More than 100 yr ago, Osler (1) noted that heart disease was almost entirely a disease of men. Fifty years ago, the most popular explanations for sex differences in heart disease were lifestyle, particularly cigarette smoking (which was mainly a male habit until World War II), or differences in stress (i.e. men in the workforce and women at home, or what I like to call the “happy homemaker hypothesis”) (2). Although behavior and occupation differences may play a role, neither has emerged as a satisfactory explanation for women’s favored cardiovascular status.

As shown in Fig. 1, women in every country, whether heart disease rates are high or low, are at lower risk of fatal coronary heart disease (CHD) compared with men, despite diverse lifestyles, diets, and workplace options (3). Indeed, within countries women have less heart disease than men when stratified by similar levels of classical heart disease risk factor levels. In the Renfrew and Paisley Survey (4), for example, sex-specific coronary death rates separately adjusted for cholesterol, blood pressure, body mass index, cigarette smoking, and social class showed that women had lower absolute rates at every age, although the relative risks associated with these risk factors were similar. This is also true for one factor that differs by sex, i.e. women’s higher high-density lipoprotein (HDL) cholesterol levels; within the same HDL strata, women have less heart disease than men (5).

These results between and within populations and consistent across risk factors support the hypothesis that women are protected by an intrinsic factor, presumably female sex hormones. Several types of epidemiological evidence suggest that women’s universal protection against CHD is explained by estrogen (6, 7). The main observations are that CHD is 1) uncommon in women before the age of menopause, 2) more common in young women who have had both ovaries removed, and 4) less common in women who take hormone therapy (HT) after menopause.

Population-based evidence that heart disease death rates increase after menopause is actually weak. An exponential increase in heart disease rates by age as observed in standard plots becomes a step-wise linear association when heart disease death rates are plotted on a semilogarithmic scale. There is no evidence of a different slope around age 50 yr, the usual age at menopause, as shown in Fig. 2 (8). As also shown in Fig. 2, this is in striking contrast to the midlife change in the semilogarithmic slope for breast cancer, an estrogen-dependent disease. These results could mean that estrogen is not such a primary player in heart disease etiology as it is for breast cancer, or that heart disease has a much more multifactorial etiology, making it more difficult to observe estrogen’s central role, or that estrogen is not the intrinsic factor.

Some prospective epidemiological studies of individual women have shown small albeit significant associations of cardiovascular disease death with age at natural menopause (9, 10). These types of studies of menopause-CHD associations could be weak because the last menstrual period may not be accurately noted or because the lag to clinically manifest CHD is highly variable, because coronary disease has many other contributing risk factors. Note that these associations could also be an artifact due to confounding by cigarette smoking, the commonest cause for premature menopause and a powerful CHD risk factor.

If premenopausal estrogen levels are cardioprotective, then premature menopause, with fewer years of exposure to premenopausal estrogen levels, should magnify risk. As reviewed elsewhere (6), many autopsy studies have demonstrated an excess of coronary artery atherosclerosis in young women who have had an oophorectomy. However, these observational studies of surgical menopause could be confounded by each woman’s reason for or reaction to early menopause, the reason for the autopsy, and possibly by the concomitant loss of other hormones (e.g. testosterone) after oophorectomy.

Theoretically, studies of endogenous estrogen levels and
CHD should offer direct evidence for a protective effect of estrogen on heart disease. With a natural menopause, estradiol levels typically begin to decline 1 yr or less before the last menstrual period, although levels may fluctuate widely during this late transition period (11). Prospective studies do show a positive association of endogenous estrogen levels and breast cancer in postmenopausal women (12). No studies have shown that endogenous estrogen predicts cardiovascular disease in postmenopausal women (13). One reason for the null cardiovascular results may be limitations of earlier estradiol assays, which were below the level of assay sensitivity in up to one half of postmenopausal women in some studies (14). Postmenopausal estradiol levels also show relatively high intraindividual variation, which reduces the ability of single measurements to characterize a woman’s usual estrogen level in epidemiological studies (15).

Studies in postmenopausal women do not exclude the possibility that premenopausal levels are above some threshold necessary to protect against CHD. One small cross-sectional study reported that premenopausal women who had more severe coronary artery disease at angiography also had significantly lower levels of circulating estrogen (16). Postmenopausal estradiol levels also show relatively high intraindividual variation, which reduces the ability of single measurements to characterize a woman’s usual estrogen level in epidemiological studies (15).

In a 1991 review, Barrett-Connor and Bush (6) reported that 11 of 24 cross-sectional and prospective studies showed a statistically significant reduced risk of CHD among women using oral conjugated equine estrogen (CEE) taken without a progestin. In this review, we wrote, “Clearly, the weight of the evidence at this time points toward a substantial reduction in CHD risk among women using estrogens. Nevertheless, it is important to recognize the limitations of the data on which this statement is based. All but one small study are observational. There is no information on why women were prescribed estrogen and no ability to contrast these women...
with others who may have never been offered or refused or stopped hormone replacement therapy. Overall, women who take estrogen after the menopause are more likely to be white, educated, upper middle class, and lean, thereby at lower risk of heart disease than women without estrogen replacement therapy.” We also wrote, “Although we have used the phrase ‘replacement therapy’ in this review, this is pharmacological, not physiological, therapy. Oral estrogens are drugs. Thus, a statement recommending hormone replacement as a method of heart disease prevention for postmenopausal women would seem to warrant a clinical trial, as required for other drugs” (6).

In the same year as our 1991 review, a Food and Drug Administration (FDA) Advisory Committee voted almost unanimously in favor of an industry request for “an indication,” i.e. permission to say that estrogen replacement therapy could be recommended to reduce the risk of CHD in postmenopausal women (Table 1). It is not clear why this recommendation was never acted on by the FDA, despite the appearance of additional observational studies suggesting cardioprotection, coupled with an ever-increasing number of biologically plausible mechanisms making advocacy almost irresistible (19).

In 1992, Grady et al. (20) published a landmark paper in the Annals of Internal Medicine using meta-analyses of observational studies from the published literature to estimate differences in four disease rates by hormone use. These summary risk estimates showed that postmenopausal hormone use was associated with about one third less fatal heart disease and calculated that this benefit would prevent more deaths than the combined increased risk of death due to breast and uterine cancer because heart disease is a much more common cause of death (Table 2) (20). Based on this analysis, the same issue of the Annals of Internal Medicine published a position statement from the American College of Physicians (21) proposing that all postmenopausal women should be offered HT to prevent heart disease. Other prominent professional organizations, including the American College of Obstetrics and Gynecology and the American Heart Association, soon followed with similar recommendations.

The last meta-analysis of observational studies designed to test the hormone-heart disease hypothesis, and published before results of clinical trials in women, was based on all published observational studies through mid-1997 (22). Among 25 observational studies, most from the United States where unopposed CEE was by far the predominant regimen, the summary estimate of the relative risk for CHD for women who ever used estrogen compared with never users was 0.70 [confidence interval (CI), 0.65–0.75] (Fig. 3). A similar relative risk [0.66 (0.53–0.84)] was observed in the seven studies that specifically reported treatment with estrogen plus a progestin, usually cyclic medroxyprogesterone acetate (MPA) (Fig. 4).

Clinical trials

The first clinical trials of the effects of a medication on CHD are often performed in persons with known heart disease who have a very high risk of recurrence and who stand to receive the most benefit; these so-called secondary prevention trials reduce the number of subjects needed and the duration and cost of the trial. Therefore, the first large trial designed to examine the cardioprotective effect of estrogen was a secondary prevention trial, conducted in persons with heart disease who were at high risk of a new event. Based on their high risk, the persons selected for this estrogen trial were men with known CHD.

Coronary Drug Project

The Coronary Drug Project (23, 24) begun in the 1960s was the first clinical trial designed to determine whether estrogen reduced the risk of coronary events. Men with known heart disease were randomly assigned to one of five active therapies or placebo. Two of the study medications were CEE at a daily dose of either 2.5 or 5.0 mg; the estrogen arms were stopped early because estrogen-treated men had an increased rate of thromboembolic events and myocardial infarction (MI) (23, 24). After estrogen failed to protect men, no large estrogen trials with CHD outcomes were initiated in men or women for the next 23 yr.

Despite the absence of clinical trial data, by the mid-1990s it was almost dogma that HT would prevent CHD in postmenopausal women. Documentation that all postmeno-
pausal patients had been offered estrogen was one of the
criteria used to evaluate the quality of medical practice. Not
to recommend estrogen therapy was thought to be unethical.

**TABLE 1. HT timeline**

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
</table>
| 1920–39 | ● Searle introduced estrogen patch for symptoms (1928)  
           ● Estradiol synthesized (1938) |
| 1940–49 | ● Premarin introduced (1942)  
           ● Albright finds the low estrogen-osteoporosis link |
| 1950–59 | ● HT prescribed (but not often) for hot flashes  
           ● Wilsons promote HT to prevent “the tragedy of menopause”  
           ● Oral contraceptives introduced to regulate menses and prevent pregnancy  
           ● Clinical trial of estrogen and heart disease begun in men |
| 1970–79 | ● Men’s HT trial stopped because of early excess clotting and cardiovascular disease  
           ● Side effects of oral contraceptives recognized, especially clotting and strokes  
           ● Excess endometrial cancer risk with estrogen recognized  
           ● Progestins added to estrogen therapy in women with an intact uterus |
| 1980–89 | ● Deluge of epidemiological studies suggesting that HT reduces heart disease |
| 1990–94 | ● Meta-analysis suggests that heart benefit of HT would exceed possible risks  
           ● Many groups recommend that HT be offered to all women to prevent heart disease  
           ● FDA Expert Advisory Committee almost unanimously approves heart disease prevention label for unopposed HT (recommendation never activated)  
           ● PEPI trial of HT and CHD risk factors begins  
           ● Hulley obtains funding for HERS clinical trial (first participant 1993)  
           ● NIH sponsors WHI trial  
           ● Premarin is the most widely dispensed prescription drug in United States (1990–95) |
| 1995–99 | ● PremPro, the first combination HT pill, introduced  
           ● PEPI results confirm improvements in LDL and HDL cholesterol  
           ● HERS (1998) reports early increased heart disease risk |
| 2002–03 | ● WHI reports increased heart disease, stroke, and breast cancer, the small risks exceed smaller benefits  
           ● WHI continues unopposed estrogen arm  
           ● FDA requires black box for all postmenopausal estrogens with or without progestin |

**TABLE 2. Estimated risk for common diseases in a 50-yr-old white woman with a uterus and no known increased risk**

<table>
<thead>
<tr>
<th>Condition</th>
<th>No treatment</th>
<th>Estrogen + progestin&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Estrogen + progestin&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart disease</td>
<td>46</td>
<td>34</td>
<td>39</td>
</tr>
<tr>
<td>Hip fracture</td>
<td>15</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>10</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>&lt;3</td>
<td>&lt;3</td>
<td>&lt;3</td>
</tr>
<tr>
<td>Life expectancy</td>
<td>82.8</td>
<td>+1.0</td>
<td>+0.3</td>
</tr>
</tbody>
</table>

<sup>a</sup> Optimistic view;  
<sup>b</sup> pessimistic view.

Fortunately, some still believed that a clinical trial with heart disease as the primary outcome was necessary. Special credit should go to Stephen B. Hulley, who obtained funding for the secondary prevention trial that came to be known as the Heart and Estrogen/Progestin Replacement Study (HERS), and to Bernadine Healy, then director of the National Institutes of Health, who obtained the support for the large primary prevention trial known as the Women’s Health Initiative (WHI).

**HERS (25)**

Begun in 1993, HERS was the first large clinical trial specifically designed to evaluate whether estrogen plus progestin therapy reduced the risk for CHD events in postmenopausal women with established coronary disease. The original proposal, submitted to the National Institutes of Health/National Heart, Lung and Blood Institute by Hulley in 1990, was a secondary prevention trial designed to study both unopposed estrogen and estrogen plus a progestin in women who already had heart disease. After the NIH declined the grant application, Wyeth-Ayerst Research agreed to fund the study, with Hulley as the principal investigator. Wyeth made one critical modification: only women with a
utens were included, apparently because Wyeth expected to receive FDA approval for a statement indicating that unopposed estrogen was cardioprotective (Table 1).

HERS was a multicenter randomized, double-blind, placebo-controlled trial that enrolled 2763 U.S. postmenopausal women (mean age, 67 yr); eligibility required an intact uterus and documented CHD. The HERS intervention was a single daily tablet containing CEE (0.625 mg) and MPA (2.5 mg) or placebo. The study closed a little ahead of schedule in July 1998, after an average follow-up of 4.1 yr. HERS results, published the same year (25), showed no overall difference in the primary CHD outcome (nonfatal MI and CHD death combined) between the HT and placebo groups. Nonfatal MI or CHD death occurred in 179 women in the hormone group and 182 women in the placebo group (relative hazard, 0.99; 95% CI, 0.81–1.22) despite significant lowering of low-density lipoprotein (LDL) and increase in HDL cholesterol in the HT group ($P < 0.001$).

Not only was the overall result null, but during the first year there was a statistically significant 52% excess risk of CHD events in the HT group (25). Benefit after longer use was suggested by a highly significant ($P = 0.009$) cardioprotective test for trend over time, based on an analysis that included the excess first year event rate. After the exclusion of the first year estrogen-induced, excess risk data, however, the test for trend was not significant ($P > 0.45$). An additional 2.7 yr of follow-up (with about half of the original cohort continuing HT) also failed to show any evidence of long-term cardiovascular benefit (26). Two secondary cardiovascular outcomes (stroke and peripheral arterial disease) also did not differ between HERS treatment groups (27, 28).

Meetings were held, and papers were written trying to explain these unexpected results. The most common complaints were that the HERS women were too old or too sick. In fact, HERS women had an average age of 67 yr, a common age for women in secondary prevention trials demonstrating the benefit of lipid-lowering medications. They were also not very sick. Most had had revascularization rather than a heart attack; women were excluded from randomization who had had a cardiac event within 6 months, New York Heart Association class IV or severe class III congestive heart failure, uncontrolled hypertension, uncontrolled diabetes, or a history of venous thromboembolic disease. The HERS participants were shown to be similar to a representative sample of U.S. women with coronary disease (29).

It was also argued that too many HERS women were noncompliant with their study medications, but the 75% 3-yr adherence rate was considerably better than the continuation rates observed for HT use in the general population. Because there were fewer cardiovascular events than expected, it was also said that the study had too little power, yet the fairly narrow 95% CI values suggest that more than 20% benefit (or harm) was unlikely to have been missed (25). An extensive search for subgroups that might explain the early harm or overall null effect was conducted (30). Nine of the 172 tests for interactions were statistically significant—about the number expected by chance at $P < 0.05$—and none of these appeared to be explanatory. The authors provide a list of all subgroups examined in their publication, allowing the reader to see whether his or her favorite explanatory hypothesis (for either early harm or overall null effect) was tested.

Although the early increase in coronary death and nonfatal MI in HERS was completely unexpected by the investigators or the medical community, it was similar to the early harm observed among the men in the Coronary Drug Project (31).

Small secondary prevention trials of HT

At the time of its publication, the most valid criticism of HERS was that it was only a single trial, possibly administering the wrong HT regimen. However, several smaller secondary prevention trials described below have also reported no benefit after HT, and some suggest harm. Several studied estrogens other than CEE and regimens without a progestin.

Papworth HT Atherosclerosis Study (PHASE). Clarke et al. (32) reported an unblinded trial in 255 women with angiographically proven heart disease (mean age, 66 yr), who were randomly assigned to no treatment or a 17 $\beta$-estradiol patch (2.5 mg). The patch was administered alone every 4 d to women without a uterus or, for women with a uterus, administered over 14 d, followed by four patches containing 3 mg of 17 $\beta$-estradiol and 4 mg norethisterone. The primary endpoint was cardiac death, proven MI, or hospitalization for unstable angina. After an average of 31 months, the data and safety monitoring board recommended early closure, based on futility. The CHD event rate for the HT group was 15.6 per 100 patient years compared with 12.6 per 100 patient years in the control group. Although there were no significant differences between the groups in the frequency of any cardiac event, beginning in the first year the active treatment groups had higher rates of all events except nonfatal MI. Because a large number of women discontinued treatment, the inten-
tion-to-treat analysis underestimated risk; the as-treated analysis showed a rate ratio of 1.49 (95% CI, 0.93–2.36) compared with the lower intention-to-treat analysis rate ratio of 1.29 (95% CI, 0.82–1.86). As expected with transdermal estrogen, there were no significant differences in LDL or HDL cholesterol by treatment assignment.

**Estrogen in the Prevention of Reinfarction Trial (ESPRIT)** (33). In this secondary prevention trial, investigators randomly assigned 1017 women (mean age, 63 yr) who had survived a first heart attack to either 2 mg of unopposed estradiol valerate daily or placebo. After 2 yr, the frequency of reinfarction or cardiac death did not differ by treatment assignment [rate ratio, 0.99 (95% CI, 0.70–1.41)], and there was also no difference in all-cause mortality. There was no evidence of early harm. This trial had a very high dropout and crossover rate, which could have obscured benefit or harm.

**Estrogen Replacement and Atherosclerosis (ERA).** Herrington et al. (34) randomly assigned 309 postmenopausal women (mean age, 66 yr) to receive 0.625 mg of CEE alone (for women without a uterus), CEE with 2.5 mg of MPA daily (for women with a uterus), or placebo. Women were required to have had at least one coronary stenosis of at least 30% of the luminal diameter measured by quantitative coronary angiography. After an average follow-up of 3.2 yr, 248 women had a repeat coronary artery angiogram; the mean minimal coronary artery diameters did not differ significantly by treatment group, despite significant reductions in LDL cholesterol and increases in HDL cholesterol in the women assigned to HT. Patterns were the same with unopposed estrogen or the combined regimen.

**Women’s Angiographic Vitamin and Estrogen (WAVE).** In this placebo-controlled factorial design, Waters et al. (35) randomly assigned 423 postmenopausal women (mean age, 66 yr) to daily continuous combined oral CEE 0.625 mg and MPA 2.5 mg with or without vitamins E and C. Eligibility required a minimum of 15% coronary artery atherosclerosis on a baseline angiogram. After an average follow-up of 2.8 yr, a repeat coronary angiogram in 306 of these women showed somewhat greater progression in each active treatment group. In a preplanned analysis that assigned death, nonfatal MI, or stroke to the worst angiographic rank, the risk was nearly doubled in women assigned to HT compared with controls (HR, 1.9; 95% CI, 0.97–3.6), but these differences were not statistically significant.

**Postmenopausal hormone replacement and carotid atherosclerosis**

**HERS B-mode ultrasound substudy** (36). Five clinical centers recruited 454 women from the HERS cohort who agreed to have a carotid ultrasound at baseline. After 3.8 yr, a follow-up scan was obtained from 362 women. Intimal medial thickness (IMT) increased in both HT and placebo groups, with no overall difference in the primary outcome, but there was slightly slower IMT progression at the bifurcation (one of two secondary outcomes) with $P = 0.06$ favoring the HT treatment (36).

**Postmenopausal Hormone Replacement against Atherosclerosis (PHOREA).** Angerer et al. (37) enrolled 321 healthy women (average age, 66 yr) who had increased IMT in at least one segment of the carotid arteries. They were randomly assigned to 1 mg 17β-estradiol daily plus 0.025 mg gestodene for 12 d every month, or the same regimen with the same dose of gestodene taken only once every 3 months, or no HT (no placebo). The trial lasted only 48 wk, with 264 women available for a second carotid ultrasound. HT did not slow progression in the carotid arteries, despite a significant decrease in LDL cholesterol and fibrinogen levels.

**Primary prevention trials of HT**

To date, all lipid-lowering medications shown to prevent CHD in persons with prevalent heart disease have also shown protection in primary prevention trials, conducted in persons without known heart disease. Nevertheless, a reasonable concern was that estrogen therapy was too late in women who already had coronary atherosclerosis. In the Clarkson nonhuman primate model (38), CEE was protective only in animals treated before atherosclerosis.

**The Women’s Health Initiative (WHI)** (39). The WHI was designed to evaluate primary prevention. Begun in 1992, WHI included a randomized double-blind placebo-controlled clinical trial designed to evaluate the effect of three separate preventive strategies (HT, diet, and calcium supplements) on disease outcomes in healthy postmenopausal women aged 50–79 yr. In one HT arm of the trial, 16,608 women who had an intact uterus were randomly assigned to receive a single daily tablet containing CEE (0.625 mg) and MPA (2.5 mg) or placebo, the same regimen used in HERS. Another 10,739 women without a uterus were randomly assigned to placebo or CEE (0.625 mg/d) without MPA. The primary outcome was fatal and nonfatal heart disease, and stopping rules were established on the basis of predicted cardiovascular benefit; harm was not expected. Breast cancer was the primary adverse outcome with stopping rules. A global risk-benefit ratio was the third primary outcome. Study completion was scheduled for 2005.

After the first 2 yr of follow-up, and again 1–2 yr later, the Data and Safety Monitoring Board (DSMB) advised investigators that there had been an early excess risk of heart attack and stroke in women assigned to opposed or unopposed estrogen (40) (Table 1). This information was sent to all participants; most WHI women continued in the trial.

After an average of 5.2 yr, the combined estrogen-progestin arm was stopped on the advice of the DSMB, because the test statistic for invasive breast cancer exceeded the stopping boundary for the adverse event, the global index showed risks exceeding benefits, and there was no reason to expect future favorable effects on cardiovascular disease. Only 6 wk later, after all the WHI participants had been informed, the main WHI outcomes paper was published in the *Journal of the American Medical Association* (39). None of the excess risks or benefits was large (all less than 10 events per 10,000 women per year) as shown in Fig. 5. The hazard ratio for CHD was 1.29 (1.02–1.63) with 286 cases; this risk was apparent almost immediately. The hazard ratio for breast cancer was 1.26 (1.00–1.59) with 290 cases; this excess risk emerged after 4–5
Many still doubt that HT has been believed at all without the HERS data. Despite the global index was 19 per 10,000 women per year.

1.03 increased relative risk of negative outcomes of 1.15 (95% CI, 1.03–1.28). The absolute excess number of events included in the global index was 19 per 10,000 women per year.

It is doubtful that the WHI cardiovascular results would have been believed at all without the HERS data. Despite WHI and the other small trials, many still doubt that HT does more harm than good. Currently, one of the most common discussion points is that the women in HERS and WHI were too old (average age, 67 and 63 yr, respectively) and already had coronary artery atherosclerosis. In fact, the WHI data show that women between ages 50 and 60 yr had fewer cardiovascular events than women in older age groups, but the highest relative risk (Table 3). This is expected in that younger women have little CHD, so any excess risk will stand out clearly. The higher heart disease rates in older women make any small HT-associated CHD excess less obvious.

Postmenopausal Estrogen/Progestin Interventions (PEPI). Although the 3-yr PEPI trial (41) of 875 women who were within 10 yr of menopause was not designed to study disease outcomes, there was a (nonsignificantly) higher incidence of cardiovascular and thrombotic events among women assigned to HT (2.1 events/100 women) than in the placebo group, which had no events.

**TABLE 3.** WHI: MI/CHD death by age (65)

<table>
<thead>
<tr>
<th>Age (yr)</th>
<th>HRT</th>
<th>Placebo</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–59</td>
<td>33</td>
<td>19</td>
<td>1.67</td>
</tr>
<tr>
<td>60–69</td>
<td>68</td>
<td>51</td>
<td>1.26</td>
</tr>
<tr>
<td>70–79</td>
<td>63</td>
<td>52</td>
<td>1.18</td>
</tr>
</tbody>
</table>

HRT, Hormone replacement therapy; RR, relative risk.

yr. Increased risks of stroke (1.41; 1.07–1.85), apparent after about 2 yr, and pulmonary emboli (2.13; 1.39–3.25), which were apparent almost immediately, were also observed. In analyses adjusted for multiple outcomes and sequential monitoring, most of the adverse or beneficial effects were no longer statistically significant, but it will be recognized that this was an extremely conservative analysis. It is a common convention that primary outcomes in clinical trials (in this case heart disease, breast cancer, and global index) are interpreted without adjustment for multiple testing.

Only 400 WHI women had known cardiovascular disease at baseline. A subgroup analysis of these women, who had a history of heart attack or coronary revascularization at baseline, showed a CHD relative risk of 1.28, similar to that reported for the cohort overall. Multiple tests for interaction, including age, body mass index, prior hormone use, smoking, diabetes, aspirin or statin use, failed to show that any subgroup was selectively helped or harmed by HT with regard to heart disease, stroke, or venous thromboembolic disease.

The extensive media reports of these mostly harmful effects undoubtedly caused patients anxiety; many physicians were overwhelmed by telephone queries. Patients could be reassured because the excess risk of cardiovascular disease was small for an individual woman: the absolute excess per 10,000 women per year was seven more cardiac events and eight more strokes, but the benefits were even smaller. As shown in Fig. 5, the number of excess cardiovascular events, pulmonary emboli, and breast cancers in the hormone-treated group was larger than the number of prevented fractures or colon cancers. Overall, the global index showed an increased relative risk of negative outcomes of 1.15 (95% CI, 1.03–1.28). The absolute excess number of events included in the global index was 19 per 10,000 women per year.

**Fig. 5.** Effect of estrogen-progestin on event rates (adapted from the WHI HT Update 2000 (http://www.nhlbi.nih.gov/health/women/upd2002.htm). VTE, Venous thromboembolism.

<table>
<thead>
<tr>
<th>Additional Events</th>
<th>Benefits</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>Stroke</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Reduced Events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Estrogen in the Prevention of Atherosclerosis Trial (EPAT).** This randomized placebo-controlled clinical trial reported by Hodis et al. (44) was designed to test whether unopposed oral
micronized 17β-estradiol (1 mg/d) reduces progression in carotid artery atherosclerosis in healthy women (average age, 61 yr) without known cardiovascular disease who were an average of 13 yr post menopause. Women whose baseline LDL was at least 160 mg/dl were treated with lipid-lowering drugs. Of the 221 women randomized, 199 had a follow-up carotid artery ultrasound after 2 yr. In the group treated with lipid-lowering medication, progression of IMT was similar in women assigned to estrogen or placebo. The 77 women who had LDL levels below 160 mg/dl and received no lipid-lowering medication showed significantly less increase in carotid IMT on estrogen therapy compared with the placebo group (P = 0.002). The authors conclude that these results show that unopposed estrogen is cardioprotective if used in younger women before they have atherosclerosis. This interesting finding clearly stands alone among the clinical trial results published to date.

The negative reaction to the WHI (and other HT trials)

The subsequent FDA proscription that health care providers prescribe HT only for symptomatic women, and in as small a dose for as short a time as possible, goes against clinical experience and the feminine forever mystique. Flaws in the clinical trials design or execution have been sought that might allow a continuation of old familiar concepts and avoid a painful paradigm shift. Although both HERS and WHI met the classic criteria for a good clinical trial (randomization, placebo-controlled, and double-blind), several common criticisms remain, including the following.

The wrong estrogen. WHI and HERS and most of the smaller clinical trials were conducted in the United States and studied CEE. This most commonly prescribed estrogen therapy in the United States had been used by most of the women in more than 80% of the published observational studies in 1990 when WHI was being designed. Estradiol is more commonly prescribed in Europe, but there was no similar body of observational data pointing to heart disease prevention. If 17β-estradiol had been the WHI trial estrogen, with the same negative results for CHD, those who designed the WHI would have been asked, “Why not use the estrogen we know works?”

The wrong progestin. MPA was chosen for WHI because it was the most commonly used progestin with combined HT in the United States, and a few observational studies had suggested cardiac protection (Fig. 4) (22). Moreover, no observational studies have reported CHD risks for women using the continuous combined regimen used in WHI and HERS, which was chosen to improve compliance and minimize unblinding by reducing bleeding. It is possible that MPA (or possibly any progestin) masks cardioprotective effects of estrogen, and the WHI unopposed HT arm continues, although these women also received two letters advising them of an increased risk of cardiovascular disease.

The wrong age. Atherosclerosis down-regulates estrogen receptors, and the thesis that the women in WHI were too old for cardioprotection cannot be excluded, although WHI is by far the largest study of HT in young postmenopausal women ever reported. Perhaps age 50 yr is not young enough, because fatty streaks begin well before menopause. If cardioprotection requires beginning HT before atherosclerosis begins, a trial of HT in still-cycling women in their 40s would be necessary. The feasibility of such a study, larger and longer than WHI, is questionable.

Unblinding. In WHI (39), the study gynecologist was unblinded for about 40% of participants to make a decision about further evaluation for sustained vaginal bleeding. The study staff and the participant were rarely unblinded, however, and those who adjudicated possible CHD events were blinded to participant symptoms and treatment assignment. In any event, diagnostic suspicion bias would be expected to increase the expected outcome, not to detect an unexpected outcome such as the observed early excess CHD rates.

Dropouts. In the WHI only 3.5% of participants were lost to follow-up (39); dropout refers to discontinuation of study medications, not loss to follow-up. Discontinuation rates were high, 42% in the HT arm and 38% in the placebo arm. When there is a high discontinuation of study medication, intention-to-treat analysis tends to underestimate both the benefit and the risk of intervention. This was demonstrated in WHI, where the relative risk for heart disease was 1.29 with intention-to-treat analysis and 1.51 with as-treated analysis.

Generalizability. Women in observational studies usually represent a broader range of age, social class, and ethnicity than women in trials, and they are more likely to receive individualized treatment. Also, it is assumed that women in observational studies often started HT for severe menopause symptoms, whereas highly symptomatic women are usually excluded from clinical trials. There are, however, no data showing that highly symptomatic perimenopausal women are at increased risk of CHD or would be more protected by HT than their less symptomatic peers who volunteer for trials.

Back to Epidemiology

Observational epidemiological studies provide the opportunity to compare the association of an exposure variable (in this case HT) with a disease outcome (in this case CHD). Nearly all major modifiable CHD risk factors for heart disease were first discovered using observational studies. Prevention trials depend on epidemiological data for part of their rationale (animal models alone would not provide sufficient evidence), and to estimate the effect size of the proposed intervention, which is used to determine feasibility and calculate sample size. Clinical trials are done because even large prospective studies show associations but do not prove causality. Only clinical trials use randomization to control for the known and unknown biases that plague observational studies.

In the absence of trial data, we should evaluate possibly causal associations in epidemiological studies using standard criteria, which include biological plausibility, consistency and strength of association, specificity, temporality, dose-response, and prevention or reversibility. To help understand unexpected trial results, we can now ask how well
observational studies met these criteria and dealt with problems of confounding and bias.

**Biological plausibility**

Biological plausibility is estrogen’s strongest suit; *in vivo* and *in vitro* studies have demonstrated that estradiol has multiple potentially beneficial effects at the molecular and cellular level, and in animal models, as reviewed elsewhere (45, 46). Small clinical trials studying CHD risk factors and vascular reactivity also support HT cardioprotection, although intermediate effects do not necessarily predict disease.

Among the first identified potentially favorable effects of HT were an increase in HDL-cholesterol when estrogen was given alone, a decrease in HDL when estrogen was given with androgenic (19-nor) progestins, and an intermediate effect when estrogen was given with a nonandrogenic 17-nor progesterin (47). In the PEPI (41) study, 875 healthy postmenopausal women who were within 10 yr of menopause were randomly assigned to one of five treatment regimens for 3 yr. Treatments were placebo, CEE; CEE + cyclic MPA; CEE + daily MPA; and CEE + cyclic micronized progesterone (MP). All active PEPI treatments significantly reduced LDL cholesterol and increased HDL cholesterol and triglycerides compared with placebo. CEE alone or with MP raised HDL significantly more than either CEE + MPA regimen. In a separate analysis restricted to women who were able to continue 80% of their study medication, unopposed CEE was associated with a significantly greater increase in HDL than in women adherent to CEE + MP, although CEE + MP remained superior to either CEE + MPA regimen (48). These lipid effects are now thought to represent a first-pass effect through the liver following oral HT, and they seem unlikely to explain the HERS and WHI CHD results.

**Consistency and strength of association**

Consistency means that different studies find the same thing. Strength of association refers to the magnitude of the risk estimate and its 95% CI values. As shown in Figs. 3 and 4, nearly all of the individual comparisons of HT in observational studies suggested benefit based on the point estimates but had wide CI values that usually included one. Taken individually, these studies are not impressive with regard to the strength of the association. However, when presented as a pooled meta-analysis, the summary estimate now excludes one and has narrow CI values, compatible with significant pooled effects. By increasing the sample size, we can show a more precise estimate of the treatment effect, narrowing the 95% CI values, and at the same time visualizing the consistency of the results.

The meta-analysis was originally designed to combine data from small clinical trials with similar but nonsignificant results and is a very useful tool for summarizing clinical trial data. In observational studies, particularly those of self-selected interventions to prevent disease, the meta-analysis may provide only consistent significant evidence of bias, because treatment was not randomly assigned (49). All sorts of HT-user biases, particularly prevention bias, have been known for years and obscured by the enthusiasm for the HT-CHD hypothesis, the verity of which was sanctified by meta-analyses (19, 50).

**Prevention bias.** Women taking estrogen tend to have more education, more money, more favorable lifestyles, better levels of several heart disease risk factors, and less diabetes than untreated women. Some or all of estrogen’s putative CHD benefits could have been spurious, reflecting “the healthy wealthy woman” bias (50). Education and social class are strongly, independently, and inversely associated with the risk of CHD in both men and women (51). This type of bias for HT was elegantly demonstrated in a prospective epidemiological study by Matthews et al. (52). Following 355 women through the menopause, they found that those who elected to take HT after the menopause had more favorable levels of HDL cholesterol, fasting insulin, and blood pressure; and they reported more physical activity, alcohol intake, and education before the menopause than untreated women. Similar healthy wealthy self-selection bias may also be the explanation for the apparently cardioprotective effects of antioxidant vitamin supplements reported in large observational cohort studies, but not confirmed in clinical trials (e.g. Ref. 35).

Could social class bias explain the HERS and WHI CHD results? Very telling is a meta-analysis published 1 wk after the WHI by the U.S. Preventive Services Task Force (53). In this meta-analysis of 22 good-to-fair-quality observational studies (of 43 published studies), pooled data from the five studies that controlled for social class, education, or occupation showed that the previously observed reduced risk for coronary artery disease among HT current users was no longer present (Fig. 6). The authors also noted that no benefit was observed in the fair-to-good-quality studies that adjusted for alcohol use and physical activity.

Prevention bias also relates to physicians’ prescribing practices. Under the U.S. private payer health care system, women without money or health insurance are less likely to use their limited resources for health care and to be prescribed medication for prevention. In addition, until recently, sick women were less likely to be prescribed estrogen because hypertension, diabetes, and heart disease were listed as contraindications on the estrogen package insert. Women without these diseases were less likely to have a heart attack, a benefit attributed to HT.

**Compliance bias.** In a representative study of U.S. women, 45% had used estrogen for at least 1 month, but only 20% continued HT for 5 or more years (54). Compliance has strong, if poorly understood, health benefits as shown in randomized double-blind clinical trials (55). For example, in the Beta-Blocker Heart Attack (BHAT) trial, women who were compliant with placebo had a 60% decreased risk of mortality compared with women who were not compliant with placebo (56). Adjustment for multiple known predictors of coronary disease did not explain the decreased risk for coronary disease associated with good adherence to medication. The CHD risk reduction observed in compliant women adherent to placebo in BHAT is similar to the 50% reduction attributed to estrogen in an early meta-analysis of observational studies (57). This compliance bias is the main reason why intention-
to-treat analyses are essential components of clinical trials; otherwise as-treated results may simply reflect compliance bias within a trial.

Dose response

Stepwise graded associations showing decreasing risk with increasing dose or duration of therapy strengthen the likelihood that a beneficial association is causal. Among the common disease outcomes reported to date from WHI, the strongest evidence for a duration effect is for breast cancer. In contrast, the detailed report available on the website for the U.S. Preventive Services Task Force failed to find a consistent association of CHD with dose or duration of HT (58).

Temporality

Causality requires that the exposure precede the disease. This time sequence can be difficult to ascertain in studies of CHD because coronary artery atherosclerosis begins years before clinical manifestations. If the exposure causes death, case-control studies will miss it. Equally important, in many prospective studies information about medication is obtained only at baseline, and discontinuation or initiation during follow-up is unknown or incompletely assessed. If a new medication is started after baseline assessment and has an untoward serious effect during the interim, the fact of the event will be ascertained more often than the fact that medication was started. This design could underestimate early harm by misclassification of interim exposure and could have masked an early CHD harm in prospective studies of HT. This may explain why Framingham (59), the only study that defines exposures by most recent status, evaluating medication every 2 yr, is also the only observational study in Fig. 3 that shows a significant excess risk of CHD in HT users.

If the overall null CHD effect is explained by bias, the early harm requires another epidemiological explanation. The lack of interim medication data for HT in epidemiological studies may explain why early harm was not suspected until shown in clinical trials.

Specificity

Specificity of effect is one of the least useful criteria for causality, because many endogenous and exogenous exposures have diverse effects on different tissues. Specificity would not be expected to be a characteristic of studies of estrogen, given that estrogen receptors are present throughout the body.

Those who disbelieve the null trial results for CHD should recall that epidemiological studies correctly predicted the effect of HT on all of the other main WHI outcomes, including the increased risk of breast cancer, stroke, pulmonary embolus, reduced risk of colon cancer and fracture, and absent effect of the continuous combined regimen on endometrial cancer, as reviewed elsewhere (60). It is ironic that only the CHD prevention results from WHI differ from those predicted by observational studies, in that the prevention of CHD was the main factor motivating more universal use of HT for prevention. I believe that all of the apparent benefit of HT for CHD seen in observational studies and not confirmed by clinical trials could be explained by prevention and compliance biases.

Prevention and reversibility

Neither has been demonstrated in observational studies or in clinical trials with CHD outcomes.

Explanations and recommendations

The failure to find cardioprotective effects in any of the several clinical trials with CHD outcomes offers little hope that postmenopausal HT will prevent heart disease. The results, unexpected and unwelcome, are nonetheless definitive and are likely to extend beyond the studied treatment regimens. The most likely reason for lack of effect in trials compared with observational studies is consistent bias in the observational studies (61). The early harm (for both CHD and stroke) may be a thrombotic effect reflecting a synergistic effect between one of the many relatively common hypercoagulable states and HT (62). Against this explanation is the observation that raloxifene, a selective estrogen receptor modulator, causes a similar risk of venous thromboembolic disease as HT, but does not appear to increase the risk of stroke or CHD (63). Another candidate for early harm is inflammation, with rupture of the vulnerable plaque. In this regard, it is intriguing that estrogen raises C-reactive protein dramatically, whereas raloxifene has no effect on this in-

**Relative risk or odds ratio**

![Relative risk or odds ratio](image_url)

*Fig. 6. Relative risk or odds ratio for coronary artery disease incidence. [Adapted with permission from L. L. Humphrey et al.: Ann Intern Med 137:273–284, 2002 (53).]*
flammary maker (64). Other possibilities abound. We have spent the last 10 yr looking at mechanisms for estrogen's expected favorable effects. We now need to give equal time to explaining its harmful effects.

Whatever the mechanisms, a paradigm shift in medical practice should follow the mounting clinical trial evidence that postmenopausal hormone regimens are not indicated to prevent heart disease. Today's heart protection plans for women should emphasize behavior changes and, where appropriate, the lipid and blood pressure medicines that have been proven in clinical trials to reduce CHD risk in postmenopausal women.

Acknowledgments

Received May 20, 2003. Accepted June 2, 2003.

Address all correspondence and requests for reprints to: Elizabeth Barrett-Connor, M.D., Department of Family and Preventive Medicine, School of Medicine, University of California, San Diego, 9500 Gilman Drive, Mail Code 0607, La Jolla, California 92036-0607. E-mail: ebarrettconnor@ucsd.edu.

References


J Clin Endocrinol Metab. September 2003, 88(9):4031–4042

401

Downloaded from jcem.endojournals.org at Univ of Cal San Diego Serials/Biomedical Lib 0699 on May 3, 2007

Barrett-Connor • Clinical Review
51. 284
52. 1991 Cardioprotective effects of estrogen replacement therapy. JAMA 267:1833–1839
55. 545
59. 1997 The effects of hormone replacement therapy and raloxifene on C-reactive protein and homocysteine in healthy postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA 287:847–857
60. 2000 Raloxifene and cardiovascular events in osteoporotic postmenopausal women: results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA 287:847–857
62. 1999 Coronary heart disease and estrogen replacement therapy: Can compliance bias explain the results of observational studies? Ann Epidemiol 11:15–118
64. 2000 The effects of hormone replacement therapy and raloxifene on C-reactive protein and homocysteine in healthy postmenopausal women: a randomized, controlled trial. J Clin Endocrinol Metab 85:214–218
65. 2002 WHI Steering Committee and Writing Group Response, Letter to the Editor. JAMA 288:2823–2824
68. 1997 Use of postmenopausal hormone therapy and the primary prevention of cardiovascular disease. JAMA 288:2823–2824
69. 1997 Use of estrogen replacement therapy, are users healthier than nonusers? Am J Epidemiol 140:771–779
70. 1997 Coronary heart disease and estrogen replacement therapy. Can J Cardiol 13:153
76. 1997 The effects of hormone replacement therapy and raloxifene on C-reactive protein and homocysteine in healthy postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA 287:847–857
77. 2000 Raloxifene and cardiovascular events in osteoporotic postmenopausal women: results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA 287:847–857
78. 2000 The effects of hormone replacement therapy and raloxifene on C-reactive protein and homocysteine in healthy postmenopausal women: a randomized, controlled trial. J Clin Endocrinol Metab 85:214–218
79. 2002 WHI Steering Committee and Writing Group Response, Letter to the Editor. JAMA 288:2823–2824