*Contains Nonbinding Recommendations*

# Evaluation of Sex-Specific Data in Medical Device Clinical Studies

# Guidance for Industry and Food and Drug Administration Staff

**Document issued on March 31, 2025.**

**Document originally issued on August 22, 2014**

For questions about this document regarding CDRH-regulated devices, contact CDRH Health of Women at CDRHHealthofWomen@fda.hhs.gov. For questions about this document regarding CBER-regulated devices, contact the Office of Communication, Outreach, and Development (OCOD) at 1-800-835-4709 or 240-402-8010, or by email at ocod@fda.hhs.gov.




**U.S. Department of Health and Human Services**
**Food and Drug Administration**
**Center for Devices and Radiological Health**
**Center for Biologics Evaluation and Research**

GRZAN

# Preface

## Public Comment

You may submit electronic comments and suggestions at any time for Agency consideration to https://www.regulations.gov. Submit written comments to the Dockets Management Staff, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD 20852-1740. Identify all comments with the docket number FDA-2011-D-0817. Comments may not be acted upon by the Agency until the document is next revised or updated.

## Additional Copies

### CDRH

Additional copies are available from the Internet. You may also send an e-mail request to CDRH-Guidance@fda.hhs.gov to receive a copy of the guidance. Please include the document number GUI00001727 and complete title of the guidance in the request.

### CBER

Additional copies are available from the Center for Biologics Evaluation and Research (CBER), Office of Communication, Outreach, and Development (OCOD), 10903 New Hampshire Ave., Bldg. 71, Room 3128, Silver Spring, MD 20993-0002, or by calling 1-800-835-4709 or 240-402-8010, by email, ocod@fda.hhs.gov, or from the Internet at https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances.

GRZAN

# Table of Contents

**I. INTRODUCTION ..... 1**

**II. SCOPE ..... 1**

**III. BACKGROUND ..... 2**

A. WHY CONSIDER SEX DIFFERENCES? ..... 2
B. PARTICIPATION OF WOMEN IN CLINICAL STUDIES ..... 4
*1. Lack of Available Data for Women ..... 4*
*2. Barriers to Enrollment of Women ..... 5*

**IV. RECOMMENDATIONS FOR ACHIEVING APPROPRIATE ENROLLMENT ..... 7**

A. CONSIDERATION OF POTENTIAL SEX DIFFERENCES ..... 7
*1. For New or Ongoing Studies (IDE study design/early enrollment stage) ..... 7*
*2. For Completed Studies (marketing application stage) ..... 8*
*3. For Postmarket Studies (PAS or 522 PS stage) ..... 8*
B. STUDY DESIGN AND CONDUCT ..... 8
*1. For New or Ongoing Studies (IDE study design/early enrollment stage) ..... 8*
*2. For Completed Studies (marketing application stage) ..... 9*
*3. For Postmarket Studies (PAS or 522 PS stage) ..... 10*

**V. CONSIDERING SEX IN STUDY DESIGN AND DATA INTERPRETATION ..... 11**

A. STATISTICAL CONCEPTS FOR ASSESSING HETEROGENEITY ACROSS SEX GROUPS ..... 12
*1. For New or Ongoing Studies (IDE study design/early enrollment stage) ..... 13*
*2. For Completed Studies (marketing application stage) ..... 13*
*3. For Postmarket Studies (PAS or 522 PS stage) ..... 13*
B. RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL ELEMENTS IN STUDY DESIGN ..... 14
*When Sex Group Differences are Anticipated ..... 14*
*Pre-specifying Assessment of Heterogeneity Across Sex Groups in Study Design ..... 14*
*Additional Considerations for Particular Study Design Types ..... 15*
*Special Considerations for Diagnostic Devices ..... 16*
C. RECOMMENDATIONS FOR ANALYSIS AND INTERPRETATION OF SEX-SPECIFIC DATA IN COMPLETED STUDIES ..... 16
*Sex-Specific Analysis ..... 16*
*Additional Considerations for Data Analysis in Particular Study Design Types ..... 17*
D. INTERPRETATION OF SEX-SPECIFIC DATA ..... 18

**VI. RECOMMENDATIONS FOR REPORTING SEX-SPECIFIC INFORMATION IN APPLICATIONS AND PUBLIC DOCUMENTS ..... 18**

A. ENROLLMENT DEMOGRAPHICS, BASELINE CHARACTERISTICS & CO-MORBIDITIES ..... 19
*1. For New or Ongoing Studies (IDE study design/early enrollment stage) ..... 19*
*2. For Completed Studies (marketing application stage) ..... 19*
*3. For Postmarket Studies (PAS or 522 PS stage) ..... 20*
B. SEX-SPECIFIC OUTCOMES (SAFETY OR EFFECTIVENESS) ..... 20
*1. For Completed Studies (marketing application stage) ..... 20*
*2. For Postmarket Studies (PAS or 522 PS stage) ..... 21*

**APPENDIX 1 – DECISION FRAMEWORK ..... 22**

RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL DESIGN ..... 22
RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL ANALYSES FOR COMPLETED STUDIES - ONE –ARM STUDIES ..... 23
RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL ANALYSES FOR COMPLETED STUDIES - COMPARATIVE STUDIES ..... 24

# Evaluation of Sex-Specific Data in Medical Device Clinical Studies

# Guidance for Industry and Food and Drug Administration Staff

***This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff or Office responsible for this guidance as listed on the title page.***

## I. Introduction

This document provides guidance on the study and evaluation of sex-specific data in medical device clinical studies. The purpose of this guidance is to outline the FDA's expectations regarding sex-specific patient enrollment, data analysis, and reporting of study information. The primary intent is to improve the quality and consistency of available data regarding the performance of medical devices in both sexes by encouraging appropriate enrollment by sex in clinical studies of devices, and that data from such studies is appropriately analyzed by sex. This information can be of benefit to patients and their medical providers, as well as clinical researchers and others.

The specific objectives of this guidance are to: 1) encourage the consideration of sex and associated covariates (e.g., body size, plaque morphology) during the study design stage; 2) provide recommendations for study design and conduct to encourage appropriate enrollment of each sex (e.g., in proportions generally representative of the demographics of disease distribution, if appropriate); 3) outline recommended sex-specific statistical analyses of study data with a framework for considering sex-specific data when interpreting overall study outcomes; and 4) specify FDA's expectations for reporting sex-specific information in summaries and labeling for approved or cleared medical devices.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

## II. Scope

This guidance is intended for devices that require clinical information in support of a marketing submission, whether a premarket notification (510(k)), premarket approval (PMA)

application, Evaluation of Automatic Class III Designation (*de novo* request), or humanitarian device exemption (HDE) application. The recommendations contained herein also apply to post-approval study (PAS) submissions and postmarket surveillance (PS) studies conducted in accordance with Section 522 of the Food, Drug and Cosmetic Act, where noted.

The impact of demographic variables[1] may apply more to certain types of products or diseases than others. For example, certain OB/GYN and urology devices may be intended for use in single-sex populations, so studies of these devices would not be expected to address the potential for sex differences in outcome. Additionally, some *in vitro* diagnostic (IVD) device clinical studies are conducted on de-identified left over specimens so it may not be possible to obtain demographic information, including sex. As a result, evaluation of sex-specific data would not be possible in these cases.

FDA recommends the use of this guidance document as a supplement to other FDA guidance, in particular, any relevant device-specific guidance. Consultation with the FDA primary reviewing division is advised.

# III. Background

Certain elements described in this guidance have been emphasized in Agency regulations and/or policy in the past. Over recent decades the Agency's views, as well as those of the medical community in general, have evolved regarding women in clinical studies.

Prior to developing the policy set forth in this guidance, FDA publicly sought input from a variety of experts and stakeholders regarding the study and evaluation of women in clinical studies for medical devices. On June 2, 2008, various government agencies, physician professional societies, and patient advocacy groups participated in a public workshop to discuss ways to overcome barriers to understanding the impact of sex differences on clinical outcomes, with a focus on clinical study conduct and statistical analysis. On December 9, 2008, FDA's Center for Devices and Radiological Health (CDRH) and an industry trade association co-hosted a second public meeting to facilitate discussion in anticipation of issuance of FDA guidance on this subject.[2] This guidance document reflects the recommendations generated in these and other public fora and in subsequent internal Agency discussions. It is intended to provide guidance on the design and conduct of clinical studies to improve sex-specific information about the safety and effectiveness of approved new medical devices.

## A. Why consider sex differences?

Certain medical products elicit different responses in women compared to men. Differences may be attributable to intrinsic factors (e.g., genetics, hormones, body size, sex-specific physiology), extrinsic factors (e.g., diet, sociocultural issues, environment) or interactions between these factors. For example, there may be medical conditions that are unique to a certain sex, which should be considered in study recruitment and in reporting of results.

[1] Consult the Agency-wide guidance for industry *Collection of Race and Ethnicity Data in Clinical Trials*, which was issued September 2016.

[2] https://www.regulations.gov/document/FDA-2008-N-0038-0089

Covariates associated with female sex (e.g., size, age, co-morbidities, past pregnancies) may be responsible for certain differences in safety, effectiveness, or design attributes such as failure mode. Fluctuations associated with hormonal changes (e.g., onset of puberty, menstrual cycle, menopause, oral contraceptive or hormone replacement therapy use) may interact with clinical outcomes. Additionally, the menstrual cycle is associated with hormonally-mediated differences in metabolism or changes in fluid balance which could lead to intra-subject variability.

Following are examples where sex differences affect FDA's regulatory considerations:

1. Ventricular Assist Devices (VADs) provide mechanical circulatory support for patients with heart failure. One study of a next-generation VAD showed that in subjects treated with the investigational device, female sex or covariates associated with sex (body surface area, BSA) were found to be correlated with a higher rate of stroke in women as compared to men (18% vs. 6%). There were also trends toward increased rates of bleeding and infection in women compared to men. There did not appear to be differences in primary effectiveness outcome of survival (to cardiac transplantation or 180 days of support while being listed as status United Network for Organ Sharing (UNOS) 1A/1B for transplant). The strength of these conclusions is somewhat limited by the sample size (150 men and 44 women). The FDA Advisory Committee recommended that a post-approval study be conducted which would include adequate collection of data regarding both sex and body surface area to determine if differences exist in device performance. (Thoratec HeartMate II, Summary of Safety and Effectiveness: http://www.accessdata.fda.gov/cdrh_docs/pdf6/P060040b.pdf)

2. Cardiac Resynchronization Therapy Defibrillators (CRT-D) provide two functions. As an implantable cardioverter defibrillator (ICD) it senses dangerous abnormal heart rhythms and then delivers a shock to stop the abnormal rhythm, allowing the normal rhythm to resume. As cardiac resynchronization therapy, it generates small electrical impulses to coordinate the beating of the left and right ventricles so that they work together more efficiently to pump blood throughout the body. In one study, the benefit of CRT-D therapy over ICD alone (benefit defined as reduction in the composite endpoint of all-cause mortality or first heart failure event) was observed to be greater in women than men (77% versus 42%). Left Bundle Branch Block (LBBB) is a marker of an electrical conduction disorder in the heart and has been associated with a greater benefit in patients receiving CRT; the proportion of subjects with LBBB in this study was significantly greater in women than men (87% versus 65%). These findings are considered exploratory since the sex-specific analysis was *post hoc*. There did not appear to be differences in primary safety outcome of system-related complication-free survival within 91 days post implant. The FDA Advisory Committee recommended that two post-approval studies be conducted that would include adequate collection of data regarding the effects of the therapy in patients fulfilling the approved indication.

3. Hip joint deterioration can lead to pain, stiffness or difficulty walking. When these symptoms do not respond to conservative treatment, such as physical therapy, patients may be advised to undergo total hip replacement (THR) or hip resurfacing. As part of this treatment, patients may receive a "metal-on-metal" (MoM) hip implant in which the "ball and socket" of the device are both made from metal.[3] In June 2012, the Orthopaedic and Rehabilitation Devices Advisory Panel met to discuss the clinical performance of MoM hip implants as well as associated adverse events, including early device failure and the need for revision surgery.[4] The THR and hip resurfacing studies that identified revision rates by sex show that the revision rate appeared higher among women 3-5 years post implant in most studies. Sex-specific revision rates in THR studies ranged between 2.7% to 19.8% for women and 0 and 14.6% for men. Sex-specific revision rates in the resurfacing studies ranged between 0 and 27.6% for women and 1.4% and 8.97% for men.[5] Differences in sex-specific revision rates and the basis for these differences were a recurring concern throughout the panel discussion. From this information, FDA recommendations for orthopedic surgeons include that women may be at risk for increased device wear and/or adverse local tissue reactions and should be followed more closely.

## B. Participation of Women in Clinical Studies

Historically, women have been under-represented in or excluded from many clinical studies. This has led to a lack of information available for women and their physicians regarding the risks and benefits of many medical treatments and diagnostic procedures.

### 1. Lack of Available Data for Women

Concerns about representation of women in U.S. clinical trials initially surfaced in the drug context. In the mid-1970s, legislation and subsequent regulations and guidelines conveyed the recommendations of FDA and many in the medical and scientific community that women "of child-bearing potential" be excluded from drug studies to protect the fetus from exposure to unknown drugs.[6] However, it

[3] Metal-on-Metal Hip Implants. http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/MetalonMetalHipImplants/default.htm (Page last updated 1/17/2013)

[4] 2012 Meeting Materials of the Orthopaedic and Rehabilitation Devices Panel. https://wayback.archive-it.org/7993/20170403223426/https:/www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/MedicalDevices/MedicalDevicesAdvisoryCommittee/OrthopaedicandRehabilitationDevicesPanel/ucm309184.htm (Page last updated 03/30/2017)

[5] FDA Executive Summary Memorandum. Metal-on-Metal Hip Implant Systems. June 27-28, 2012. https://wayback.archive-it.org/7993/20170404140134/https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/MedicalDevices/MedicalDevicesAdvisoryCommittee/OrthopaedicandRehabilitationDevicesPanel/UCM309302.pdf

[6] U.S. Department of Health, Education, and Welfare, "General Considerations for the Clinical Evaluation of Drugs, HEW (FDA) 77-3040" (Government Printing Office, Washington, September 1977).

soon became apparent that this policy contributed to "compromis[ing] the quality of health information available to women as well as the health care they receive."[7]

The Government Accountability Office (GAO) audited clinical study information submitted to FDA in support of drug marketing applications, and concluded in a 1992 report that women were significantly underrepresented, and sex-specific data analysis was performed in less than 50% of drug studies. A 2001 report by the GAO on FDA-reviewed drug studies found that women accounted for 52% of total study enrollees, but approximately 30% of the study documents examined did not report outcomes by sex, and almost 40% did not report enrollment demographics.[8] Since then, studies have shown improvements in the inclusion of women and sex-specific analysis and reporting in drug studies.[9] In medical device studies, an evaluation of cardiovascular PMAs reported in 2009 showed pivotal studies that reported sex enrolled an average of 33.9% women.[10]

## 2. Barriers to Enrollment of Women

Women may be less likely to enroll in clinical studies. There are myriad suspected reasons for the continued lower participation rates of women in clinical studies in certain product areas. Some of the key reasons suggested at the June 2008 FDA workshop include:

- Lack of understanding about main obstacles to participation of women in clinical research;
- Fear of fetal consequences if a female participant becomes pregnant (e.g., effects of radiographic assessments or concomitant drug therapy);
- Inclusion/exclusion criteria potentially not needed to define the study population may unintentionally exclude women (e.g., upper age limit);
- Lack of understanding about differences in disease etiology and pathophysiology may lead to under-diagnosis and under-referral of women;
- Investigator and sponsor avoidance of female patients due to the perception that it takes more time and money to recruit them; and

[7] Women's health. Report of the Public Health Service Task Force on Women's Health Issues. Public Health Rep. 1985 Jan–Feb; 100(1): 73–106.
[8] United States General Accounting Office. Women's Health. Women Sufficiently Represented in New Drug Testing, but FDA Oversight Needs Improvement. (http://www.gao.gov/new.items/d01754.pdf, ed,: 2001).
[9] Poon R, Khanijow K, Umarjee S, Fadiran E, Yu M, Zhang L, Parekh A. Participation of women and sex analyses in late-phase clinical trials of new molecular entity drugs and biologics approved by the FDA in 2007-2009. *J Womens Health (Larchmt)*. 2013 Jul;22(7):604-16. Epub 2013 Jun 14.
[10] Kramer DB, Mallis E, Zuckerman BD, Zimmerman BA, Maisel WH. Premarket Clinical Evaluation of Novel Cardiovascular Devices: Quality Analysis of Premarket Clinical Studies Submitted to the Food and Drug Administration 2000–2007. *Am J Therapeutics.* 2009.

- Family responsibilities limiting women's ability to commit time for study follow-up.

Lower rates of participation by women in device clinical studies may also be attributable in part to limitations of manufacturing certain medical devices to accommodate anatomical differences between women and men; for example, technology may not yet be developed to manufacture smaller sizes or certain configurations of some devices which could increase use in women.

Where ongoing enrollment data demonstrate an underrepresentation of women enrolling in the study, sponsors are encouraged to investigate the reason for lack of enrollment and consider the approaches in Section IV.B.1. to enhance enrollment. It may be informative to evaluate whether the demographic distribution varies at different key time points (e.g., at screening, after evaluation of study inclusion/exclusion criteria, after consent, and at various follow-up time points). For example, if the proportion of women drops significantly after screening for inclusion/exclusion criteria, this may suggest that the study criteria may need to be examined to reduce inappropriate, unintentional exclusion of women. Similarly, cutoffs excluding patients with smaller body surface area (BSA) may exclude large proportions of female patients who may otherwise benefit from treatment. Removing such exclusions (entirely or through parallel cohort studies) could improve the participation rates of women in the overall study. Information regarding changes in demographic distribution at key time points in study screening, enrollment, and follow-up can provide insight into methods to substantially lower barriers to enrollment of women, as well as other subgroups of study participants, (e.g., flexibility in follow-up visit scheduling with consideration of child care or elder care services during appointments). Changes to a study protocol and informed consent may be made based on demographic distribution information with appropriate notification to and approval from the IRB and FDA, where necessary.

Sponsors may also wish to consider resources the National Institutes of Health developed,[11] or discussion with academic and contract research organizations, and high-enrolling clinical study sites, in determining practices best suited to achieve appropriate enrollment with respect to demographic groups, and to provide investigator training about these techniques.

Some specific examples of strategies to increase inclusion are discussed in Section IV.B below.

[11] The National Institute of Mental Health developed a resource document ("Points to Consider about Recruitment and Retention While Preparing a Clinical Research Study"), which outlines common issues that can impact clinical recruitment and retention, and strategies to address these issues. https://www.nimh.nih.gov/funding/grant-writing-and-application-process/points-to-consider-about-recruitment-and-retention-while-preparing-a-clinical-research-study

# IV. Recommendations for Achieving Appropriate Enrollment

It remains important that clinical trials include populations which reflect the intended population, whenever possible and appropriate. In general, to achieve an unbiased estimate of treatment effect in the general population, sponsors should plan to enroll representative proportions of women and men (e.g., consistent with disease prevalence). However, in cases where disease science or prior clinical study results suggest treatment effect in only one sex, sponsors may need to intentionally enroll sufficient numbers to support valid analysis (i.e., a sample size sufficient for sex-specific claims).

Historically, many medical device clinical studies have not enrolled proportions of women that reflect the underlying disease distribution in the affected population. This can be problematic because the ability to detect differences in response to treatment is markedly diminished if there is no or limited clinical experience with the product in the subgroup of interest. This has contributed to a substantial lack of available data regarding the risks and benefits of medical device use in women.

## A. Consideration of Potential Sex Differences

To understand potential sex differences that may be relevant to the clinical evaluation of your device, we recommend that you investigate whether sex differences may or may not exist for the disease or condition which your device is intended to treat or diagnose in the following areas:[12]

- sex-specific prevalence;
- sex-specific diagnosis and treatment patterns;
- identification of proportions of women included in past studies for the target indication;
- identification of any known clinically meaningful sex differences in outcomes related to either safety or effectiveness.

If information demonstrating sex differences is available, it should be included in your study and submission documents as described in the following sections. FDA recognizes that such information is limited in some device development programs (e.g., those based on testing of de-identified non-annotated specimens).

### 1. For New or Ongoing Studies (IDE study design/early enrollment stage)

You should include the information described above as part of the risk analysis section of your investigational plan (see 21 CFR 812.25(c)). We also recommend that you summarize this information in your study protocol and investigator training materials to explain the importance of enrolling appropriate proportions of women.

[12] The intent is to provide context based on disease science. Sponsors may consider providing similar information related to other demographic groups such as age, race, ethnicity, co-morbidities, etc.

For studies which are already enrolling under an approved (or conditionally approved) IDE where there is inadequate enrollment of women, FDA and the sponsor should discuss an appropriate path to communicate this new information to investigators without introducing bias to the study.

### 2. For Completed Studies (marketing application stage)

Where available background information or clinical study results suggest there are clinically meaningful sex differences, you should include this information as part of your marketing application in sections containing results of clinical investigations. A summary of this information should also be included in your draft PMA Summary of Safety and Effectiveness or 510(k) Summary, and in your labeling (see Section VI below for more details).

### 3. For Postmarket Studies (PAS or 522 PS stage)

Where available background information or clinical study results suggest there are clinically meaningful sex differences, you should include this information in interim reports and in the results section of your final report. If warranted, you should also submit revised labeling to include this information.

## B. Study Design and Conduct

As discussed in Section III.B., women have been historically under-represented in clinical studies of medical devices; therefore, the approaches described below are aimed at increasing enrollment of women in your study. However, in fields where men may be under-represented (e.g., breast cancer diagnosis, bone density scans) we recommend that you adapt these or other methods to increase enrollment of men if the intended population also includes men.

### 1. For New or Ongoing Studies (IDE study design/early enrollment stage)

You should develop and describe your plan to prospectively enroll proportions of each sex in your study which are appropriate based on the contextual information provided in Section IV. A. (e.g., consistent with the sex-specific prevalence of the disease or condition which your device is intended to treat or diagnose). To enhance enrollment of women, the approaches described below may be considered, with appropriate caution designed to avoid introducing bias or jeopardizing data validity.

a. Target investigational sites where recruitment of women can be more easily facilitated (e.g., women's clinics).

b. Consider alternative communication strategies (as used in the Women's Health Initiative study[13]) for study recruitment, informed consent documents, and patient materials.

[13] J. Hays, et al. The Women's Health Initiative Recruitment Methods and Results. *Ann Epidemiol 2003;13:S18–S77.*

c. If women are likely to benefit from your device but may not meet certain study enrollment criteria, consider revising the enrollment criteria, when appropriate, or consider parallel cohorts for collecting data on device use in women.

d. Responsibly enroll women of child-bearing age with appropriate risk reduction to avoid pregnancy during clinical trial participation.

e. Include provisions to encourage certain target enrollment for women (e.g., maintain open enrollment for women until pre-specified proportion is reached).

f. Investigate reasons for under-enrollment or non-enrollment of women or other key demographic groups (e.g., periodically evaluate screening logs for all patients who are screened but not ultimately enrolled in studies).

g. Plan focused efforts to enroll women under a continued access study.[14]

h. Consider factors that generally increase recruitment and retention such as community or local health care practitioner involvement in recruiting or referring patients, incentives or compensation (e.g., for transportation costs), and presentation of the benefits of participating in the study (e.g., send a newsletter to subjects to maintain interest).

i. Consider flexibility in follow-up visit scheduling with provision of child care or elder care services during appointments or to allow various opportunities that match subjects' schedules, which may include evenings and weekends.

j. For in vitro diagnostic tests and diagnostic devices, include samples from both women and men at the cutoff selection and validation stages.

## 2. For Completed Studies (marketing application stage)

If available evidence suggests that there may be clinically meaningful sex differences in outcomes (related to safety and/or effectiveness) with your device, results should then be discussed within your marketing application and considered in the context of available alternative treatments to determine whether additional data collection (for men and/or women) are needed to address a clinically important question before the device is marketed. Consideration should also be given to whether results support market approval in one sex, with additional pre-market data collection in the other sex; or whether market approval is supported for both sexes, with post-market studies to gain further information regarding any observed sex differences. The FDA team may recommend that you consider:

[14] see *FDA Decisions for Investigational Device Exemption Clinical Investigations: Guidance for Sponsors, Clinical Investigators, Institutional Review Boards, and Food and Drug Administration Staff available* at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-decisions-investigational-device-exemption-clinical-investigations for more information on IDE supplements.

a. Planning focused efforts to enroll women or men under a continued access study

b. Including provisions to encourage certain target enrollment for women or men (e.g., maintain open enrollment for women until a pre-specified proportion is reached).

## 3. For Postmarket Studies (PAS or 522 PS stage)

You should develop and describe your plan to enroll and retain proportions of women and men in your study that are consistent with the sex-specific prevalence of the type of disease or condition that your device is intended to treat or diagnose. For PAS designed for continued follow-up of the pivotal study cohort, FDA may determine that additional study of one sex is warranted if the pre-market study data suggest there are clinically meaningful sex differences. To enhance enrollment of women or men, we recommend that you undertake the following:

a. Consider whether outstanding questions warrant specific post-market evaluation in female-only or male-only studies based, for example, on sex-specific signals observed in pre-market clinical studies or known sex differences in the underlying disease or the response to concomitant treatment or therapies that may affect safety or effectiveness.

b. Target investigational sites where recruitment of needed populations can be more easily facilitated (e.g., women's clinics).

c. Consider alternative communication strategies (as used in the Women's Health Initiative study[15]) for study recruitment, informed consent documents and patient labeling.

d. Periodically evaluate screening logs to identify reasons for under-enrollment of women or men or other key demographic groups.

e. Consider factors that increase recruitment such as community or local health care practitioner involvement in recruiting or referring patients, incentives or compensation (e.g., for transportation costs), and presentation of the benefits of participating in the study (e.g., send a newsletter to subjects to stimulate interest).

f. Consider flexibility in follow-up visit scheduling with provision of child care or elder care services during appointments or to allow various opportunities that match subjects' schedules, which may include evenings and weekends.

[15] J. Hays, et al. The Women's Health Initiative Recruitment Methods and Results. *Ann Epidemiol 2003;13:S18–S77.*

We also recommend that sponsors and clinical study investigators consider the approaches described below, which can help avoid or minimize loss-to-follow up of subjects (regardless of sex).

Sponsor Responsibilities

a. Develop a follow-up plan that details follow-up goals, frequency of contacts, and number and type of contact for patients missing a follow-up visit.

b. Demonstrate interest in the subjects (e.g., send newsletter to subjects to maintain interest).

c. Monitor follow-up rates closely so that follow-up problems can be identified and addressed as soon as possible.

d. Report subject accountability data as part of the study report.

Investigator Responsibilities

a. Counsel subjects about the importance of returning to follow-up during informed consent and follow-up visits.

b. Remind subjects of upcoming scheduled follow-up visits.

c. Attempt to locate/return patients who miss scheduled clinic visits.

d. Obtain proxy information to use when unable to contact a study subject.

e. Ask subjects who withdraw during the study to provide the reason for withdrawal and ask them whether the investigator may contact them once more at the end of the study follow-up to assess the experience with device

f. Demonstrate interest in the subjects (e.g., telephone follow-up after surgery, particularly if the device is implantable).

# V. Considering Sex in Study Design and Data Interpretation

Differences between men and women range from the obvious (e.g., sexual organs, body fat distribution) to the less obvious (e.g., bone density, blood viscosity). Genetic sex can affect all levels of biological organization (cell, organ, organ system, and organism), including susceptibility to disease. Differences across the sexes in the incidence and severity of certain diseases may be related to differences in exposures, routes of entry and processing of a foreign agent, and cellular responses. Therefore, unless the investigational device is intended for use in only one sex (e.g., pregnancy test, or PSA testing), it is important that the variation in data across sex be considered in both study design and interpretation of study data.

## A. Statistical Concepts for Assessing Heterogeneity Across Sex Groups

There may be a substantial difference in how a device performs in women versus men in terms of safety or effectiveness. Thorough investigation of heterogeneity across sex groups, especially for primary safety and effectiveness endpoints, should be conducted. Heterogeneity here refers to a difference in outcome across sexes. Statistical hypothesis tests can be performed to detect heterogeneity, and methods of statistical inference for estimating its magnitude are also available.

When multiple treatment groups are considered, a form of heterogeneity is treatment by sex interaction, which measures the magnitudes of differences in outcome across treatments in one sex compared with the other. The concept of treatment by sex interaction applies to a study endpoint (such as probability of survival, adverse event rate) involving the comparison between two treatments. It is important to distinguish between qualitative versus quantitative interactions. Qualitative treatment by sex interaction for a parameter refers to the situation where one treatment is superior to the other in one sex, but not in the opposite sex. Quantitative treatment by sex interaction refers to the situation where one treatment is superior to the other in both sexes but by different amounts (See Figure 1 below).

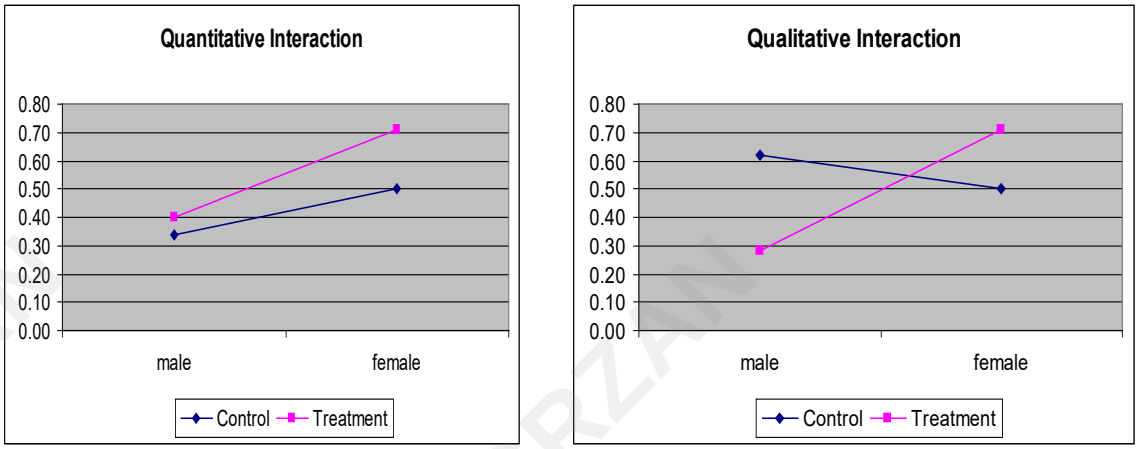



Figure 1. Illustrations of quantitative (left graph) and qualitative (right graph) interactions .

Statistical hypothesis tests of treatment by sex interaction have been widely utilized to detect treatment effect heterogeneity across sex. Most of the tests of interaction in common use have as their null hypotheses the absence of treatment by sex interaction. As statistical tests, their significance levels should be pre-specified in the investigational plan. Note, however, that the power of such tests may be unspecified. Therefore, lack of statistical significance for a test of treatment by sex interaction may not convincingly evidence the absence of clinically relevant interaction. By the same token, moderate statistical significance may not convincingly evidence the presence of clinically relevant interaction. While statistically significant interactions will be investigated for their clinical meaningfulness, interactions

without associated statistical significance may also be examined for clinical reasons specific to the design and endpoint.

For studies involving a single treatment with a single device (one-arm study), heterogeneity across sex groups can be assessed only for that single treatment and device. The concept of treatment by sex interaction has no direct applicability in such studies. To assess heterogeneity, statistical hypothesis tests comparing two sex groups under the (single) study treatment may be utilized, and in this specific context they are often subject to limitations similar to those besetting the aforementioned statistical tests of treatment by sex interaction.

Other patient characteristics (e.g., body size, co-morbidities, age) correlated with sex sometimes might explain apparent sex differences in clinical outcomes. If differences between men and women are observed, FDA recommends that a sponsor investigate potential explanation of the differences by other patient characteristics.

### 1. For New or Ongoing Studies (IDE study design/early enrollment stage)

- The Statistical Analysis Plan (SAP) in the protocol should include pre-specified plans for addressing the issues described in the sections below.
- It remains important that clinical trials include populations that reflect the intended population, whenever possible and appropriate. In general, to achieve an unbiased estimate of treatment effect in the general population, sponsors should provide a strategy to enroll representative proportions of women and men (e.g., consistent with disease prevalence).
- Sponsors should make an effort to identify in advance any key covariates that might explain possible differences across sexes, to plan to collect data on these covariates, and to pre-specify a modeling approach to investigate the extent to which these covariates can explain the observed differences.

### 2. For Completed Studies (marketing application stage)

In general, all studies should report descriptive statistics for outcomes of interest by sex as detailed in Section C below. After overall effectiveness and safety have been investigated, the influence of sex on primary endpoints for both safety and effectiveness should be assessed. If any clinically meaningful sex differences are suspected, either based on pre-specified or exploratory *post hoc* analyses, sponsors should discuss with FDA to determine whether additional data are needed to address any remaining sex-specific questions of safety or effectiveness.

### 3. For Postmarket Studies (PAS or 522 PS stage)

For PAS involving continuing data collection on PMA cohort patients, we recommend that you conduct the analyses described in Section C below for all follow-up time points.

For PAS (or 522 PS studies) involving newly enrolled patients, you should include the analyses described in Section C below as part of a pre-specified statistical analysis plan in

your protocol. Furthermore, if results from sex-specific analyses of pre-market data suggest there may be a clinically meaningful difference in outcomes, you should consult with the Division of Epidemiology to determine whether this should also be incorporated into the study design and hypothesis for your PAS.

When exploring sex-related differences during analysis of data from a PAS or 522 PS study, we recommend you address the issue of confounding by using multivariate analyses adjusted for patient characteristics that may confound the relationship between sex and study outcomes (e.g., smaller size, diabetes, etc.).

## B. Recommendations for Sex-Specific Statistical Elements in Study Design

When Sex Group Differences are Anticipated

- If, based on previous studies, literature, or disease science, important differences in the benefit-risk profile of a medical device are anticipated between men and women, clinical study design should take this into consideration.
- For devices that are appropriate for both men and women, where background information or previous clinical study results point to the potential existence of a clinically meaningful difference by sex, the study may need to be powered to evaluate treatment effect for both sex groups if the intended claim is for both sexes. In other words, sponsors may need to intentionally enroll sufficient number of patients in each sex group to allow valid analysis (i.e., a sample size sufficient for sex-specific claims); a stratified study design with outcome analyses by sex may be needed.
- A single study can be designed to support marketing approval for the combined population of men and women or one sex only. A common key element of all such study designs is successful control of Type 1 error rates at the desired levels, taking into account the multiplicity due to the two ways to claim study success. Just as with any study having a complex design, the sponsor is encouraged to talk to FDA early.
- Although rarely done, it is possible to plan a study that simultaneously investigates the overall treatment effect and the effect on only one subgroup such as women (or men). This would be done if the claim were for the entire population or just one pre-identified sex. One approach would be to allocate some fraction *f* of the overall Type I error rate (*alpha*) to the investigation of the overall inferential procedure and the rest to investigating the particular subgroup. In the hypothesis testing framework, the study would then be successful if either the overall test were significant at level *f* times *alpha* or the subgroup was effective at level (*1-f*) times *alpha*.
- Studies may be designed to investigate overall treatment effect in the combined population, and if positive, conduct pre-specified secondary analyses in one sex or another.

Pre-specifying Assessment of Heterogeneity Across Sex Groups in Study Design

- Unless a device to be studied is intended for use in only one sex (e.g., pregnancy test, PSA testing), it is important that variability in data across sex groups and its

interpretation be considered in the study design even if no substantial sex difference is expected at the design stage.

- The statistical analysis plan should include a strategy for assessing heterogeneity across sexes, since FDA recommends such an assessment as an integral part of interpreting study results for every submission. In particular, the heterogeneity assessment can serve as the basis for poolability conditions for studies with pre-specified success criteria expressed in terms of data pooled across sex groups. Such poolability conditions bear some resemblance to those commonly used for determining whether data can appropriately be pooled for analysis across different clinical sites. Poolability conditions may be specified as statistical hypothesis tests, which, for studies involving the comparison of two treatments, would typically be tests of treatment by sex interaction. The interaction tests should ideally be able to detect interaction of relevant magnitude measured on pertinent parameters with a reasonably high probability, and this goal should guide the choice of appropriate significance level.

Additional Considerations for Particular Study Design Types

- For one-arm studies:
    - Sponsors should provide strategy for assessing heterogeneity across sex groups. [16] The specific methodology could vary; if the methodology requires any assumptions, the validity of these assumptions should be investigated.
    - Sponsors may also consider sex-specific objective performance criteria (OPC) or performance goals. It may be used for sex-specific claims. It is important to control overall type 1 error rate to support any multiple claims based on hypothesis testing.
- For comparative studies:
    - Sponsors should pre-specify interaction testing. The validity of any assumptions should be investigated.
    - Sponsors may consider powering for sex-specific claims when sex-subgroup differences are anticipated. If seeking multiple claims based on hypothesis testing, it is important to control overall type 1 error rate.
    - If the control is non-randomized or historical and patient-level data exist, then the interaction can be investigated in conjunction with a propensity score data analysis.
    - For randomized controlled trials, sponsors may consider sex as a stratification variable in the randomization process if clinically meaningful sex difference is anticipated.

[16] This type of analysis is currently conducted for the purposes of determining whether data can appropriately be pooled for analysis.

Special Considerations for Diagnostic Devices

For *in vitro* diagnostic assays, imaging devices, and diagnostic devices in which a cutoff is used, sponsors should include data from both women and men both at the cutoff selection and cutoff validation stages. An assay or device involves a cutoff whenever a continuous or ordinal measurement is used to separate patients into two or more categories (for example, diseased and non-diseased). Separate cutoffs for men and women should be used only when there is reason to believe separate cutoffs are needed based on previous evidence or if the data in the current clinical study provide evidence for different cutoffs. The use of separate cutoffs may affect study design and sample size calculations. Analysis by sex of clinical performance measures such as sensitivity, specificity, positive and negative likelihood ratios, and positive and negative predictive values should be performed. Analysis of reference intervals with regard to mean (median) values, standard deviation and percentiles should be performed for men and women separately. Separate reference intervals for men and women should be considered only if they will be clinically useful and when there is reason to believe such intervals are needed based on previous evidence. For new measurands, if the information necessary to decide these questions is not available, but the data of the reference interval study indicate sex-specific differences, reference intervals should be presented for men and women separately and for combined data. Situations may arise in which an assay or device has high overall accuracy (e.g., very high sensitivity and specificity); when this occurs, subgroup analysis may not be warranted.

## C. Recommendations for Analysis and Interpretation of Sex-Specific Data in Completed Studies

Sex-Specific Analysis

In general, all studies should report descriptive statistics for outcomes of interest, including the estimate of variance or standard deviation (as applicable), by sex. At the primary follow-up time-point, regardless of the potentially limited statistical power of these sex-specific subgroup analyses, data should be examined for clinically meaningful sex differences in each of the following:

- primary effectiveness endpoint(s);
- primary safety endpoint(s); and
- key secondary endpoints.

- After overall effectiveness and safety have been investigated, the influence of sex on primary endpoints for both safety and effectiveness (and in some cases for important secondary endpoints as well) should be assessed.
- It is important to carry out all analyses set forth in the Statistical Analysis Plan (SAP). FDA expects sponsors to plan and conduct analyses to evaluate heterogeneity by sex, including treatment by sex interaction when applicable, as described in previous sections.

- In some cases the test for treatment by sex interaction (or heterogeneity in general) may have adequate power to detect only a very large interaction (or heterogeneity) but may fail to detect a smaller yet clinically important interaction (or heterogeneity). Such situations may arise when the number of patients in one or both of the sex groups is small, in which case additional data from men or women (or both) may be required. Observed heterogeneity could exist across sexes due to large variability associated with small sample sizes; interpretation of clinical meaningfulness may be premature in those cases. Consultation with FDA is recommended.
- For recommendations on interpreting data, see Section D below.

Additional Considerations for Data Analysis in Particular Study Design Types

- For one-arm studies:
    - If overall treatment effect is neither statistically significant nor clinically meaningful, subgroup analyses are not recommended. In such cases, analysis likely raises questions about data to support marketing application.
    - If no significant difference is observed across sexes, data may be poolable across sex.
    - If a significant difference is observed across sexes, it is important to explore whether the difference remains significant after adjusting for other covariates. If not, data may be poolable across sex.
    - If difference remains significant after adjusting for other covariates, data may not be poolable across sex. Additional data may be required to appropriately evaluate the effect of sex on the study endpoints. In these cases, discussion with FDA is advised.
- For comparative studies:
    - If overall treatment effect is not statistically significant and clinically meaningful, subgroup analyses are not recommended. In such cases, analysis likely raises questions about data to support marketing application.
    - If there is evidence of an interaction of treatment by sex, it is important to describe the nature of interaction (qualitative or quantitative) and assess the clinical importance of the differences. In some cases, the interaction effect could be statistically significant but not clinically meaningful, or clinically meaningful but not statistically significant. In these cases, discussion with FDA is advised.
    - If no significant interaction effect between treatment and sex is observed for the outcome of interest, data may be poolable across sex. However, the decision about the validity of pooling the data should be based on the size of the observed interaction effect as well as its clinical importance.
    - If a treatment effect difference is noted across sexes, it is usually helpful to perform additional analyses to investigate possible explanations for this difference using variables such as body size (e.g., body mass index), bone density or concomitant illness (e.g., diabetes). If significant interaction effect

between treatment and sex is observed, explore whether this remains significant after adjusting for other covariates. If not, data may be poolable across sex.

- If the interaction effect remains significant after adjusting for other covariates, data may not be poolable across sex. Additional data may be required to appropriately evaluate the effect of sex on the study endpoints. In these cases, discussion with FDA is advised.
- If a significant treatment by sex interaction has been identified, it may be helpful to investigate if there is a sex difference in treatment group only, control group only, or both. Alternately, the interaction could be explored by assessing whether there is a treatment difference in women only, men only, or both.

### D. Interpretation of Sex-Specific Data

If any clinically meaningful sex differences are found, either based on pre-specified or exploratory *post hoc* analyses, you should discuss with FDA whether additional data are needed to address any remaining sex-specific questions.

If results of your analysis suggest that there is insufficient data to assess whether sex is associated with clinically meaningful differences in outcome, FDA may determine that clinical data from additional subjects in one or both sexes may be needed pre- or post-market to address potential sex-specific questions related to safety or effectiveness.

Although expected to be rare, in cases where clinically meaningful differences between the sexes are observed in safety or effectiveness, FDA may request additional confirmatory studies in one or both sexes, implement specific pre- or post-approval study conditions, and/or modify the design of subsequent studies.

There are limitations to interpreting clinically meaningful differences in small data sets. Mean differences could exist between sexes due to small samples sizes; interpretation about whether they are clinically meaningful may be premature in many cases.

## VI. Recommendations for Reporting Sex-Specific Information in Applications and Public Documents

Confidential submissions to FDA contain detailed analyses of clinical study data, which may include a variety of sex-specific analyses. However, public documents, including labeling and FDA summaries of review (e.g., SSED) for medical devices approved in the past, are inconsistent with regard to the degree of information reported on device performance in demographic subgroups. Although sponsors may be most interested in the generalizability of the findings, individual patients and their medical providers may benefit from more data regarding effectiveness and potential adverse events associated with device use in a particular demographic subgroup.

# A. Enrollment Demographics, Baseline Characteristics & Co-Morbidities

The strength of the conclusions of your clinical study(ies) with respect to device performance in women and men is linked to the proportions of each sex in your study(ies). FDA recommends that you report the number and proportion of subjects by sex who were treated or diagnosed with your device as part of a clinical study as follows:

- You should report study demographics in terms of proportion enrolled by subgroup. You should discuss whether the proportions enrolled are consistent with the sex-specific prevalence of disease, if known. For studies with multiple arms, you should report enrollment proportions for each sex in each arm.

- If co-morbidities and/or other baseline characteristics are collected, we recommend that you report these by demographic subgroup as well as overall.

- For per protocol analyses, we recommend a comparison and discussion of sex-specific differences in follow-up compared to at enrollment, for the overall study sample and for each study arm.

You may choose to adapt the example language below, or you may use similar language which incorporates the contents described above.

Example Language:

*Women represented [34%] of the total patients enrolled in the overall study. This is similar to the prevalence of [coronary artery disease] in the general U.S. population [citation]. Among subjects in the treatment group, m1/n1 (p1%) were women, and m2/n2 (p2%) of subjects in the control group were women.*

*Women were more likely to have diabetes compared to men (35% vs. 22%) and less likely to have prior history of myocardial infarction (24% vs. 36%).*

Additionally, we recommend that you include this type of information in any applicable tables and charts.

## 1. For New or Ongoing Studies (IDE study design/early enrollment stage)

You should report this information as part of your annual progress reports.

## 2. For Completed Studies (marketing application stage)

You should report this information as part of your marketing application in sections containing results of clinical investigations, including the labeling. A summary of this information should also be included in your draft PMA Summary of Safety and Effectiveness, 510(k) Summary, or *de novo* decision summary.

### 3. For Postmarket Studies (PAS or 522 PS stage)

You should report this information in interim reports and in the results section of your final report.

## B. Sex-Specific Outcomes (Safety or Effectiveness)

Information regarding sex-specific outcomes analyses should be described in the labeling and summaries of review, regardless of whether the analyses are pre-specified or *post hoc.* Covariates that might explain possible outcome differences between sexes should be described.

- If outcome differences by sex are statistically significant and clinically meaningful, you should report the results of the outcome analyses.
- If results of these analyses suggest a sex difference in an endpoint or event that is clinically meaningful but not statistically significant, you should report the findings descriptively.
- If results of these analyses suggest no sex differences in outcomes, you should report which analyses were conducted and that no differences were found.

### 1. For Completed Studies (marketing application stage)

When presenting results of *prespecified* sex analyses, we recommend the following:

- Clearly state which analyses were conducted
- Specify statistical methods used to assess for heterogeneity of treatment differences by sex (as described above)
- You may include inferential statistics, including p-values and/or confidence intervals. To provide appropriate context, describe prior scientific evidence suggesting that clinically meaningful differences by sex are expected, or describe statistical limitations of analyses.

When presenting results of *post hoc* sex-specific analyses, we recommend the following:

- Clearly state that the analyses were unplanned
- Clearly state which analyses were conducted
- Specify statistical methods used to assess for heterogeneity of treatment differences by sex (as described above)
- Use descriptive statistics only (mean, standard deviation, etc.). Results in confidential submissions to PMA can include inferential statistics, with a disclaimer that these are from *post hoc* analyses.

If clinically meaningful sex differences in safety or effectiveness are observed, or if there are potential differences that might require follow-up studies, data on benefits and risks should be described separately for women and men in labeling and review summaries.

## 2. For Postmarket Studies (PAS or 522 PS stage)

When presenting results of sex-specific analyses of PAS or 522 PS data, the recommendations above should also apply.

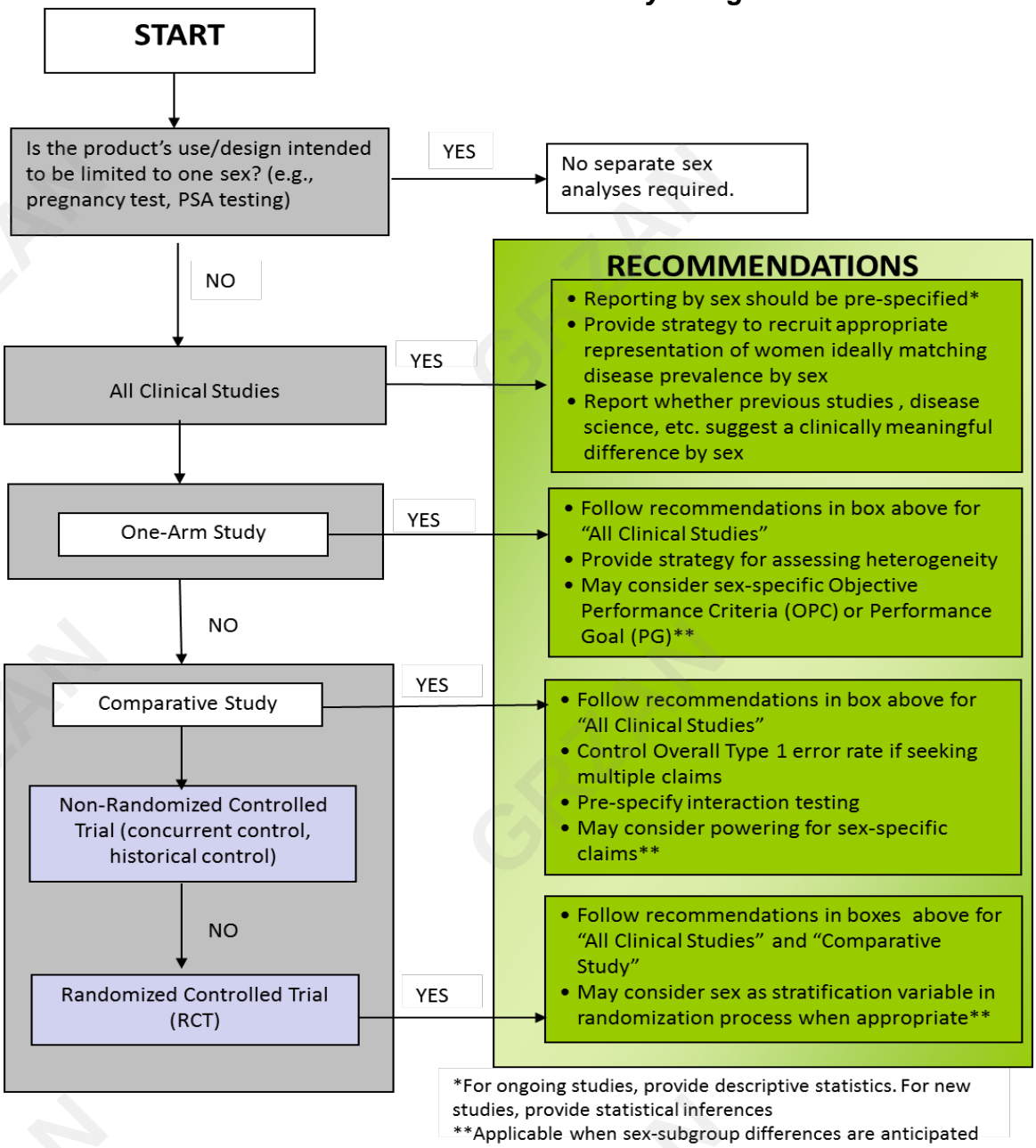
If a clinically meaningful signal is detected in your final analysis, FDA may recommend changes to your approved labeling and review summaries, which you should submit with your final study report.

# APPENDIX 1 – DECISION FRAMEWORK

We encourage the use of existing scientific data (e.g., previous studies, disease science) to determine whether there is a hypothesis for a clinically meaningful sex difference for your device. When there is a hypothesis for a clinically meaningful sex difference, the following decision trees provide a framework in deciding when various sex-specific statistical recommendations apply for different clinical study designs.

## RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL DESIGN

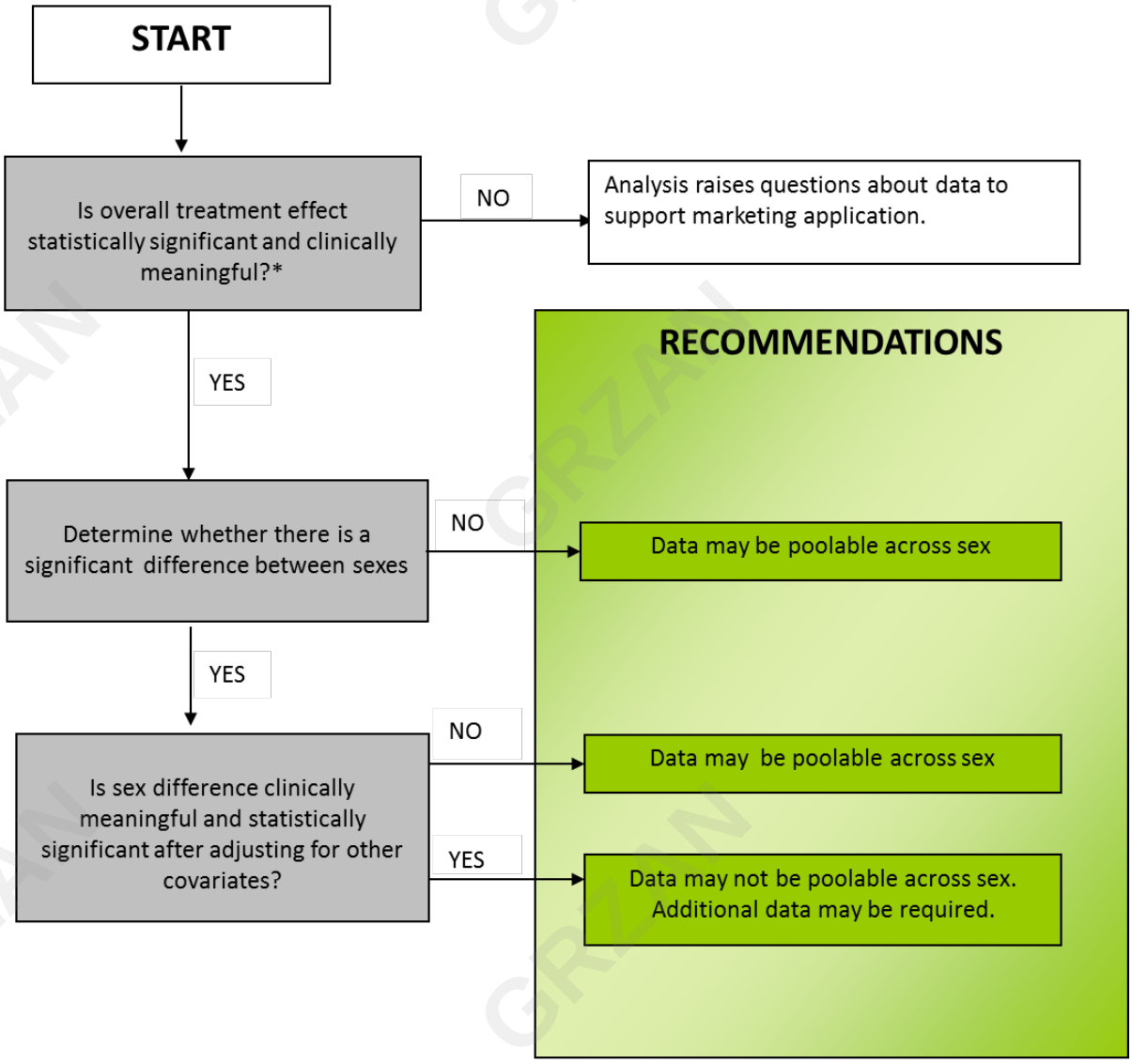
### *Follow Recommendations associated with study design*



# DECISION FRAMEWORK

## *RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL ANALYSES FOR COMPLETED STUDIES - ONE –ARM STUDIES*

**(Objective Performance Criterion, Performance goal, Observational Study)**

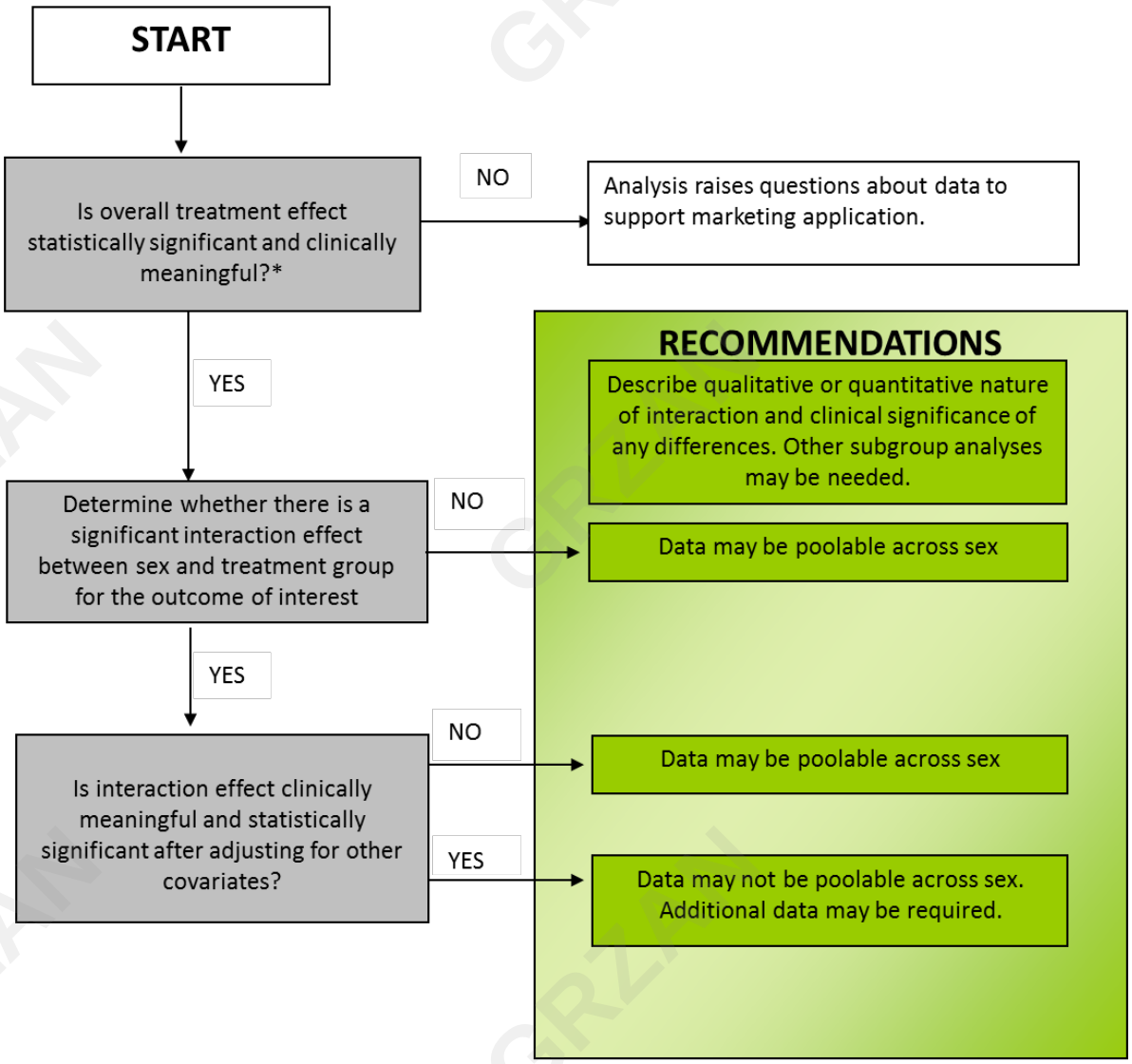



*Subgroup analyses are not recommended if overall treatment effect is not statistically significant and clinically meaningful.

Note: In some cases, the sex difference could be statistically significant but not clinically meaningful or clinically meaningful but not statistically significant. In these cases, discussion with FDA is advised.

# DECISION FRAMEWORK

## *RECOMMENDATIONS FOR SEX-SPECIFIC STATISTICAL ANALYSES FOR COMPLETED STUDIES - COMPARATIVE STUDIES*




*Subgroup analyses are not recommended if overall treatment effect is not statistically significant and clinically meaningful.

Note: In some cases, the interaction effect could be statistically significant but not clinically meaningful or clinically meaningful but not statistically significant. In these cases, discussion with FDA is advised.