



# Women in Clinical Trials: An FDA Perspective

Linda Ann Sherman, Robert Temple, Ruth B. Merkatz\*

Since the early 1980s, the U.S. Food and Drug Administration (FDA) has been interested in the individualization of therapy, that is, determining whether and how treatment should be modified for various demographic groups within the population. After several years of discussion, a guideline was published on the study of drugs in the elderly (1), and later a similar guideline on the study and evaluation of gender differences in the clinical evaluation of drugs was published (2). Both guidelines suggested that a representative sample of the patients likely to receive a drug should be included in clinical studies, and that data should be analyzed to determine whether drug effectiveness or safety responses in relevant demographic groups differ. Such analyses had already been requested by FDA in the content of the clinical and statistical sections of new drug applications (3).

In addition to stressing the need to include a broad population in clinical trials and to analyze different gender and age groups for safety and efficacy responses, the two guidelines placed strong emphasis on the need to identify pharmacokinetic (PK) differences between groups (differences in the blood concentrations of a drug after similar doses). These differences are common and an important cause of different responses, and they are relatively easy to detect compared with pharmacodynamic differences among groups (different responses to the same blood concentrations). One way to identify PK differences is to do separate studies in each group of interest, for example, in women, men, elderly individuals, and people with renal failure. As an alternative to numerous trials, both guidelines also suggested the use of a PK screen, a method of searching for PK differences by examining steady-state blood concentrations of drugs in most participants in clinical trials. The PK screen, by examining a large number of individuals, is a means of detecting, inexpensively, a wide range of PK differences due to demographic and other influences, such as metabolic or excretory differences. The guidelines also recommended additional small studies of particu-

lar relevance to specific populations, such as drug interaction studies between a test drug and oral contraceptives (2) and cognitive function studies in the elderly (1).

Although the 1993 gender guideline reflected a general FDA interest in the individualization of treatment, it also stated the agency's concern about the participation of women in studies of medical products. Beginning in the mid-1980s, strong views were expressed that a lack of participation of women in clinical studies was resulting in inadequate prescribing information related to the specific use of drugs in women (4). Analyses of published cardiovascular drug trials lent support to the conclusion that drug responses of women were not being assessed in several important areas (5). In the early 1990s, attention was focused on the importance of including women in early phases of trials for critical drugs so that they would have access to experimental therapies, especially for treatment of human immunodeficiency virus (HIV), which was becoming recognized as a major threat to women as well as to men (6).

## Prelude to the 1993 Gender Guideline

When the FDA analyzed the participation of the elderly and both sexes in the total clinical trial database used to support marketing applications, it was found, with respect to sex, that the proportions of men and women in clinical studies usually reflected the prevalence of the condition under study in the general population (2, 7). Thus, the majority of participants (60 to 65%) in studies of antidepressants, anxiolytics, and anti-inflammatory drugs were women, reflecting the greater use of psychiatric agents in women and the female predominance of rheumatic diseases. In contrast, drugs for cardiovascular diseases were studied primarily in men (60 to 65% of the participants were male), reflecting the greater prevalence of these diseases in men within the age range of most cardiovascular clinical trial participants (that is, usually less than 65 years of age).

Despite a gender predominance for various conditions, the representation of the minority gender was still substantial. Thus, in stating explicitly FDA's belief that a drug should be studied in the full range of people who would be exposed to it, including both

men and women, the gender guideline translated what was already the common practice into a more formal expectation. Nevertheless, although women and men were included in studies in numbers that could be evaluated, there were few attempts to use the trial data to examine whether responses were different by gender. The guideline explicitly sought to correct this deficiency.

Apart from the overall participation of women in clinical trials, there has been particular concern about their participation in early trials, especially when the drugs involved are medically important, such as drugs to treat acquired immune deficiency syndrome (AIDS) and its complications. This concern arose in part because of the 1977 FDA guideline entitled "General Considerations for the Clinical Evaluation of Drugs," which explicitly prohibited the participation of women of childbearing potential in phase 1 and early phase 2 trials (8). These women could not be included until animal teratogenicity studies had been completed and until early phase 2 studies revealed that the drug was effective in men, older women, or both. The phrase "women of childbearing potential" was defined very broadly, essentially as any premenopausal women physically capable of becoming pregnant. It excluded only women who were surgically sterile or postmenopausal. There was a clear exception to the 1977 ban with respect to the study of treatments for life-threatening illnesses.

The 1977 policy had many critics, both inside and outside the agency, who felt that it was both rigid and paternalistic and that women, physicians, and institutional review boards should have more say in the decisions about participation in clinical trials. Furthermore, although the restriction applied only to phase 1 and early phase 2 studies, which represent a small fraction of total new drug applications (NDAs), the ban has had a significant effect on drug evaluation. Although FDA analyses of NDA databases showed that women, including young women, were well represented overall in studies of new drugs, a 1994 survey of phase 1 (PK) studies of new drugs approved between 1985 and 1991 revealed that women were entirely excluded from more than 51% of PK protocols (9). In addition, an internal review of 152 HIV protocols (covering a period from 1988 through 1994) revealed that 24 (16%) of the studies did not enroll women. None of the 24 protocols contained specific exclusionary criteria for women (10). To the agency, it appeared that the 1977 guideline had a spillover effect, resulting in a lack of attention to the participation of women throughout the drug development process, at least in some cases. Because HIV infec-

L. A. Sherman is at the Office of Operations, HF 29, Office of the Commissioner, R. Temple is at the Center for Drug Evaluation and Research, FDA, and R. B. Merkatz is director at the Office of Women's Health, FDA, HF-8, 5600 Fishers Lane, Rockville, MD 20857, USA.

\*To whom correspondence should be addressed.

tion is a life-threatening illness that affects many women, making it essential that women participate in trials of HIV-related drugs, these findings also suggested that the issue of the participation of women in clinical trials was more complex than initially realized.

Persuaded that the 1977 guideline was indeed paternalistic, that studies in women of childbearing potential could be carried out without risk of fetal exposure, and firm in its commitment to promote women's health and to enhance knowledge about the effects of drugs in women, the FDA removed the 1977 restriction and issued a new guideline (11).

### Perceptions of the Guideline

Comments submitted to FDA and other public discussions indicate that there is general agreement with the purpose and direction of the gender guideline. There have, nonetheless, been areas of discussion and controversy. For example, because the guideline does not specify what constitutes an adequate sample of patients with the disease to be treated, nor exactly how to conduct the requested demographic group analyses, some have felt it does not go far enough in guaranteeing adequate data on usage of new drugs for women (12). Others, however, presumed that FDA was insisting that an NDA database be increased to provide separate answers to all questions for both sexes (13). Moreover, although the guideline described what information was needed in a premarketing application, it left to sponsors the choice of how and when to obtain these data. This suggested to some that there was insufficient attention to assuring participation of women in clinical trials, but to others it suggested a rigid requirement for equal gender participation in all studies.

It may be that the flexibility built into the guideline has created exaggerated concerns about FDA's real demands and expectations as well as concerns about our seriousness in proposing them. Certainly, there are many legitimate concerns about how the requested subset analyses will be conducted, evaluated, and interpreted. It is worth reiterating, therefore, what the guideline is and what it is not, and providing a brief analysis of some of its major points.

### Aims of the Guideline

First, the gender guideline states that participation of both sexes, taking into account the prevalence of the disease in each sex, is expected and that this is part of a general FDA expectation that drugs will be evaluated in a reasonable sample of the people who will receive them. It is not implied in the guideline that the overall

number of women or men currently in clinical trials is inadequate, but it is made clear that a drug development program that excludes any clinically relevant group, for example women or men, would present a problem in a drug's overall evaluation. It is not expected that most, or even many, drugs will behave differently in men and women, nor is their evidence to support such a conclusion; nonetheless, it is surprising how little analysis there has been on this important question.

The FDA believes that inclusion of both sexes in drug development trials and that analysis of efficacy and safety results by sex in large individual studies and in pooled data can provide insight into whether gender differences are present. Caution is required in these analyses because of the risk that the comparisons involved will lead to spurious conclusions. This is not, however, a reason not to explore the data to see whether there are, in fact, significant differences. Some differences will surely exist, and some have been found. For example, analysis of the calcium channel blocker amlodipine, revealed a higher rate of edema, flushing, and palpitations in women (13) (information which appears on the product label). Although this finding may reflect the smaller size of women and the relatively higher dose that they receive, this factor alone does not seem to account entirely for the gender difference. Similar overview analyses of data submitted in marketing applications to the FDA for beta-blocking agents and angiotensin-converting enzyme (ACE) inhibitors have revealed variations in antihypertensive responses between African-Americans and caucasians. These differences were not unexpected in view of the greater prevalence of high-renin-related hypertension in caucasians, but the ability to detect the differences support the usefulness of analyses of pooled controlled trial data. Similar analyses with a Medicaid database showed higher relative rates of angioedema, a rare but potentially life-threatening adverse reaction, in African-American patients receiving ACE inhibitors (14).

The pooled databases of most of the NDAs are already large enough to permit appropriate group analyses in most cases without a need to increase the number of participants from particular demographic groups. Even 30% of a 1000-patient database is potentially large enough to reveal subset differences in effectiveness or rates of common side effects. These approaches, of course, will generally not be helpful in the consideration of rare events. Moreover, in those NDAs based on one or two large intervention studies, for example, mortality studies (treatments after a heart attack such as thrombolysis), it can be expected that (15, p. 94)

Generally, trials adequate for detecting an overall treatment effect cannot be expected to detect effects within even relatively large subgroups. . . . Since most currently preformed clinical trials are barely large enough to detect overall effects, the numbers of patients in even the larger subgroups examined within these same trials would hardly be expected to provide reliable estimates of treatment effects.

Despite the 1988, 1989, and 1993 guidelines, the drug industry has been erratic in examining subsets of data from clinical trials with respect to age, race, and gender. A General Accounting Office (GAO) survey completed in 1992 (9) found that the data on fewer than half of the drugs were analyzed for gender-related differences in drug response. Because the GAO study included many applications submitted before the publication of the 1988 guideline (3), FDA conducted a second survey of NDAs submitted from June 1991 to July 1992 (16). This study found that of the 28 NDAs examined, 64% of the integrated summaries of safety contained analyses by age, 54% by gender, and 32% by race. Only 25% had safety analyses of all three subsets. With respect to effectiveness, 54% of the NDAs contained analyses by age, 43% by gender, and 32% by race. Only 11% contained analyses of all three subsets. More recently, a preliminary review of new molecular entities approved during 1993 revealed that 4 of the 15 submissions reviewed thus far had no analysis for gender, age, or race (17). Since 1993 the FDA has stated and instructed sponsors that without these analyses or an agreement to conduct them promptly, the review of new marketing applications will not be initiated.

Perhaps the most controversial aspect of the FDA's changes was the withdrawal of the 1977 ban on the inclusion of women of childbearing potential in phase 1 and 2 trials and our urging that women be included. Concerns have been expressed about potential liability in the event of a fetal abnormality, especially if women are allowed to participate in clinical trials before the completion of reproductive toxicology studies, as allowed under FDA's new policy. It has also been argued that recommendations in the document with regard to contraception were incompletely developed (18).

In its 1993 gender guideline, FDA did not abrogate its responsibility to the fetus, nor did it specify what the community must do. Rather, the agency withdrew a virtual federal ban on the inclusion of women in early studies and entrusted decisions about what to do to internal review boards, patients, and their physicians. This is not to say that the agency has no view on the matter. The FDA believes that the possibility of fetal exposure can be minimized by appropriate contraception, laboratory test-



ing, and behavior. It is expected that study participants will be counseled about the importance of preventing pregnancy during a clinical trial. It is also expected that laboratory screening for pregnancy will be conducted before and at appropriate intervals during the study. As more preclinical information becomes available, it is expected that the sponsor and investigator will provide the relevant information to the study participant. Nonetheless, in so sensitive an area strong, diverse opinions will lead to further debate about differences in practice. In the case of serious and life-threatening diseases, however, the agency remains committed to the inclusion of women in all phases of clinical trials.

In summary, the FDA expects sponsors to study the full range of patients likely to receive a drug, including both genders, and to analyze the data to determine whether responses in various groups are different. This expectation is not new and implementing it is not likely to add significantly to drug development costs. The FDA recognizes, however, that not all aspects of how to analyze data in population groups are settled and that the best way to obtain population pharmacokinetic information is still a matter of debate (19). Nevertheless, the FDA

believes that the changes outlined in the 1993 guideline, changes that the agency continues to implement and develop, will not only have a positive effect overall in fostering women's health but will improve the ability of physicians and other health providers to prescribe drugs safely and effectively for both men and women.

## REFERENCES AND NOTES

1. Center for Drug Evaluation and Research, *Guideline for the Study of Drugs Likely to be Used in the Elderly* (FDA, Washington, DC, 1989).
2. ———, *Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs* (FDA, Washington, DC, 1993).
3. ———, *Guideline for the Format and Content of the Clinical and Statistical Sections of New Drug Applications* (FDA, Washington, DC, 1988).
4. *Report of the Public Health Service Task Force on Women's Health Issues. Vol. 2* [publication no. (PHS) 85-50206, Department of Health and Human Services, Washington, DC, 1985]; B. A. Ray and M. C. Braude, Eds., *Women and Drugs: A New Era for Research* (research monograph 65, National Institute on Drug Abuse, Rockville, MD, 1986).
5. N. K. Wenger, *J. Am. Med. Assoc.* **268**, 1460 (1992); J. Froom and P. Froom, *J. Clin. Epidemiol.* **44**, 1127 (1991); U.S. Preventive Services Task Force, *Guide to Clinical Preventive Services* (Williams & Wilkins, Baltimore, MD, 1989), pp. 258–260; J. W. Rich-Edwards, J. A. E. Manson, C. H. Hennekens, J. E. Buring, *N. Engl. J. Med.* **332**, 1758 (1995).
6. AIDS Service Center HIV Law Project, NOW Legal Defense and Education Fund, American Civil Liberties Union, *AIDS Project Citizens Petition*, 15 December 1992; D. Murphy, *Food Drug Law J.* **48**, 175 (1993).
7. General Accounting Office, "Women's Health; FDA Needs to Ensure More Study of Gender Differences in Prescription Drug Testing" (Publication #GAO/HRD-93-17, Government Printing Office, Washington, DC, 1992).
8. Center for Drug Evaluation and Research, *General Considerations for the Clinical Evaluation of Drugs* (FDA, Washington, DC, 1977).
9. V. Hale, unpublished material.
10. T. A. Toigo, K. Struble, D. Birnkrant, R. E. Behrman, unpublished material.
11. R. B. Merkatz, R. Temple, S. Sobel, K. Feiden, D. A. Kessler, *N. Engl. J. Med.* **329**, 292 (1993).
12. Comments submitted to docket no. 93D-0236, Dockets Management Branch (HFA)-305, FDA, Rockville, MD; Proceedings of the *National Task Force on AIDS Drug Development*, Crystal City, VA, 19 January 1995.
13. *Physicians' Desk Reference* (Medical Economics Data Production Company, Montvale, NJ, 1995), p. 1877.
14. G. A. Burkhardt, M. R. Griffin, W. A. Ray, personal communication.
15. S. Yusef, J. Wittes, J. Probstfield, H. A. Tyroler, *J. Am. Med. Assoc.* **266**, 93 (1991).
16. R. Temple *et al.*, unpublished material.
17. S. Machado, unpublished material.
18. L. R. Johnson and J. Bush, "Activities of the Pharmaceutical Industry Relative to the FDA Gender Guideline," *Proceedings of the Workshop on Gender Guidelines*, Washington, DC, 18 January 1995, in press.
19. On 6 and 7 November 1995, the FDA will sponsor a workshop, *Gender Studies in Product Development: Scientific Issues and Approaches*, to probe further the scientific aspects of carrying out gender analysis.

# The Inclusion of Women in Clinical Trials

Curtis L. Meinert

The U.S. National Institutes of Health (NIH) Revitalization Act of 1993 requires in the case of any clinical trial involving treatment of diseases common to both genders that the trial is (1, p. 134)

designed and carried out in a manner sufficient to provide for a valid analysis of whether the variables being studied in the trial affect women or members of minority groups, as the case may be, differently than other subjects in the trial.

The "valid analysis" mandate came about because of the perception that the clinical research enterprise in the United States is biased in favor of male participants. The perception has come to be an accepted fact by inductive reasoning from specifics to the general. Most arguments intended to establish the perception as true focus on a 1977 guideline from the U.S. Food and Drug Administration that dis-

couraged the participation of women of childbearing potential in phase 1 and 2 drug trials (2) and in a few large-scale heart trials, principally, the Physicians' Health Study (3) and the Multiple Risk Factor Intervention Trial (4).

## New Guidelines

The NIH Revitalization Act has given rise to guidelines from NIH for implementing the valid analysis requirements of the legislation (5) and to a new bureaucracy for review of trials in relation to phase 3 trials (NIH has interpreted the Act to pertain exclusively to phase 3 trials; phase 1 and 2 trials have been exempted from the requirements of the Act). The Act has been challenged in a petition from the Society for Clinical Trials (6), in a directive from the membership of the Society to the director of NIH (7), and by various writers, including myself (8).

For a treatment to be of value for general use, it must first be shown to be of

value in some limited setting. There is no point in worrying about whether a treatment works the same or differently in men and women until it has been shown to work in someone.

Every trial involves a select, nonrepresentative study population. The requirement of consent alone is sufficient to ensure that fact. Hence, the strength of a trial lies in its internal validity. A comparison of treatment within a trial is valid as long as the demographic composition of the treatment groups is the same. There is no requirement for demographic coverage or representativeness for internal validity. Generalizations from the study population to the broader universe of patients are a matter of judgment and is always open to question, even when the trial involves a demographically heterogeneous population.

A preoccupation with subgrouping leads to a quagmire of confusion and to a mosaic with ever more parts. That the United States is headed in this direction seems apparent by the increasingly strident voices from constituent groups for their place in the mosaic. Each group argues that it is different from all others and, hence, must be represented in sufficient numbers to provide a valid analysis for them.

If we want to know more about the treatments we use in regard to demographics, we have to be prepared to pay the

The author is at Johns Hopkins University, School of Hygiene and Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA.