of the circles represents the sample size of the available data only at 4- and/or 12-weeks.

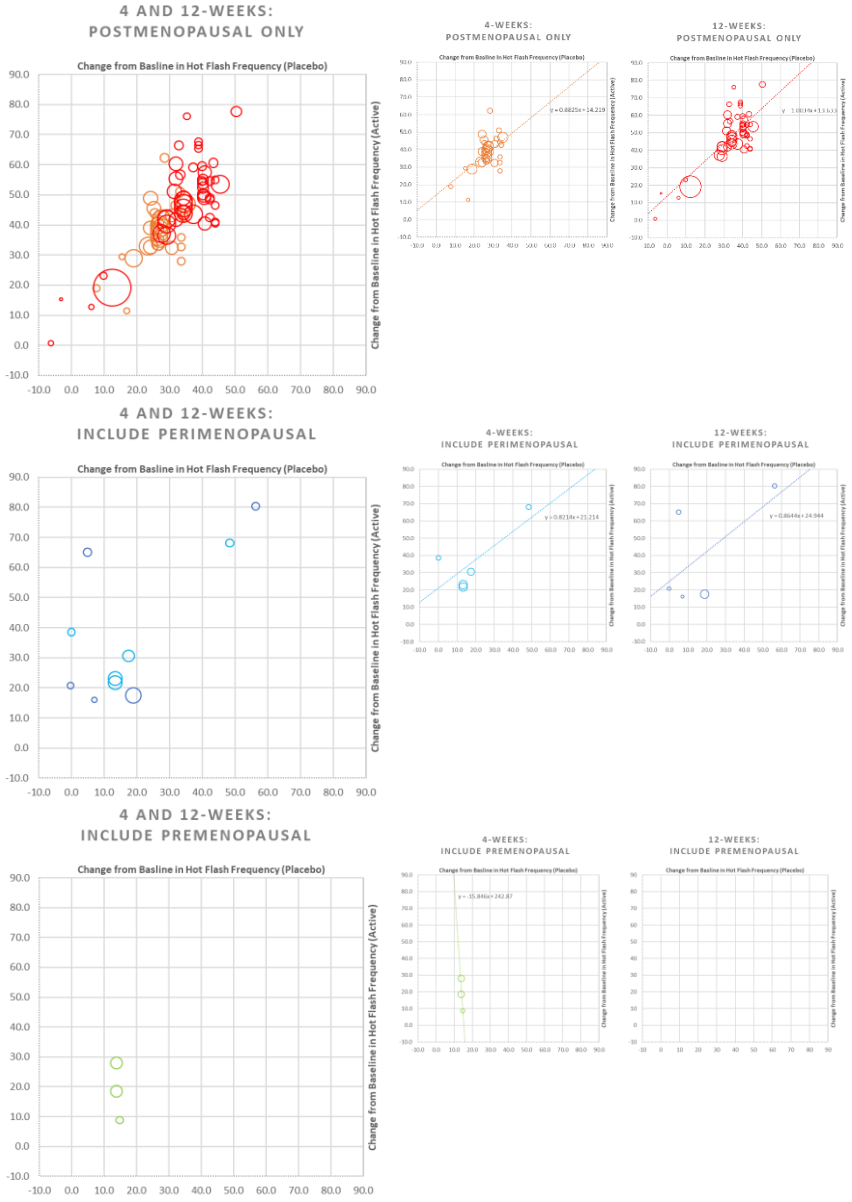




Appendix 3. Correlation of change from baseline in hot flash frequency between placebo and active treatment seen in the different active treatment type

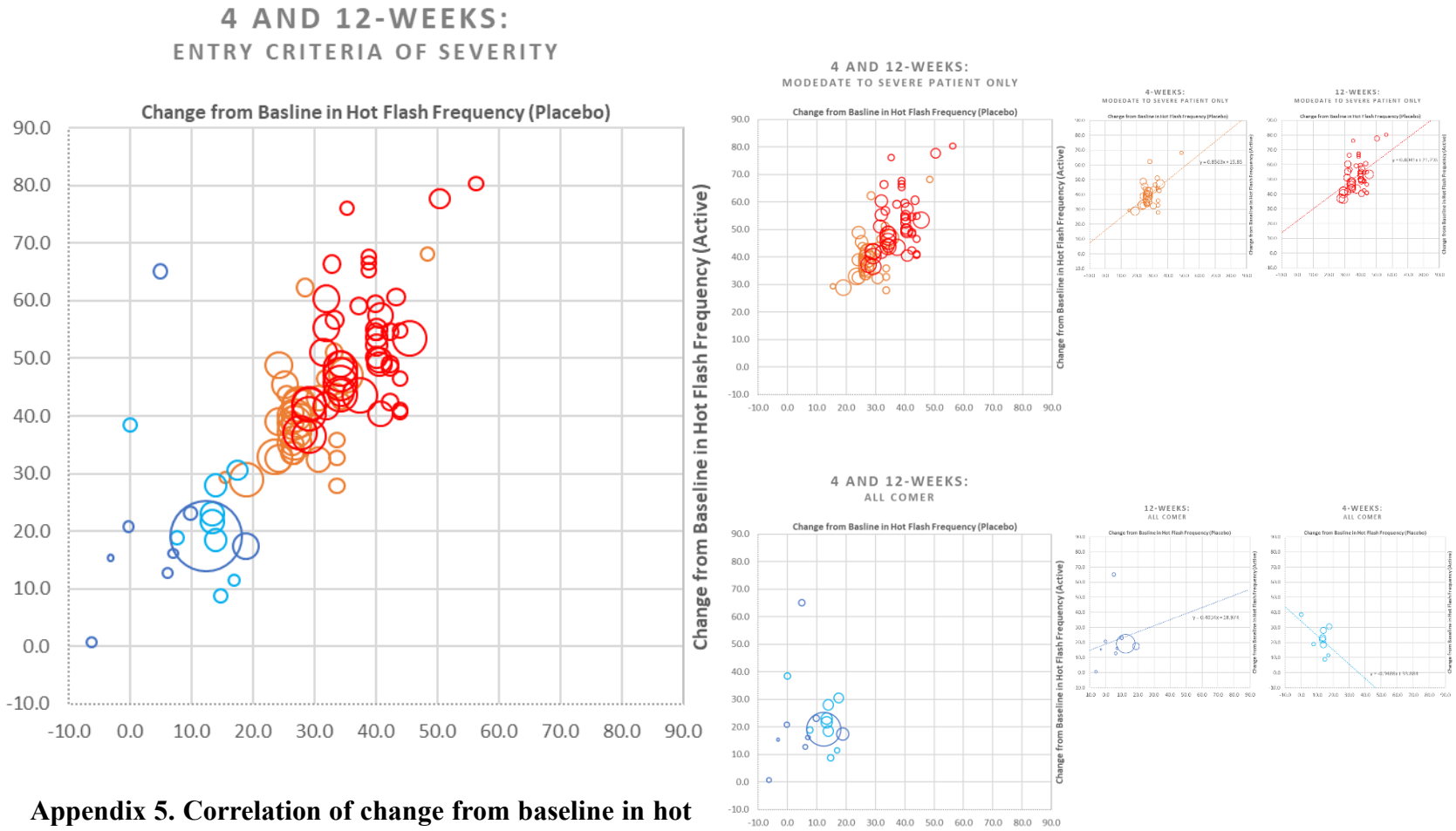
The size of the circles represents the sample size of the available data only at 4- and/or 12-weeks.





Appendix 4. Correlation of change from baseline in hot flash frequency between placebo and active treatment seen in the different menopausal status

The size of the circles represents the sample size of the available data only at 4- and/or 12-weeks.



Appendix 5. Correlation of change from baseline in hot flash frequency between placebo and active treatment seen in the different entry criteria of severity

The size of the circles represents the sample size of the available data only at 4- and/or 12-weeks.