Participation of Women in Clinical Trials Supporting FDA Approval of Cardiovascular Drugs

Pamela E. Scott, PhD, MA,a Ellis F. Unger, MD,b Marjorie R. Jenkins, MD, MEHP,a Mary Ross Southworth, PHARM D,b Tzu-Yun McDowell, PhD,b Ruth J. Geller, MHS,a Merina Elahi, BS,a Robert J. Temple, MD,b Janet Woodcock, MDb

ABSTRACT

BACKGROUND Concerns exist that women are underrepresented in trials of cardiovascular medications.

OBJECTIVES The authors sought to examine women’s participation and the reported safety and efficacy by gender for pivotal cardiovascular disease (CVD) trials submitted to the U.S. Food and Drug Administration (FDA) supporting marketing applications.

METHODS On the basis of publicly available FDA reviews, the authors assessed enrollment of women in trials supporting 36 drug approvals from 2005 to 2015. Prevalence-corrected estimates for the participation of women were calculated as the percentage of women among trial participants divided by the percentage of women in the disease population (participation to prevalence ratio [PPR]), with a range between 0.8 and 1.2 reflecting similar representation of women in the trial and disease population. Sex differences in efficacy and safety were assessed.

RESULTS The proportion of women enrolled ranged from 22% to 81% (mean 46%). The calculated PPR by disease area was within or above the desirable range for atrial fibrillation (0.8 to 1.1), hypertension (0.9), and pulmonary arterial hypertension (1.4); PPR was <0.8 for heart failure (0.5 to 0.6), coronary artery disease (0.6), and acute coronary syndrome/myocardial infarction (0.6). The authors found little indication of clinically meaningful gender differences in efficacy or safety. Gender differences in efficacy or safety were described in labeling for 4 drugs.

CONCLUSIONS Women were well represented in trials of drugs for hypertension and atrial fibrillation, and over-represented for pulmonary arterial hypertension. Representation of women fell below a PPR of 0.8 for trials in heart failure, coronary artery disease, and acute coronary syndrome. Minimal gender differences in drug efficacy and safety profiles were observed. (J Am Coll Cardiol 2018;71:1960–9) Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Since the mid-1980s, inclusion of women in clinical trials and analyses of potential gender differences in treatment response have been integral to the drug approval process (8). The U.S. Food and Drug Administration (FDA) continues to advance these efforts (9) by implementing regulations (10), issuing guidances, assessing demographic inclusion, and conducting of gender analyses (8,9,11). Discussion of demographics of trial participants and subset analyses are built into reviewer templates and addressed in reviewer training (12). The FDA Office of Women’s Health coordinates a lecture series on topics pertaining to inclusion of women, subpopulation analysis, and sex and gender differences in disease areas. As a result of these efforts, analyses for potential gender differences in drug trials have increased from 47% (13) (as reported by a survey of sponsors) to >90% (12) (based on FDA’s publicly available documents, e.g., reviews and product labeling) since the 1990s.

Over the past several decades, women’s participation in clinical trials has improved (14,15) in some (16), but not all, CVD areas (17,18). Hypothesized obstacles to participation of women include difficulty accessing study sites, familial responsibilities, cultural barriers, socioeconomic barriers, and concerns about risk (19-21). The prevalence of CVD is higher in older women (22), and previous studies have suggested that focus on recruitment of younger patients decreases overall enrollment of women (23). Inclusion criteria that would tend to select men and exclusion criteria more common in women have also been proposed as contributors (16). Despite these barriers, examples exist in which women and men showed comparable willingness to enroll in hypothetical (24,25) and actual (25-27) CVD trials.

We studied women’s participation in CVD trials supporting new drug application (NDA) approvals, as well as gender differences in trial results. For trials with available data, exploratory analyses were conducted to assess the potential impact of study screening criteria on gender differences in enrollment.

METHODS

SELECTION OF APPROVALS. The FDA granted 45 NDA approvals for cardiovascular indications between January 1, 2005, and September 15, 2015. Drugs for the following disease areas were included: acute coronary syndrome/myocardial infarction (ACS/MI), atrial fibrillation (AF), and coronary artery disease (CAD) including angina, heart failure (HF), hypertension, and pulmonary arterial hypertension (PAH) (Online Table 1). Data were collected from trials used to support drug approval. Of the 45 approvals, 9 were excluded for one of the following: pediatric indication, indication not within the 6 cardiovascular therapeutic areas, trial aimed to demonstrate bioequivalence only, or enrollment of <50 subjects (Online Figure 1). Of the remaining 36 approvals, 1 drug (ticagrelor) was approved for 2 indications (ACS and CAD). Hence, the analyses include 36 approvals (57 trials) for 35 drugs.

PARTICIPATION TO PREVALENCE RATIO. The participation to prevalence ratio (PPR) is a metric used to describe representation of women in a trial relative to their representation in the disease population (28,29). Trial representation was calculated by dividing the number of women in the trial by the total trial enrollment. For each disease area, the percentage of women in the disease population was estimated by dividing the prevalence or incidence of the disease among women by the total prevalence or incidence (Online Table 2). If a definitive, gender-stratified estimate was not available, 2 references for the percentage of women in the disease population were used to calculate a range for PPR. The PPR was calculated as follows:

\[
PPR = \frac{\text{percentage of women among trial participants}}{\text{percentage of women among disease population}}
\]

A PPR close to 1 indicates that the gender composition of the trial approximates that of the disease population. A PPR <0.8 or >1.2 indicates that women were underrepresented or overrepresented, respectively, relative to the disease population (28,29). For example, consider a disease where 40% of patients are women. If a clinical trial in this disease were to enroll 24% women, the PRR would be: percentage of women in the trial (24%)/percentage of women in the disease population (40%) = 0.6.

CLINICAL TRIAL PARTICIPATION. Data on participation were obtained from product labeling and medical and statistical reviews available publicly at Drugs@FDA.

EFFICACY AND SAFETY RESULTS BY GENDER. Many trials used binary endpoints (event yes/no), usually a composite endpoint that included events such as MI, stroke, and cardiovascular death. These were generally analyzed as time to first occurrence of an event. For these trials, gender-stratified hazard ratios and 95% confidence levels were calculated by CVD area
Hazard ratios (HRs) and 95% confidence intervals (CIs) for the primary efficacy endpoint by gender for targeted cardiovascular (CV) trials and categories. The HRs and 95% confidence limits that are shown do not take into account how many comparisons were made, nor do they reflect the effect of gender after adjustment for all other factors. Apparent homogeneity or heterogeneity among gender should not be overinterpreted. The primary efficacy endpoint and the reference group for each cardiovascular disease (CVD) area/drug are as follows. Acute coronary syndrome (ACS): cangrelor: the composite of all-cause mortality, myocardial infarction (MI), ischemia-driven revascularization, or stent thrombosis; prasugrel and ticagrelor: the composite of CV death, MI, or stroke (clopidogrel as the reference group for all 3 drugs). Atrial fibrillation (AF): apixaban, dabigatran, edoxaban, rivaroxaban: the composite of stroke or noncentral nervous system systemic embolism (warfarin as the reference group for all 4 drugs); dronedarone: the composite of first CV hospitalization or all-cause mortality (placebo as the reference group). Coronary artery disease (CAD): ticagrelor: the composite of CV death, MI, or stroke; vorapaxar: the composite of CV death, MI, or stroke; vorapaxar: the composite of CV death, MI, or stroke, or urgent coronary revascularization (placebo as the reference group for both drugs). Heart failure (HF): isosorbide/hydralazine: all-cause mortality was presented in this figure instead of the primary efficacy outcome, which was a composite score of clinical outcomes (placebo as the reference group); ivabradine and sacubitril/valsartan: the composite of hospitalization for worsening heart failure or CV death (enalapril as the reference group). Pulmonary arterial hypertension (PAH): macitentan: the composite of death, a significant morbidity event or other worsening PAH (placebo as the reference group). F = female; M = male.
For the trials with continuous endpoints, that is, variables that measure the magnitude of an effect and can assume any value within a range (e.g., blood pressure or 6-min walk distance), and/or trials without any pre-specified safety endpoints, we examined FDA reviews and product labeling to identify and describe gender differences in efficacy and safety. The primary efficacy endpoints for the trials with continuous endpoints are described in Online Table 3.

SCREENING FAILURES BY GENDER. To examine the potential impact of study eligibility criteria on gender differences in study enrollment, 5 NDAs (with available screening data) were examined to determine the numbers of patients not enrolled because of failure to
TABLE 1  Clinical Trial Participation of Women Across CVD Areas

<table>
<thead>
<tr>
<th>Cardiovascular Area</th>
<th>Number of Drug Approvals</th>
<th>Total Enrollment, N</th>
<th>Women Enrolled, n</th>
<th>Percentage of Women Among Trial Participants, % (Range)</th>
<th>Percentage of Women Among Disease Population*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute coronary syndrome/myocardial infarction</td>
<td>3</td>
<td>43,377</td>
<td>11,932</td>
<td>28 (26-28)</td>
<td>43</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>5</td>
<td>76,311</td>
<td>28,884</td>
<td>38 (35-47)</td>
<td>36; 49†</td>
</tr>
<tr>
<td>Coronary artery disease (including angina)</td>
<td>3</td>
<td>49,190</td>
<td>11,777</td>
<td>24 (24-25)</td>
<td>43</td>
</tr>
<tr>
<td>Heart failure</td>
<td>3</td>
<td>15,997</td>
<td>3,802</td>
<td>24 (22-40)</td>
<td>40; 53†</td>
</tr>
<tr>
<td>Hypertension</td>
<td>15</td>
<td>35,779</td>
<td>16,560</td>
<td>46 (27-53)</td>
<td>52</td>
</tr>
<tr>
<td>Pulmonary arterial hypertension</td>
<td>7</td>
<td>3,763</td>
<td>2,907</td>
<td>77 (74-81)</td>
<td>57</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>224,417</td>
<td>75,862</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

*References and calculations are presented in Online Tables 1 and 2. †Gender-stratified prevalence of atrial fibrillation in a representative population was not available. Two percentages were provided, one based on age-adjusted prevalence using published population-based studies (36%) and the other based on a cohort of atrial fibrillation patients within Kaiser Permanente of Northern California (53%). ‡Trials for heart failure drugs were conducted only among patients with reduced ejection fraction. Gender-stratified prevalence of heart failure with reduced ejection fraction was not available in a representative population. Two percentages were provided, one based on prevalence of all heart failure patients in the United States (53%) and the other based on the Framingham cohort with heart failure with reduced ejection fraction (46%).

CVD = cardiovascular disease.

meet study entrance criteria (screening failures) by gender.

RESULTS

PARTICIPATION OF WOMEN IN CLINICAL TRIALS.

Trials supporting 36 cardiovascular drug approvals between January 1, 2005, and September 15, 2015, included 224,417 participants, of whom 34% were women. Participation of women varied by trial (22% to 81%, mean per trial 46%) and cardiovascular area. The lowest enrollment of women was in HF (24%) and the highest in PAH (77%) (Table 1).

Women were represented at a rate similar to or greater than their share of the disease population in trials in PAH (PPR 1.4), AF (PPR 0.8 to 1.1), and hypertension (PPR 0.9), with representation below the pre-defined range in HF (PPR 0.5 to 0.6), CAD (PPR 0.6), and ACS/MI (PPR 0.6) (Central Illustration).

Efficacy Results by Gender. There were 14 drugs approved on the basis of clinical trials with a binary efficacy endpoint. Results of such trials can be expressed as a hazard ratio with 95% confidence intervals, and it is convenient and informative, therefore, to display the results of such trials by gender and other baseline characteristics in a single forest plot. These drugs were generally approved on the basis of efficacy data from a single trial. Results for these 14 drugs actually showed overlapping 95% confidence intervals for men and women (Figure 1), indicating similar effects for both. Results were not adjusted for other factors (e.g., age, weight, renal function). For the 22 drugs approved on the basis of trials with continuous endpoints, as described in Online Table 3, there were no indications of a gender difference, with 1 exception: for ranolazine, a drug indicated for angina, reductions in angina frequency and nitroglycerin use were less for women, and this is described in product labeling (Online Table 3).

Safety Results by Gender. Safety results by gender were obtained for 31 of 36 drugs (86%), and were not identified for 2 hypertension drugs and 3 PAH drugs. Although FDA generally assesses safety results by gender, for many drugs, the numbers of individual adverse events are too small to conduct meaningful subgroup analyses. Figure 2 shows bleeding results by gender for anticoagulation and antiplatelet drugs. The 95% confidence intervals for men and women generally overlapped, indicating similar effects in both genders. Of the remaining drugs, clinically significant gender differences in safety were described in the labels for 3 hypertension drugs (Online Table 4).

Screening Failures by Gender. Table 2 summarizes the numbers and percentages of women and men who were screened, screened out, and ultimately enrolled (Online Table 5). Overall, the percentage of women participating in a screening visit was similar to the percentage of women ultimately enrolled in the trial. Although a higher percentage of women than men were screened out for all 5 trials, differences were modest except for 1 ACS trial where 32% of women were screened out compared with 23% of men.

DISCUSSION

The proportion of women enrolled in cardiovascular trials supporting drug approvals ranged from 22% to 81% (mean per trial 46%) (Table 1). Women’s participation varied by disease area. Although participation approached (hypertension, AF) or exceeded (PAH) disease prevalence in some areas, women were underrepresented in others (HF, CAD, and ACS/MI) (Central Illustration). Our results were consistent with previous studies that found underrepresentation in some cardiovascular device clinical trials (30,31).

In the 3 HF trials, women’s participation ranged from 22% to 40%, and the PPR was 0.5 overall. These results are consistent with previous research, with low enrollment of women (range 21% to 29%) (16,23,32-34). Thus, enrollment has not increased over time. One HF trial conducted among African American participants achieved 40% enrollment of
women, demonstrating that enrolling larger numbers of diverse women is possible.

In the CAD and ACS/MI trials, 24% and 28% of enrolled patients were women, respectively, with a PPR <0.6 in both areas. In recent surveys of CAD and ACS trials, participation of women ranged between 25% and 33% (16,17,33). Underenrollment of women in these areas has been attributed to underenrollment of elderly patients (35) and the presence of comorbidities such as diabetes (17).

Gender differences exist in the clinical presentation of ACS and its diagnostic criteria, which could affect screening and enrollment in trials. Women tend to present at an older age, and in some cases without ST-segment elevation (17,36). Historically, it has been accepted that women are more likely to present with atypical symptoms of ischemic CVD (36-38), although recent research indicates similar symptoms in both genders (39,40). Further, data indicate that women with ACS may be less likely to undergo coronary angiography or percutaneous coronary intervention, which would impact eligibility for trials that require coronary anatomy documentation or percutaneous coronary intervention for entry (41,42). With respect to HF trials, HF with preserved ejection fraction (EF) is thought by some to be more common among women (43); therefore, inclusion criteria such as EF <40% might disproportionately exclude women with HF with preserved EF.

Some have hypothesized that inclusion and exclusion criteria disproportionately exclude women from cardiovascular studies, such that screening...
TABLE 2 Screened and Enrolled Participants by Gender

<table>
<thead>
<tr>
<th>Trial</th>
<th>Total Screened Women</th>
<th>% Women Among Total Screened (n/N)</th>
<th>Screened Out* (Number Screened Out/Number Screened)</th>
<th>% Screened Out Among Screened (Number Screened Out/Number Screened)</th>
<th>Enrolled† (Number Enrolled)</th>
<th>% Women Among Total Enrolled (n/N)</th>
<th>PPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute coronary syndrome #1 (prasugrel)</td>
<td>5,218</td>
<td>13,128</td>
<td>28 (5,218/18,346)</td>
<td>1,695 3,043</td>
<td>32 (1,695/5,218)</td>
<td>23 1,695 3,043</td>
<td>0.6</td>
</tr>
<tr>
<td>Atrial fibrillation #3 (rivaroxaban)</td>
<td>6,981</td>
<td>10,251</td>
<td>41 1,321 1,647</td>
<td>19 16 5,660 8,604</td>
<td>40 5,660 8,604</td>
<td>1.0 23 1,695 3,043</td>
<td></td>
</tr>
<tr>
<td>Atrial fibrillation #4 (apixaban)</td>
<td>7,507</td>
<td>13,491</td>
<td>36 1,091 1,706</td>
<td>15 13 6,416 11,785</td>
<td>35 6,416 11,785</td>
<td>1.0 8,184 22 0.4</td>
<td></td>
</tr>
<tr>
<td>Heart failure #3 (sacubitril/valsartan)</td>
<td>9,896</td>
<td>15,601</td>
<td>39 1,856 2,536</td>
<td>19 16 8,040 13,065</td>
<td>38 8,040 13,065</td>
<td>1.1 8,184 22 0.4</td>
<td></td>
</tr>
</tbody>
</table>

Values are n unless otherwise indicated. *Screen failures include participants who had a screening visit but were not randomized for various reasons. †Participants who were randomized in the trial. ‡This trial had run-in periods before a double-blind treatment period. The numbers indicate all participants eligible for the run-in period before randomization.

PPR = participation to prevalence ratio.

accounts for underparticipation of women. Although our screening data are limited to only 5 trials, our analyses do not suggest that gender-biased study entry criteria are the main reason for lower enrollment of women in cardiovascular trials (Table 2). Our data suggest that lower enrollment of women reflects the lower number of women referred for screening. Screening did not exclude nearly enough patients to account for the differences in participation that were observed. Thus, factors before screening, for example, identification of potential trial participants and ability of the candidate to participate, may be more important contributors to low enrollment of women (Online Figure 2). The data suggest that women are less likely than men to consider participation in trials, and/or they are less likely than men to be considered for screening in trials. We presume that both factors are operational for a variety of reasons, but their elucidation is beyond the scope of this paper. On the basis of our limited data, study inclusion and exclusion criteria appear to exert relatively minor effects on women’s participation.

We found few clinically meaningful gender differences in efficacy and safety in the drugs assessed. Although the 95% confidence intervals include the null value for some subgroups (Figure 1) (e.g., ticagrelor in women for CAD, isosorbide/hydralazine in men for HF), interpretation of apparent homogeneity or heterogeneity deserves caution. Results in the forest plot do not take into account multiple comparisons or adjust for other factors that may have affected outcomes. If there is a strong signal indicating a potential difference, FDA conducts additional analyses to evaluate the observed heterogeneity. For example, in the case of ticagrelor for CAD (Figure 1), FDA evaluated both clinical and pharmacokinetic data, and concluded that this apparent gender difference was probably a chance finding rather than a true gender effect (44). Gender differences that are found to be well-supported by the underlying data are described in labeling.

FDA examines both efficacy and safety data by demographic subgroups, including gender. In our assessment, for purposes of regulatory decision making, participation of women was sufficient to assess possible gender differences in safety and efficacy. For trials in cardiovascular areas in which women were most underrepresented on the basis of comparison to disease prevalence, the overall trial size was large, so that there were adequate numbers of women upon which to base assessments of efficacy and safety. With respect to efficacy, we recognize that studies are powered to demonstrate the treatment effect in the overall study population and almost never powered to demonstrate statistically significant treatment effects in subgroups. Nonetheless, when there are similar positive trends in females and males, and the numbers of females is reasonable, we are reassured that there are no clinically important treatment differences. Safety analyses differ inherently from efficacy analyses, in that they are generally descriptive, without planned statistical analysis, and include large numbers of potential signals. Once safety signals are identified, interpretation of subgroup analyses is particularly challenging. First, relatively few patients contribute data (i.e., only those with an adverse drug reaction), and second, multiple analyses are conducted, such that there is danger of overinterpretation of apparent small differences among subgroups.

FDA encourages companies to enroll patients who reflect the makeup of the population most likely to use the product, leading to a wide range of subgroups.
of interest (e.g., race, ethnicity, many age groups, renal function, and concomitant disease). Companies have an obligation specified in regulations to assess demographic subgroup differences in safety and effectiveness in their integrated analyses including gender differences (10), and FDA reviews these data to ensure that drugs are safe and effective for all of the intended population, recognizing that small differences can occur by chance. If, during review, FDA identifies gender differences in treatment response, FDA can require companies to develop and submit additional data, or present the differences in labeling, or may make approval decisions that reflect the difference.

There is no legal requirement for clinical trials to be powered to identify effects for subgroups based on gender, age, or other characteristics, and FDA has not identified or required specific numbers or percent-ages of patients for particular subgroups, including gender. When clinical trials are planned, they are sized to demonstrate a statistically significant treat-ment effect in the overall population; if they were sized to demonstrate statistically significant treat-ment effects in both genders or in all subgroups of interest, the number of patients in cardiovascular trials would increase dramatically.

FDA is aware that various groups have suggested target enrollment of women equal to the composition of the disease population (PPR of 1); others have encouraged powering clinical trials to detect statistically significant gender differences in the primary endpoints. Others have suggested that gender parity (50/50 enrollment) should be the standard. Although efforts led by key stakeholders, including FDA (45), have moved the discussion forward, none has led to a definitive recommendation regarding optimal repre-sentation of women in clinical trials because of vari-ations in disease areas and indications. In an effort to enhance transparency, FDA has implemented the Drug Trial Snapshots (46). Snapshots present the participation of patients in pivotal trials by age, gender, and race, and highlight whether there was any difference in benefits or side effects among these subgroups.

**STUDY LIMITATIONS.** This study included only pivotal studies and does not reflect the inclusion of women in all studies reviewed by the FDA for an NDA (such as early-phase studies). We focused on pivotal studies because these well-controlled studies most directly assess efficacy and safety of FDA-approved drugs.

A limitation of the PPR calculation is that the population used to derive prevalence or incidence estimates may be dissimilar to the population included in the trial(s) for a particular NDA or cardiovascular area. For example, HF with preserved and reduced EF are recognized as distinct diseases, but gender-stratified epidemiological data for each HF subgroup in a representative population were not available. The prevalence ratio also assumes that the percentage of women in the disease population is accurately known, which may not be the case. Another limitation of the PPR calculation is that it does not take age into account. Most CVD tends to affect women at an older age than men (16); therefore, both numerator and denominator of the PPR probably vary by age. High-quality, age- and gender-stratified epidemiological data for all 6 disease areas were not available; therefore, age-stratified PPR were not calculated in the present analysis.

Because it is not required for NDA submission, data on screening failures were available from only a small number of NDAs; therefore, our ability to draw conclusions about the effects of screening criteria in clinical trials is limited.

**CONCLUSIONS**

Based on prevalence-corrected estimates for repre-sentation of women, trials in hypertension and AF, were within, and trials in PAH were above the pre-defined range for similar representation of women, indicating success in enrolling women in some CVD clinical trials. However, representation of women was below the prevalence estimate for trials in HF, CAD, and ACS.

Based on this work, future research is needed to identify factors leading to underparticipation of women in cardiovascular clinical trials, particularly those occurring before screening (Figure 2). Research is needed to better define barriers that limit participation of diverse populations, not only of women, but also of minorities and the aged. Although some have postulated that inclusion/exclusion criteria have led to underparticipation of women in cardiovascular trials, we did not find evidence to support this concept. Disease prevalence data often lack age stratification. Because age affects the reported prevalence of CVD by gender, an area for future inquiry is exploration of the prevalence-adjusted representa-tion of women in cardiovascular clinical trials across relevant age categories, which will require high-quality epidemiological data stratified by age group and gender.

As we move into the era of precision medicine, that is, assessing the impact of a wide range of
patient and disease characteristics on drug effects, it is imperative that clinical trial participants represent the full spectrum of patients for whom the drug will be prescribed. Clinical researchers, patient advocacy groups, federal agencies, and industry must work together to ensure that representative patient populations are enrolled. These steps will move us closer toward the goal of providing the best information possible about the use of drugs for every patient.

ACKNOWLEDGMENT The authors thank Dania Shafei, MS, for providing editorial support for this paper.

ADDRESS FOR CORRESPONDENCE: Dr. Pamela Scott, U.S. Food and Drug Administration Office of Women’s Health, 10903 New Hampshire Avenue, Building 32, Room 2326, Silver Spring, Maryland 20993. E-mail: Pamela.Scott@fda.hhs.gov.

REFERENCES

KEY WORDS cardiovascular diseases/therapy, clinical trials, drug efficacy, drug safety, Food and Drug Administration, women’s health

APPENDIX For supplemental tables and figures, please see the online version of this paper.