SHORT COMMUNICATION

Controversy concerning the Women's Health Initiative trial

Norman F. Angell, M.D., Ph.D.

I have read with great interest the article by Risk (1) and the response by Elhelw (2) in this journal. Few published papers have produced the furor that resulted from the publication of the principal results and subsequent papers produced from the Women's Health Initiative trial of estrogen and progestin (3). Initially there was a flurry of publications to support the conclusions of this study, but more recently others have questioned the initial interpretations. How can well-informed and honest scientists and physicians have such divergent opinions concerning the results of this study? I believe that these opinions differ in important areas about the trial. Of importance are: the design of the trial in relation to the questions posed and the manner in which the data were presented, the confidence in the degree of certainty produced by the results of this trial, and the contribution of learned societies, the media, and society to the dissemination of the results.

The design of this study and the description of the population studied have been well described (3). The WHI trial was described as a primary prevention trial and has been presented as such. There are controversies about the question of whether this is a primary prevention trial, about the choice of the population, and problems with the trial itself. The analysis of the degree in which this trial can be accepted as a primary trial allow us to evaluate several aspects that are crucial to the evaluation of the data.

1. A trial should study clinically important diseases which affect large numbers of people and produce significant morbidity/mortality.

Coronary artery disease and breast cancer clearly meet these tests.

2. There should be significant uncertainty concerning the proposed intervention. While a large proportion of obstetrician/gynecologists accepted hormone replacement therapy, other specialists and an important proportion of laywomen were uncertain of the efficacy and safety of hormones.

3. The trial should be large enough to have adequate power to answer the questions posed. The trial initially was large with 16608 subjects but there was more than 50% of the subjects lost during the trial by drop out and drop in. The outcomes measured were relatively uncommon so that absolute differences in the variables were small.

4. Precursors of disease should be easily measured and predictable of clinical disease. Heart disease and breast cancer are difficult to detect prior to clinical disease. Detection of coronary heart disease prior to clinical disease is expensive and invasive. Mammography is sensitive but not very specific. Unfortunately, mammographically-detected breast cancer may not be curable.

5. A primary prevention trial should begin before the disease is present. The average age of the population was 63 and a large proportion of the subjects were more than 10 years post-menopausal. 26% of women in the study had pre-existing heart disease. These factors seriously undermine the idea of WHI producing a primary prevention trial.

6. The trial should study within the same time frame as the disease studied.
Heart disease and breast cancer evolve over decades, but the trial was planned to be 8.5 years and was stopped at 5.2 years. Women more than a few years post-menopausal without hormone replacement therapy likely have significant plaques that cannot be reversed with HRT. Effects of estrogen on cholesterol, LDL, and HDL will need years to show protection from clinical disease. Data indicates that 10 years or more require the production of detectable breast cancer after initiation.

7. The population studied should be closely related to those who are most likely to be treated. Few women in this trial were studied early in menopause and symptomatic women were excluded. Treatment is most commonly prescribed early in menopause because of symptoms.

8. The trial should avoid contamination by previous treatment.

26% of women were using or previously used HRT. When HRT is used long-term a short washout period does not adequately address previous exposure.

The interpretation of results is subjective in general. However differing interpretation is more likely when the findings show small differences. The authors presented nominal and adjusted confidence intervals but chose to discuss only the nominal intervals. The adjusted confidence intervals are a more conservative approach to the data. When multiple comparisons are made, the adjusted confidence intervals decrease the risk of accepting Type I errors. The use of nominal intervals allows statistical significance of smaller hazard ratios and increases the chance of Type I errors. The choice of how conservative to view the data is arbitrary. The Writers Group chose to accept a greater chance of making a Type I error by choosing to discuss only the nominal intervals.

The hazard ratios found in this study were quite small. Of the many outcomes measured only thromboembolic disease showed a HR >2. The statistical significance for CHD is minimal. The HR for CHD was 1.29 with a nominal CI of 1.02-1.53 based on 286 cases. Of perhaps greater importance is that with the exclusion of women with pre-existing heart disease there is no statistical significance (HR 1.28, CI 1.00-1.59). Similarly, the hazard ratio for breast cancer was very small (HR was 1.26 with a CI of 1.00-1.59 with 290 cases). The statistical significance depended not on an increase in breast cancer in the treated group but in a decrease in cancer cases in the placebo group in year 4. The small HR and the inclusion of 1.00 in the CI decreases the confidence that there is causality in the increase seen in breast cancer. In addition, the absolute risks are similarly small (CHD absolute risk increase of 7 events per 10,000 women/years and breast cancer absolute risk increase of 8 events per 10,000 women/years). One must consider that 583 subjects were lost to follow-up and that the drop-out and drop-in resulted in greater than 50% loss of study participates.

No absolute determination of how much difference is important can be made. Many have trouble accepting such HRs as evidence of causality. However, the users of the data (health care provider, patients, and regulators) will utilize their own experience, fear, prejudices, and political background in interpretation.

The responses from the scientific community, the media and the legal profession have been numerous. With the exception of certain venues seen by subspecialists in reproductive endocrinology and menopause, opinion have been uniformly that of complete acceptance of the opinions of the authors of WHI. If the treatment of the menopause was less emotional and political I believe that the WHI report would have been accepted more as just another piece of evidence rather than the final truth about treatment of the menopause.

The societies of obstetrics and gynecology, particularly the American College of Obstetrics and Gynecology, have reported the findings of the WHI as definitive (4), although more recent publications have shown a more balance view (5). Observational studies, studies of animal and human basic science of the effects of estrogen deficiency and the clinical experience of many clinicians have often been swept away. It is not surprising that the media report simplifications which negate the subtleties of science when our societies choose such a simplification. Each report copies the statements of previous reports and the science is viewed as a much firmer foundation than the data deserves. The plaintiffs' bar has responded...
vigorously. A simple search of Google.com with the search words "Prempro and Legal" resulted in 46 pages in which attorneys offer to evaluate the client's case for litigation if she took Prempro. The politics of the menopause are complicated but probably predicts the response to the WHI. Society wants questions answered quickly and cheaply that trials can't provide. The trial that is needed would require decades and would cost even more than the approximately $800 million that was spent on the WHI. The exclusion of symptomatic women is probably necessary to include a placebo so that the most sensitive population cannot be studied. There is widespread fear of hormones among lay women and health professionals. In a society in which substantial numbers of people view hormones as "bad" a regulatory agency will respond to soothe the fear of that population. The Federal Drug Administration of the United States of America will err on the side of safety. The response of the FDA was the "Black Box" warning on the basis of WHI. The FDA accepted the results of the WHI as the definitive statement. The label advises that prescription of estrogen and progestin should not be prescribe for the purpose of protection from heart disease. The label also advises that estrogen and progestin be prescribed at the "lowest effective dose for the shortest duration" which has no scientific basis in general and the question of dose or duration of treatment was not addressed by WHI. Opinions from societies such as ACOG and governmental bodies such as FDA are produced by consensus to be safe and non-controversial. These bodies cannot ignore the worries of society and their opinions are placed in that context. One also cannot ignore the importance of competitors in the marketplace. WHI studied only one product produced by a single company. The competitors might see an opportunity to increase their market share. Similarly, the producers of SERMs and statins are clearly the winners of the wholesale acceptance of the conclusions of WHI. Those who have never accepted menopause as an endocrinopathy will be happy with these results. There are those who will believe that women will magically alter their diet and begin regular exercise because of the publicity about WHI and its apologists. Finally, the media has found a simplistic and sensational evaluation story which is useful for its business. It is naïve to think that science exists in a pure setting without societal pressure. Perhaps it is wrong and dangerous to ignore societal pressures, as ultimately, it is society that asks the questions for scientists and pays for the quest to answer those questions. I fear that the acceptance of such weak science as the WHI will result in harming women that will not be seen for decades.

REFERENCES

4. Questions and Answers on Hormone Therapy. ACOG.COM 12/19/02.
5. Report to Women's Health Initiative Study results by the American College of Obstetricians and Gynecologists ACOG.COM 10/22/03.