The Heart and Estrogen/Progestin Replacement Study Revisited

Hormone Replacement Therapy Produced Net Harm, Consistent With the Observational Data

John A. Blakely, BA, MD, FRCPC

Lower coronary event rates in women receiving hormone replacement therapy (HRT) have led to a presumption of benefit. The Heart and Estrogen/Progestin Replacement Study, a large randomized trial, observed a 1.4% first year excess of coronary events, well beyond the plausible play of chance on the expected effect. Over the duration of the study, event totals were similar, but patients treated with HRT experienced them earlier, with a net loss of patient-months of event-free survival. The point at which the lower event rate in hormone-treated patients would fully repay the first year loss, with constant rates, is almost double the trial duration (of 4.1 years). Since patients in the trial were preselected for satisfactory adherence to therapy, the net benefit in practice is likely to be even less. Had the patients in the Heart and Estrogen/Progestin Replacement Study been recruited to an observational study at various intervals over the first 5 years after starting HRT, the apparent risk reduction over 5 years would have been between 21% and 34%. A previous meta-analysis of trials of HRT for other indications also shows net harm. Women with or at high risk of coronary heart disease should not start HRT. There is a risk that women without coronary heart disease might experience even greater net harm from HRT. The late benefit is necessarily limited, as it cannot exceed the event rate. The mechanism of the early loss is unknown; if it were reduced proportionately less than the late benefit, considerable net harm could result.

From the Department of Cardiology, Sunnybrook and Women’s College Health Sciences Centre, Toronto, Ontario.

Hormone Replacement Therapy: The Observational Evidence

Observational studies have consistently shown that the rate of coronary heart disease (CHD) in postmenopausal women who take hormone replacement therapy (HRT) is only half to two thirds that in women who do not take HRT. Such studies would be expected to overestimate benefit, as compliant patients and women who take HRT tend to be healthier than those who do not, but the effect seems too large to be explained by “prevention bias.” Although an association is not in itself evidence of causality, these studies have resulted in the widespread prescription of HRT to postmenopausal women in the expectation of cardiovascular benefit.

The Trial

The Heart and Estrogen/Progestin Replacement Study (HERS) was a prospective, randomized trial of HRT in women with CHD that was undertaken to provide definitive evidence of the causal relationship between the administration of HRT and low rates of cardiac events in postmenopausal women. HERS appears to have been a powerful, well-designed and well-executed trial in which 2763 women participated for a mean duration of 4.1 years. The total number of nonfatal myocardial infarctions and cardiac deaths was almost identical in the treated and the untreated groups. However, the group receiving HRT experienced a 50% higher car-
The important issues are the credibility of the trial and its result, the appropriate analysis of trials of “present pain for future gain,” the consistency of the trial results with the observational evidence, and the relevance of the HERS results, obtained in older women with CHD, to a younger and lower risk population.

CREDIBILITY OF THE TRIAL RESULT

The trial appears to have been well designed and well executed. No patients were unavailable for follow-up, and treatment crossovers were not excessive. There was an initial period of monitoring of adherence to treatment during treatment with placebo alone, with enrollment restricted to those taking at least 80% of prescribed medication. Adherence to treatment was likely a good deal better than in practice.

Lipid lowering by HRT resulted in a few more patients taking statins in the control group (22% vs 18%), but the low- and high-density lipoprotein values were the same at entry, and better in the HRT group at the end of 1 year’s treatment. Triglyceride levels were higher by 0.04 mmol/L (4 mg/dL) at entry and 0.11 mmol/L (10 mg/dL) at 1 year in patients who were taking HRT; however, to attribute the result to lipid levels would require that a 5% higher triglyceride level produce a 50% higher 1-year event rate in the face of a 10% advantage in both low- and high-density lipoprotein values. None of these factors seems a plausible explanation for a 50% increase in early risk, particularly in the face of the authors’ expectation of a 24% reduction in risk. There are no obvious grounds for dismissing the trial result on the basis of its design, execution, or analysis.

METHODS

ANALYSIS OF “CROSSED CURVE” TRIALS

HERS resembles a trial of surgery in which “present pain produces later gain.” A patient dying in the first year loses more survival than that gained by a patient who is saved later, and, at the point where the event rates are equal, there is a net loss of patient-years of event-free survival in the treated group. Eventually, the early loss may be repaid by an equivalent later gain, but the time required for repayment may exceed the trial duration, despite a favorable late event gradient. In HERS, HRT did not just fail to help. Over the period of observation, the treated group experienced net harm.

The HERS result has been reanalyzed incorporating the value of time, using patient-months of event-free survival as the unit of effect. Net treatment effect has been extrapolated past the period of observation, assuming constant event rates. For this analysis, events in the first year have been assumed to be evenly distributed throughout the year, and the mean effect after year 1 has been used for the remainder of the trial. Early and late events have been assumed to be of similar average severity. Because of the nonrecurring early losses, all estimates of net benefit are specific to the duration of calculation.

One could argue that, because not all the patients took all the medicine all the time, the true value of HRT has been underestimated. Although it is certainly true that treatments work better when they are taken, the issue of adherence to treatment worked the other way in HERS. Because women in HERS were subject to screening before the study, in which all potential recruits were given placebo and only those adhering satisfactorily were recruited, adherence to treatment is unlikely to be as high in practice as it was in HERS. This means that the larger number of women beginning but not continuing HRT would be exposed to the early risk but not to the later benefit, and the net benefit in practice would be proportionately less than was observed in the HERS trial.

RESULTS

The excess of events in the first year in the treated group was 1.4%, a relative hazard of 1.5, large enough that chance is an unlikely explanation (57 vs 38, \(2P = .046\)). However, the nominal \(P\) value is hardly the issue, since the expectation was for a lesser, not an equal, number of events. The observational studies suggested as much as a 50% reduction, and the trial was designed to detect an intention-to-treat reduction of 24%. The design anticipated a 1-year delay in half the benefit, so that the expected result at the end of the first
year can be construed to be a reduction in primary events of 12%. If this were the “true” value, a result as contrary as that observed would occur only once in 258 such trials. To produce the expected net benefit of 24% over the trial duration, the late event rate in the treated group would have had to be about 62% less than was observed. Over the entire trial, ignoring the earlier occurrence of events in the HRT group, the play of chance on a true relative risk reduction of 24% would produce at least the overall HERS result, on average, only once in 104 trials. Neither the first-year excess nor the overall HERS result can plausibly be dismissed as the play of chance on the expected effect.

Observation of the patients in HERS for 4 years, beginning 3 years after the initiation of therapy, results in an apparent risk reduction of 34%. If 25% of the group were observed during initiation of therapy, the apparent risk reduction would be 21%.

In a “present pain for future gain” trial, losses that are attributable to the higher initial event rate treated may eventually be repaid by the favorable rate of later events. In HERS, with constant event rates, the initial net loss would be repaid after 6 to 7 years. The trial was designed around an expected benefit of 24% over 4 years. (The expected “true” benefit used in the power calculation was 35%, but, after crossovers and dropouts, a benefit of 24% was anticipated in the trial.) If event rates were to remain constant, the estimated net benefit at 10 years would be about 1 month of event-free survival per patient randomized to HRT. This estimate is subject to both statistical and biological uncertainties. The confidence interval on the benefit observed during the trial, extrapolated to 10 years, is 4.7 to –2.7 months per patient; there are too few observations for a “statistically credible” 10-year estimate. The biological uncertainties are even more important. The mean period of observation in HERS was only 4.1 years, and some effects of HRT, such as malignancy, may appear only after a latent period.

Of course, not all patients would experience benefit; if the estimate is correct then, over 10 years, roughly 4% of the patients receiving HRT would receive an average benefit of 14 months each. But eventual benefit is contingent on event rates remaining constant into the future, and the estimate is far from reaching “statistical significance.”

Had HERS been an observational study, the modeling exercise suggests that as the proportion of patients whose first year of therapy was included decreases from 25% to 0%, the apparent relative risk reduction over 4 years of observation rises from 21% to 34%. With due allowance for the absence of selection bias, the results of HERS are consistent with those of the observational studies when applied to these patients.

Most observational studies would only include patients actually taking treatment. The effect on the HERS conclusions of lesser adherence to treatment in practice has been roughly estimated. If, in the model observational trial, the population on treatment after year 1 were to be reduced by 20%, the already uncertain benefit extrapolated to 10 years would be reduced by 50%, to an average of 2 weeks per patient randomized to HRT. Many women prescribed HRT stop treatment after a few years.7

---

**COMMENT**

**PRINCIPAL FINDINGS**

HERS, the first large prospective study of HRT on cardiovascular disease, did not just fail to demonstrate the benefit expected from observational studies. The treated group experienced net harm, with fewer patient-months of event-free survival in those taking HRT. Contrary to intuition, this has been demonstrated to be consistent with the observational data.

**STRENGTHS AND WEAKNESSES OF THIS ANALYSIS**

**Strengths**

Benefit, in trials involving “present pain for future gain,” is a function of the relative loss of patient-months of event-free survival in the groups, not of the proportion of patients who survive to the end of the trial. The treated group experienced net harm, a net loss of patient-months of event-free survival. The issue is whether the harm observed in the study would be expected in practice, or whether it can be explained away by defects in the design, conduct, or analysis of HERS.

Various rationales for ignoring the HERS results have been advanced. It has been pointed out that the initial increase in events was not “statistically significant,” that not all patients were taking acetylsalicylic acid or β-blockers, that statin use was greater in the control group, that the authors confined their conclusions to patients similar to those in HERS, and that the HERS data are in contrast to the extensive data from the observational studies.8

But “statistical significance” vs the null hypothesis is hardly the issue. The initial harm is sufficiently large that it cannot plausibly be attributed to the play of chance on the expected effect. The rationale for modification or reversal of the effect of HRT by omission of acetylsalicylic acid or β-blocker therapy is unstated, and it is unlikely that a higher proportion of patients in the community or in the observational studies were taking these medications. A review of the lipid data in HERS shows that the higher consumption of statins in the control group was a reflection of their less favorable lipid profile, and is not a plausible cause of their unexpectedly better survival. That the authors refrained from considering the possible implications of their data for patients at lower risk is hardly in itself valid criticism of further analysis or comment. And reanalysis of HERS as a typical observational study shows that there is no discrepancy between the results of HERS and those of previous observational studies, when allowance is made for exclusion from observational studies of patients who are beginning therapy. Also, the results of HERS are consistent with the previous prospective data; a published meta-analysis of the existing prospective HRT trials that was performed to evaluate other effects shows almost the same increase in risk as observed in the first year of HERS.9

**Weaknesses**

The results of HERS were unexpected. While perverse results are
commonly analyzed by the same methods as those predicted, no “statistical” analysis can establish a causal relationship between HRT and the observed early increase in cardiac events. One cannot conclude that harm from HRT has been proved. But one can conclude that the presumption is unwarranted that the observed association necessarily implies benefit.

STRENGTHS AND WEAKNESSES IN RELATION TO OTHER STUDIES

Strengths

The authors of HERS, and commentators since,10 allude to the HERS results, and comments imply benefit. The authors of HERS, and commentators since, allude to the HERS results, and comments imply benefit. Since the HERS result was unexpected, it constitutes an observation, to be tested prospectively. The-wow-HRT and similar trials that are assessing disease progression by angiography will be sensitive only to the late effects of HRT, rather than to the net benefits. The issue is likely to be resolved with the report of the Women’s Health Initiative, an ongoing randomized trial of estrogen and combined HRT in the primary prevention of CHD, which is expected in 2005. Until then, it would be prudent to consider the probability of an increase in cardiovascular risk, rather than benefit, in the prescription of HRT, for any purpose.

Accepted for publication May 25, 2000.

The author wishes to thank Barbara Huth for critical review of the manuscript and for assistance in its preparation.

Corresponding author: John A Blakely, BÁ, MD, FRCP, Sunnybrook and Women’s College Health Sciences Centre, 2075 Bayview Ave, Room H121, Toronto, Ontario, Canada M4N 3M5 (e-mail: john.blakely@utoronto.ca).

REFERENCES