Hormone Replacement Therapy: Real Concerns and False Alarms

Avrum Z. Bluming, MD, and Carol Tavris, PhD

Abstract: From 2002 to 2008, reports from the Women’s Health Initiative (WHI) claimed that hormone replacement therapy (HRT) significantly increased the risks of breast cancer development, cardiac events, Alzheimer disease, and stroke. These claims alarmed the public and health professionals alike, causing an almost immediate and sharp decline in the numbers of women receiving HRT. However, the actual data in the published WHI articles reveal that the findings reported in press releases and interviews of the principal investigators were often distorted, oversimplified, or wrong. This review highlights the history of research on HRT, including a timeline of studies that have or have not found a link between HRT and breast cancer; discusses how to distinguish important, robust findings from those that are trivial; closely examines the WHI findings on HRT and breast cancer, most of which are weak or statistically insignificant; reviews the current thinking about possible links of HRT with cardiovascular disease and cognitive functioning; and reports research on the benefits of HRT, notably relief of menopausal symptoms, that affect a woman’s quality of life. On these complicated matters, physicians and the public must be cautious about accepting “findings by press release” in determining whether to prescribe or take HRT.

Key Words: hormone replacement therapy (HRT), estrogen, breast cancer, women’s health initiative (WHI), risk assessment

(Cancer J 2009;15: 93–104)

Hormone Replacement Therapy (HRT) is the term used for the administration of estrogen, or estrogen plus progesterin, to women who have reached menopause. Estrogen is most commonly given with progesterin to women who still have a uterus, because as early as 1975 investigators had found that estrogen, taken alone, increases the incidence of uterine cancer. This increased risk is eliminated when progesterin is added. Estrogen replacement therapy (ERT) alone is thus generally given only to women who have had hysterectomies.

The majority of American women do not take any form of HRT during or after menopause; of those who do, most take it for fewer than 5 years. A minority of women—percentages vary across community studies—take HRT for the rest of their lives. HRT is highly effective in alleviating the most common menopausal symptoms, including hot flashes, night sweats, emotional lability, palpitations, insomnia, uncomfortable and frequent urination, and painful sexual intercourse. Some women who are at high risk of osteoporosis are also prescribed HRT, because estrogen decreases the incidence of osteoporotic hip fracture by 25% to 50%. Given that alternative medications—bisphosphonates such as Fosamax, Actonel, and Boniva—offer a similar protective benefit, most physicians no longer recommend using hormones simply to prevent hip fracture. These alternatives, however, can have unpleasant side effects, including esophageal and stomach irritation and, in rare cases, jaw-bone damage (osteonecrosis).

For decades, researchers, physicians, and women’s health advocates have debated the risks and benefits of estrogen, with or without progesterin. In the 1950s and 1960s, when Ayerst Laboratories aggressively began marketing their estrogen preparation, Premarin, supplementary hormones for postmenopausal women were heralded as a panacea that would, in the seductive words of New York gynecologist Robert Wilson, keep women “feminine forever.” Yet, by the 1980s, various critics began arguing that supplementary hormones were a serious and unnecessary risk to women’s health. The critics’ greatest concern was breast cancer, a disease that many women understandably fear (although heart disease causes far more deaths than breast cancer does).

In July 2002, with the first publication of findings from the Women’s Health Initiative (WHI), headlines across the country began trumpeting the dangers of HRT—not only breast cancer but also heart disease and stroke. The news was particularly alarming because the WHI is the largest prospective study in which women were randomized to take hormones or a placebo and then followed over time. An earlier prospective, randomized, double-blind study had found no increased risk of breast cancer in women on HRT, even after 22 years, but this small study never made headlines. The WHI research has cost nearly a billion dollars; the investigators consist of eminent physicians, statisticians, and epidemiologists across the country, and the findings have been published in medicine’s most prestigious journals. Accordingly, the WHI’s findings received, and continue to receive, worldwide attention. It is no wonder that its claims of the dangers of HRT caused the prescription rate for HRT to fall by some 50%.

DATA DREDGING, RISK REPORTING, AND OTHER PROBLEMS IN RESEARCH

Should women who have menopausal symptoms deny themselves its benefits, whether in the short term or over many years, because they fear breast cancer, heart disease, or stroke? Are their concerns warranted by the data? When we took a close look at the findings in the published WHI articles, placing them in the context...
of research on HRT over the past decades, we were surprised by the enormous discrepancy we found between the belief that hormones are dangerous and the lack of supporting data.

Science is a process; it is rare that a single study gives us a definitive answer.\(^{26}\) Yet the news-hungry media crave “breakthroughs” and thrive on scare stories. Thus, it is essential to look behind the headlines to the actual data, to try to get a sense of the larger picture that emerges over time and across studies. Sometimes that larger picture yields a clear image; sometimes, as with HRT, it becomes foggier than ever. Two statistical errors common to research on HRT have contributed to that fog: one has to do with how risks are reported; the other has to do with the often inappropriate “mining” of data, when researchers retrospectively hunt around in their findings for something, anything, that might seem to be a significant risk factor.

Consider, first, the difference between absolute risk and relative risk. The media, following the example of many researchers themselves, tend to report relative risks, which are expressed in percentages that can seem more important than they are. For example, if we tell you that the relative risk of breast cancer is increased by 300% in women who eat a bagel every morning, that sounds serious, but it is not informative. You would need to know the baseline absolute number of new breast-cancer patients. If the number shifted from 1 in 10,000 women to 3 in 10,000 women, that is a 300% increase, but it is very likely a random artifact. If the risk had jumped from 100 to 300 in 10,000, also a 300% increase, we might reasonably be concerned. In large epidemiological studies that generally include tens of thousands of people, it is very easy to find a small relationship that may be considered “significant” by statistical convention but which, in practical terms, means little or nothing because of the low absolute numbers.\(^{29}\)

This is why scientists who are working to promote statistical literacy, especially in helping the public and physicians understand actual versus inflated risks of diseases and treatments, emphasize that knowing the baseline of absolute numbers when comparing two groups is essential.\(^{30}\) Two major consensus projects on the reporting of clinical trials concluded that stating relative risks alone is often misleading. A reliance on relative risks can also create misleading, faulty comparisons. For example, let us say that 3% of the people who eat chocolates develop cavities, and 2% of people who do not eat chocolates develop cavities. The absolute difference between these populations is only 1%. That means that for every 100 people who eat chocolates, 1 extra person will develop cavities (in addition to those who do not eat chocolates). The absolute difference between these groups is essential.\(^{30}\) Consider, for example, if we told you that the relative risk of breast cancer is increased by 300% in women who eat a bagel every morning, that sounds serious, but it is not informative. You would need to know the baseline absolute number of new breast-cancer patients. If the number shifted from 1 in 10,000 women to 3 in 10,000 women, that is a 300% increase, but it is very likely a random artifact. If the risk had jumped from 100 to 300 in 10,000, also a 300% increase, we might reasonably be concerned. In large epidemiological studies that generally include tens of thousands of people, it is very easy to find a small relationship that may be considered “significant” by statistical convention but which, in practical terms, means little or nothing because of the low absolute numbers.\(^{29}\)

A reliance on relative risks can also create misleading, faulty comparisons. For example, let us say that 3% of the people who eat chocolates develop cavities, and 2% of people who do not eat chocolates develop cavities. The absolute difference between these populations is only 1%. That means that for every 100 people who eat chocolates, 1 extra person will develop cavities (in addition to those who do not eat chocolates). This is not a particularly frightening risk if you enjoy chocolate. But suppose we report the identical conclusion as a relative risk: 1 additional case divided by the 2 baseline cases gives us an increased relative risk of 0.5 or 50%. A 50% increase in cavities if you occasionally eat chocolates! Stop at once!

Many of the studies of HRT and risk of disease, especially breast cancer, have produced statistically modest or borderline results that have been made to look more impressive than they actually are by reporting them as relative risks. Consider Table 1, which lists the reported increases in relative risks associated not only with HRT\(^{33,34}\) and ERT\(^{35}\) but also with birth weight,\(^{36}\) fish intake,\(^{37}\) eating 1 additional serving of French fries per week during preschool years,\(^{38}\) eating grapefruit,\(^{39}\) working on a night shift,\(^{40,41}\) working as an airline flight attendant in 2 different airlines,\(^{42–44}\) suffering from severe caloric restriction during the 1944–1945 Dutch famine,\(^{45}\) taking antibiotics,\(^{46}\) and the use of electric blankets by African-American women.\(^{47}\) You can see at a glance how weak these associations are; to put them in perspective, we included the

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Relative Risk</th>
<th>95% Confidence Interval (CI)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conjugated equine estrogen</td>
<td>0.77</td>
<td>0.59–1.01</td>
<td>34</td>
</tr>
<tr>
<td>Birth weight</td>
<td>1.09†</td>
<td>2.00–17.00</td>
<td>35</td>
</tr>
<tr>
<td>Fish intake</td>
<td>1.14</td>
<td>1.03–1.26</td>
<td>36</td>
</tr>
<tr>
<td>Premarin/Progestin</td>
<td>1.24</td>
<td>1.01–1.54</td>
<td>37</td>
</tr>
<tr>
<td>Premarin/Progestin</td>
<td>1.26</td>
<td>1.00–1.59</td>
<td>25</td>
</tr>
<tr>
<td>French fries</td>
<td>1.27</td>
<td>1.12–1.44</td>
<td>38</td>
</tr>
<tr>
<td>(1 additional serving per week)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grapefruit</td>
<td>1.3</td>
<td>1.06–1.58</td>
<td>39</td>
</tr>
<tr>
<td>Night shift work</td>
<td>1.51</td>
<td>1.36–1.68</td>
<td>40, 41</td>
</tr>
<tr>
<td>Flight attendant (Finnish)</td>
<td>1.87</td>
<td>1.15–2.23</td>
<td>42, 43</td>
</tr>
<tr>
<td>Dutch famine†</td>
<td>2.01</td>
<td>0.92–4.41</td>
<td>44</td>
</tr>
<tr>
<td>Antibiotic use§</td>
<td>2.07</td>
<td>1.48–2.89</td>
<td>45</td>
</tr>
<tr>
<td>Flight attendant (Icelandic)</td>
<td>4.1</td>
<td>1.70–8.50</td>
<td>46</td>
</tr>
<tr>
<td>Electric blanket use§</td>
<td>4.9</td>
<td>1.50–15.6</td>
<td>47</td>
</tr>
<tr>
<td>Tobacco smoking and lung cancer</td>
<td>26.07</td>
<td>6.58–103.3</td>
<td>48</td>
</tr>
</tbody>
</table>

*A CI provides a range (an interval) with a specified probability that a given result will, with continued replications, fall within it 95% of the time (90% or 99% are also used). In the case of large-scale epidemiological studies, if the spread of the confidence interval includes the number 1.0, the result is not considered statistically significant even in the presence of a low P value. Generally speaking, the lower limit of the CI should be at least 3.0 before the finding is considered a strong, reliable one.\(^{29}\)

†The odd finding that the relative risk is lower than the lower limit of the confidence interval is not a typo. The RR of 1.09 is a calculated increase per 1000 g (2.2 pounds) of weight at birth.

‡RR was maximum for those exposed to severe famine only between ages 2 and 9.

§1–50 d of antibiotic use = RR 1.45; more than 1001 d of antibiotic use = RR 2.07. But apparently this result depends on why a woman was taking the antibiotic: There was no increase in RR for women using tetracycline or macrolide for acne or rosacea.

*Increased RR most pronounced for more than 10 yr of use, especially when women who used the device for more than 6 mo per yr were excluded.

Results for a real finding—smoking and lung cancer—at the bottom of the table.\(^{48}\) The relative risks in almost all cases are very low, and the use of HRT is virtually the lowest, being less risky than eating fish or grapefruit, using antibiotics, or being a flight attendant.

Another way of misrepresenting findings comes from the practice, severely frowned upon in research, of retrospective stratification, commonly known as “data mining” or “data dredging.”\(^{49–56}\) Data mining occurs when researchers, having failed to find the statistically significant associations that they had originally hypothesized would exist between a possible risk factor and a disease, go back into their data and rummage around, looking for other factors that might show a statistical link to the dependent variable in question. This effort might yield interesting questions or hypotheses for future research, but the problem is that in a data set of many thousands of people, some relationship that is unearthed retrospectively will turn out to be statistically significant (ie, P < 0.05) just by chance.\(^{57}\) In Against the Gods: The Remarkable Story of

\(^{A}\) A consensus article on how best to report findings from randomized trials cautioned authors to “especially resist the temptation to perform many subgroup analyses. Analyses that were prespecified in the trial protocol are much more reliable than those suggested by the data.” The authors did not mince words: “The strategy for reporting study results should be specified before the results are known, and selective reporting or emphasis of statistically significant results based on ex post factosubgroup analyses should be discouraged.”\(^{57}\)
Risk, the economist Peter Bernstein put it this way: “If you torture the data long enough, the numbers will prove anything you want.”56

A now-famous example of the spurious results that can emerge from data mining can be found in an article that was submitted to the Lancer in 1988, reporting that men hospitalized for acute heart attacks who had been taking an aspirin daily had a better survival rate than similarly hospitalized men who had not been on aspirin. This was clearly an important finding, and the editors agreed to accept this article with 1 condition: The authors would have to retrospectively stratify the 17,187 men in their study according to a variety of factors, including the men’s age, weight, and race. Now, it would certainly be good to know whether the benefit of taking aspirin (or any other drug) is affected by how old you are or whether you are overweight or Asian, or other possible demographic factors. But the authors refused to do this reanalysis, explaining that the benefit or risk for these subcategories would best be assessed by a new prospective study. The editors insisted: no stratification, no publication. And so the authors eventually turned in a revised article with the additional findings, including a slightly adverse effect of aspirin on mortality in patients born under the astrological signs of Gemini or Libra, in contrast to a strikingly beneficial effect of aspirin for patients born under all other astrological signs. The editors agreed to publish the article if the astrological results were omitted. “You wanted retrospective stratification, we gave you retrospective stratification,” the authors said (in effect), and demanded that the Lancer stick to the deal. And so this landmark article was published, with a new title that began: “Aspirin’s effect on myocardial infarct mortality and astrology.”77

Richard Feynman,90 a Nobel Laureate in physics, had a good test for truth in science. He said: “If something is true, really so, if you continue observations and improve the effectiveness of the observations, the effects stand out more obviously. Not less obviously.” The relationship between cigarette smoking and lung cancer is an example of the truth becoming clearer with repeated observations: an association between them is noted, then confirmed with replications, and further understood when the biologic mechanism accounting for the association is identified. In the case of lung cancer, the epidemiologic data have been consistent across many replications, and further understood when the biologic mechanism underlying the association is identified. In the case of lung cancer, the epidemiologic data have been consistent across many replications, and further understood when the biologic mechanism underlying the association is identified.

The strength and consistency of these data are sufficient to draw a conclusion about a causal relationship: cigarette smoking causes a significant increase in the risk of lung cancer.

In contrast, the relationship between HRT and breast cancer is still not clear despite a vast amount of research, study, and reporting over many decades.61–81 Table 2 reviews the highlights of research from the first manufacture of estrogen tablets (Premarin) in 1942 to the most recent studies in 2008; as you can see, the list is a jumble of positive findings, negative findings, and meaningless findings. Let us see why.

HRT AND BREAST CANCER: IS THERE A LINK?

On July 9, 2002, the National Institutes of Health issued a press release: “The National Heart, Lung, and Blood Institute of the NIH has stopped early a major clinical trial [the Women’s Health Initiative] of the risk and benefits of combined estrogen and progestin in healthy menopausal women due to an increased risk of invasive breast cancer.” The WHI investigators reported that women who had been randomly assigned to take a combination of estrogen and progestin had a small increased relative risk of breast cancer (relative risk = 1.26) when compared with women who were randomly assigned to a placebo.25 (1.26 means a 26% increase in risk.)

What few noticed in the published article was this little sentence: “The 26% increase in breast cancer incidence among the HRT group compared with the placebo group almost reached nominal statistical significance.” Almost means it did not reach statistical significance. Of course, any increase might be of legitimate concern, warranting further investigation. But were this finding valid, one would have expected to see an even greater increased incidence of early, noninvasive breast cancer, the kind that precedes invasive breast cancer. There was, however, no difference between the 2 groups in the incidence of this early breast cancer, nor in deaths from breast cancer.

Yet many reporters and physicians treated that 26% increase in relative risk as being not only statistically significant but also medically significant. In an editorial published in the June 25, 2003, issue of JAMA, Peter Gann and Monica Morrow96 wrote: “A statistically significant 26% increase in breast cancer incidence contributed to the overall negative effect of estrogen plus progestin.”

In 2006, in another update of this same cohort of patients, the WHI reported no increased risk of breast cancer among women randomized to combined estrogen-progestin treatment. The “significant” relative risk had completely vanished.93 This news did not make headlines.

One of the studies that is still frequently cited by those striving to find an association between HRT and breast cancer is a 1989 Swedish study (Table 2), which reported a 440% increased risk of breast cancer among women taking combined estrogen and progestin for more than 6 years.84 This sounds impressive, until we learn that it was not statistically significant (the confidence interval was 0.9–22.4) and that it was based upon only 10 patients in the study who developed breast cancer while taking HRT. The baseline study population consisted of the 23,244 women in Uppsala, Swe-

---

Remarkably, another article in that same issue of JAMA cited the same 2002 article correctly: “After 5.2 years of follow-up the WHI reported that combined HRT was associated with a statistically nonsignificant 26% increase in breast cancer risk.”77

We calculate this increase based on data from a 2007 publication from the National Center for Health Statistics, showing that the average risk of breast cancer across 3 age groups (40–59, 60–69, 70, and older) is 4.82%, about 5 in 100. Even if the WHI finding of a 1.26 increased relative risk because of hormones had been statistically significant, then the risk will increase from 4.82% to a total additional 0.26% of the 5%. This increase in incidence is obtained by multiplying the baseline, 4.82% average risk, by 1.26, which yields a revised risk of just under 6%, or 6 in 100. (In absolute numbers, from 182,500 cases to 184,325 cases.)
TABLE 2. HRT and Breast Cancer: A Timeline of Relevant Events and Studies

1942: Researchers develop methods to extract large quantities of estrogen from the urine of pregnant mares, and Ayerst Laboratories launches Premarin (from PREGnant MARes’ urINe), the first estrogen tablets.

1950s: Ayerst Laboratories begins a marketing campaign promoting the use of Premarin to lessen menopausal symptoms.

1966: New York gynecologist Robert Wilson publishes Feminine Forever, a best seller that promises youth, beauty, and a “full sex life” for menopausal women through the use of hormone therapy.

1975: Postmenopausal women on estrogen are found to have a 4- to 8-fold increase in the risk of uterine cancer.¹

Mid 1980s: The addition of progestin to estrogen negates the increased risk of uterine cancer associated with estrogen alone.² By 1986, over 20 million prescriptions for non-contraceptive hormones are dispensed.⁸²

1982: A study finds that estrogen is not associated with an increased risk of breast cancer.⁸³

1974–1989: Of the 26 most cited reports during this period that have investigated the association between HRT or ERT and breast cancer
5 report an increased risk
7 report a decreased risk
14 report no significant association.⁸⁴

1984: A study reports that estrogen does not increase the risk of breast cancer, even when taken for many years.⁸⁵

1989: A Swedish study reports a 440% increased risk of breast cancer associated with the administration of combined estrogen and progestin for more than 6 yr. However, this risk is based upon only 10 patients in the study who developed breast cancer while taking HRT.⁶⁴

1989: A 17-yr follow-up study of more than 3000 women who had had benign breast lesions and were taking estrogen finds the women had no increased risk of developing breast cancer. In fact, estrogen slightly lowered the breast cancer risk in women with atypical hyperplasia and several other conditions.⁸⁰

1992: The Nurses’ Health Study, an observational study that followed 121,700 female registered nurses from 1976 through 1992, finds no increased risk of breast cancer when women who had ever used HRT are compared to women who never took HRT, and no increased risk of breast cancer even when HRT users for over 10 yr are compared to never users.⁸⁷

1996: A study reports that ever-use of estrogen replacement therapy is associated with a slightly decreased risk of fatal breast cancer.⁶⁹

1997: A prospective cohort study of nearly 42,000 Iowan women with a family history of breast cancer reports that HRT use is not associated with a significantly increased risk of breast cancer, but is associated with a significantly reduced mortality rate from all causes.⁷⁰

2000: A study reports a 40% increased risk of breast cancer associated with HRT.⁷⁶ However, this risk is limited to women weighing no more than 90 pounds.

2000: A retrospective study finds an increased risk of breast cancer among estrogen-only users, but only after 15 yr of use. A barely significant increased risk of breast cancer among estrogen-progestin users is found after 5 yr.⁷⁸

2002: The Women’s Health Initiative (WHI) terminates the estrogen-progestin arm of their study prematurely because they claim to have found an increased risk of breast cancer. This increased risk is, however, not statistically significant.²⁵

2003: The WHI “confirms” its 2002 finding, reporting a small increased risk of breast cancer among women on HRT. Analysis of the data according to kind of cancer finds either no statistical significance or barely a 1% increased risk.⁷⁷

2003: A British study, published in The Lancet, entitled “The Million Women Study,” reports an increased risk of breast cancer in women taking ERT or HRT. However, they also find:
● No increase in risk of breast cancer in past users of either estrogen or estrogen-progestin, regardless of duration of use.
● The increased risk of breast cancer is found only in current users.
● The average period of follow-up is only 2.6 yr.⁸⁸

2004: The WHI reports no increased risk of breast cancer associated with the use of estrogen alone.⁸⁹

2004: The WHI reaffirms its 2002 finding of no increased risk of breast cancer among women taking estrogen, even after 7 yr. This time, they report that women who had used HRT in the past had a lower rate of breast cancer than women who had never taken hormones.⁹⁰

2006: A study of 9000 Japanese women finds that HRT users are less likely to develop breast cancer than never-users.⁹¹

2006: A study reports no increase in breast cancer incidence among women who have been taking estrogen, even after 8 or more years.⁹²

2006: The WHI now reports no increased risk of breast cancer even among women randomized to take combined estrogen and progestin.⁹³

2008: The WHI reports that the risk of cardiovascular events, malignancies, breast cancers, and deaths from all causes was higher in the HRT group than in the placebo group even 3 yr after the women stopped taking HRT. However, none of the associations between HRT and breast cancer or mortality rates is statistically significant.⁹⁴

2008: An observational study of 472 postmenopausal women who have a genetic predisposition to breast cancer, the BRCA1 mutation, finds that hormone use, either as HRT or ERT, is not associated with increased risk of breast cancer. On the contrary, it is associated with a decreased risk.⁹⁵

---

den, who received prescriptions for HRT in a 3-year period. The researchers took a much smaller subset of that population to analyze, calculated that 2.2 women would be expected to get breast cancer, and found that 10 actually did—hence the “440% increased risk.”

In addition, the Swedish study found no statistically significant increased risk of breast cancer among women who used estrogen alone, which might make us wonder why estrogen was seen as the villain. Elizabeth L. Barrett-Connon,¹¹ in an editorial accompanying the report, concluded: “For the average North American woman, who will be postmenopausal for one third of her life, the benefits of estrogen seem strongly established. In my opinion, the data are not conclusive enough to warrant any immediate change in

© 2009 Lippincott Williams & Wilkins
the way we approach hormone replacement.” The Harvard Medical School Health Letter reviewed the Swedish study and concluded: “The most striking ‘result’ was in women who took estrogen combined with progestin for more than 6 years, and this was what made headlines. These women seemed to have 4.4 times the average risk of developing breast cancer. But there is a very important reason not to take this figure literally. There were only 10 women in this group, too few to provide a statistically stable result, the true value had a 95% chance of being 10% below the average or as high as 22.4 times the average (an incredible figure), or somewhere in between. Earlier research has given us no reason to expect a strong association between estrogen replacement and breast cancer.” Yet in today’s climate, that very same study is used to support the argument that HRT causes breast cancer.

To further their case, some investigators have turned to retrospective analysis. Several of the “significant” associations in Table 1 were a result of data mining: the use of antibiotics increases relative risk, but not among women using tetracycline or macrolide for acne or rosacea (apparently breast cancer needs to know why a woman is taking an antibiotic); the increased risk of surviving the Dutch famine occurs only among women who were between 2 and 9 at the time; and, in the most unintentionally funny result, the breast cancer risk associated with using electric blankets increased among African American women who used the blankets for more than 10 years—but only if those who used them for more than 6 months per year were excluded from analysis. Table 2 includes some good examples of data mining too, such as a 2000 study that found a 40% increased risk of breast cancer associated with HRT. It took some determination to get that result, because the increased risk applied only to women weighing not more than 90 pounds.

The WHI and the observational (nonrandomized) Nurses’ Health Study, which followed 121,700 female registered nurses from 1976 through 1992, are both guilty of data mining. When the association between HRT and breast cancer repeatedly failed to reach statistical significance, the investigators did not say, “Good news! Looks like HRT is probably safe, at least on the breast cancer question.” Rather, they jumped back into the data pool, trying to find something that was significant—maybe that some form of HRT is harmful for some women, or is related to some kinds of breast cancer, or is hazardous after some length of time. These are all serious possibilities, of course, and might warrant a new prospective study. But we repeat: when you get results from retrospective analysis, rather than as a premeditated focus of investigation under controlled conditions, the findings are likely to be confusing, unreplicable, and difficult to interpret. The WHI was explicit in 2002: “The WHI and the observational (nonrandomized) Nurses’ Health Study have failed to confirm the results of the WHI randomized controlled trials.”

In contrast, here is what you learn if you initially set out to study a question involving a subset of women at risk for breast cancer. If HRT were a significant risk factor, then it should surely pose particular risks for women who have a BRCA1 or BRCA2 genetic mutation, which predisposes them to develop the disease. When these women have their ovaries removed, their risk of developing breast cancer falls by half. But the surgery induces menopause, and some patients subsequently take HRT to alleviate menopausal symptoms. Are they in special danger? Does taking estrogen negate the benefit of the surgery?

To find out, researchers compared the use of HRT and ERT in 236 breast cancer patients and 236 matched controls, all of whom carried a mutation in BRCA1. (There were not enough women with BRCA2 to be included.) The results, reported in the October 1, 2008 issue of the Journal of the National Cancer Institute, found no increased risk among women taking hormones, whether they had undergone natural, age-related menopause, or surgically induced menopause. This finding is so important, and so counter-intuitive, that we want to underscore its message. If lowered estrogen levels following removal of the ovaries were the reason for the drop in breast cancer risk in women with the BRCA1 mutation, then it is irrational to give them supplemental estrogen to alleviate symptoms, right? Yet, according to another large-scale study and its own follow-up, administering estrogen to women with BRCA1 mutations, following removal of the women’s ovaries, did not nullify the benefits of the surgery. Their risk of breast cancer remained just as low.

Moreover, the majority of observational studies have found no increased risk of breast cancer associated with HRT, including studies in which HRT was given to women with a family history of breast cancer. Many researchers today are inclined to dismiss observational study conclusions, even though the Nurses’ Health Study is among them. What is wanted, they say, is a prospective, randomized study, the gold standard for determining the validity of findings in clinical trials. Yet a review of the medical literature, comparing results from observational studies and from randomized controlled trials, found that both methods usually produce similar outcomes. Another review found that the persistent validity of conclusions 20 years after initial publication was 87% among nonrandomized (observational) trials and 85% among randomized clinical trials. One reason is that randomized controlled trials are like the 10 Commandments—a fine ideal, but very difficult to execute in practice. For starters, most participants know if they are taking an active medication or an inert placebo, which affects their subsequent behavior; true randomization and double-blinding are often difficult, if not impossible, to achieve. We are not saying that observational studies are “as good” as the gold standard, but rather that both methods have strengths and weaknesses, and that, again, it is important to consider the overall pattern of evidence rather than any single study.

Some investigators who believe that the relative risks of HRT are serious enough to warrant concern acknowledge that the absolute risks from this treatment are small. In one worst-case analysis, researchers calculated that a 50-year-old woman taking estrogen and progesterin for 10 years has only a 4% risk of breast cancer. Without HRT, her risk would be 2%. (An alarmist headline writer might report this finding by stating that a woman’s risk is “doubled” if she takes HRT for 10 years, whereas a reassuring statistician would say that she has a 96% chance of remaining free of breast cancer versus 98% if she does not.) Moreover, even if HRT increases the risk of breast cancer by this modest increment, research suggests that women on HRT live longer than those not taking HRT, and that HRT-treated women have a lower death rate from breast cancer. How can the very hormones that allegedly increase the

---

6 When one of us (AB) directly asked the investigators in print why they had retrospectively subdivided their sample into current and past users, they did not answer then or since.
risk of breast cancer also be responsible for a better survival rate from that cancer?

**BUT DOES ESTROGEN CAUSE CANCER?**

The hypothesis that hormones are linked to breast cancer was originally derived from 2 well-documented facts: the incidence of breast cancer is 100 times greater in women than in men, and the earlier a woman’s menarche and the later her menopause, the greater her risk of breast cancer.123 These observations suggested, reasonably, that perhaps having more years of circulating estrogen was the culprit. As we noted in the example of cigarette smoking and lung cancer, the first step in the scientific process is to document a reliable association, and the second step is to demonstrate the biologic mechanism that might account for it. In the case of the hypothesis that estrogen causes breast cancer, not only has the association turned out to be weak or nonexistent, but also the second step has been contradicted by various lines of evidence:

- Birth control pills, which used to contain far more estrogen than HRT does, should therefore increase the risk of breast cancer. Although controversy continues on this question,124 most published studies find that oral contraceptives do not increase the risk.125–141
- Women taking estrogen alone (ERT) should have a higher risk of breast cancer. They do not. The WHI itself found that they have no increased risk, even after an average follow-up of 6.8 years; if anything, these women have a slight decrease in breast-cancer risk.90
- The incidence of breast cancer increases as women grow older. If taking estrogen is part of the reason, the breast cancer rate among postmenopausal women who do not take HRT should decrease with age, along with their naturally declining estrogen levels. It does not.142–144
- According to the National Cancer Institute’s Division of Cancer Etiology, estrogens are not direct carcinogens for mammary cells.145 Estrogen can, however, induce cell proliferation. So the WHI modified their original hypothesis into this version: mutation-inducing agents are all around us, and the higher the rate of cell proliferation, the more possible it is that a proliferating cell will be exposed to a mutagen and become malignant. The problem with this argument is that the endometrium, the lining of the uterus, is more sensitive to the proliferative effect of estrogen than is the breast. As we mentioned, women who take estrogen alone have a 5– to 6-fold increased risk of uterine cancer, but no increased risk of breast cancer. If prolonged stimulation of estrogen solely from an early menarche and a late menopause predisposed women to cancer, we would see its effects on rates of endometrial cancer. We do not.146

Mammary gland cells are divided into those that have an estrogen receptor (ER) molecule on their surface, ER positive, and those that do not have this estrogen receptor, ER negative. Some researchers hypothesize that HRT might cause an increase in breast cancer by stimulating the estrogen-receptor-positive cells. The problem with this seemingly logical notion is that ER positive cells are, most often, not the ones proliferating in breast cancer. The really dangerous cells are the 5% that constitute the cancer stem cell, and they are not ER positive. ER expressing cells of the mammary epithelium are distinct from the stem cell population, and any effect of estrogen on the stem cells is mediated indirectly.147,148

But what about drugs like tamoxifen, an “estrogen antagonist,” which are given to breast-cancer patients to reduce the chance of recurrence? One of the arguments that estrogen causes or promotes breast cancer is that tamoxifen helps to reduce or retard the growth of ER positive breast cancer by competitively blocking the binding of estrogen to the estrogen receptor on breast cancer cells.149 However, several lines of research dispute this belief. For one thing, when tamoxifen is given to premenopausal women, their natural estrogen levels increase up to 5-fold.150 This rise in estrogen should block any competitive binding of tamoxifen, yet tamoxifen’s effect against breast cancer works as well in these premenopausal as in postmenopausal women.151–153 Second, approximately 40% of ER positive patients fail to respond to tamoxifen.154 Third, laboratory studies have shown that tamoxifen inhibits the stimulatory effects of growth factors involved in breast cancer155–158 even in the absence of estrogen.159 In addition, after treatment with tamoxifen, some breast cancer cells actually acquire the ability to proliferate160—and low doses of estrogen have been shown capable of killing them.161–164 Finally, tamoxifen has also been shown to have a therapeutic effect on ER negative breast cancer cells, both in laboratory studies and in human patients.165

In other words, tamoxifen works in a variety of ways that are exclusive of its action on estrogen receptors. Because the precise mechanisms responsible for its therapeutic effect remain unknown,166,167 it seems inadequate and simplistic to claim that the success of tamoxifen supports the view that estrogen causes breast cancer or stimulates cellular proliferation in breast cancer. The overall picture to date, therefore, persuades us that HRT is not a major risk factor for breast cancer.

**ATHEROSCLEROTIC CARDIOVASCULAR DISEASE AND STROKE**

Heart-disease deaths exceed breast cancer deaths in every decade of a woman’s life including her 30s, and as women age, their risk of death from heart disease is more than 5 times as great as that from breast cancer.144 Understanding the role of HRT in the possible development or progression of heart disease, or protection against it, is therefore crucial.

Throughout the 1980s, many studies found a cardiovascular benefit of taking HRT. A 1989 review of 19 published studies of ERT’s effect on heart disease reported that ERT was associated with at least a 30% reduction in clinical coronary artery disease. This conclusion was consistent among 90% of the cohort studies, 63% of the case control studies, and the only double-blind randomized trial that had been done to date.168 By 1991, a New England Journal of Medicine editorial reported that a consensus of epidemiological studies had shown that women who are given postmenopausal estrogen have a 40% to 50% reduction in the risk of coronary artery disease in comparison with women who do not receive such therapy.169 In 2000, the Nurses’ Health Study reported that HRT reduced the development of primary cardiovascular disease by nearly 40%,170 a condition responsible for more than 300,000 deaths among US women per year.144

Subsequently, a large randomized study, the Heart and Estrogen/Progestin Replacement Study, found a statistically significant increase in heart events in women with known coronary artery disease receiving HRT—but only during the first year of use.171 In 2002, the WHI reported that women on HRT, but not women on estrogen only, had a slightly increased relative risk of “heart events”—coronary heart disease (including acute myocardial infarction requiring hospitalization or silent myocardial infarction), death because of heart disease, angina, or indications for revascularization surgery.22 But, as in the Heart and Estrogen/Progestin Replacement Study, this increased risk occurred only among women in the first year of taking combined HRT.172 In 2007, the WHI investigators

Technically, they also found a significant increase in cardiovascular events among women in their fifth year of taking hormones, but that seems to be a result of a fluke—a surprisingly low incidence of coronary heart disease in the comparison placebo group.
revised their 2002 findings, now concluding that women who start HRT within the first 10 years following menopause actually reduce their risk of coronary artery disease, whereas those who start after that period slightly increase their risk.173 In another confluence of a randomized trial and an observational study, the Nurses’ Health Study reached the same conclusions.174

Why should HRT increase cardiovascular risk only in the first year, and only among older women? We know from primate data that continuous estrogen keeps blood vessels healthy; we also know that estrogen replacement after a hormone-free interval cannot reverse vascular damage.175 The Estrogen Prevention of Atherosclerosis Trial and Estrogen Replacement and Atherosclerosis studies are consistent with the animal data.176,177 One leading explanation is that among women who do not have heart disease, HRT reduces oxidation of low-density lipoproteins and causes blood vessels to dilate, thereby inhibiting the development of atherosclerosis. However, in women who do have underlying heart disease, HRT can be potentially harmful, because it can induce inflammation in existing arterial plaque, causing a stable plaque to rupture, and can also promote bleeding into the plaque, both of which can lead to blockage of a critical coronary artery.

This analysis would explain why studies that have enrolled women of a younger age, like the Nurses’ Study, found that HRT had a protective effect: these women were less likely to have arterial plaques. But in the WHI, only 10% of the women were between 50 and 54 years old, ages at which HRT might have played a beneficial role; 70% were 60 to 79 years old, in an age range where we would expect to find previously formed plaques.25 Although HRT was effective in reducing LDL, total cholesterol, and glucose, and in raising high-density lipoprotein levels, these benefits did not result in a reduced incidence of cardiovascular disease in the older women, consistent with preexisting atherosclerotic disease in this population.

Atherosclerosis was probably present in the WHI population because, in addition to the median age of 63, fully 70% of the women were overweight and half of them were obese. Nearly 50% were either current or past cigarette smokers and more than 35% had been treated for high blood pressure.178 Yet women with these well-established risk factors for cardiovascular disease were not excluded from the analysis of HRT and cardiovascular events. The WHI investigators have repeatedly stated that all of the women they recruited were healthy, a prerequisite for participation in this primary-prevention study; these assertions are difficult to reconcile with the actual medical histories of so many of the women.

Our conclusions are:

- HRT may have beneficial effects on the heart for women who start taking hormones early in menopause (around age 50) because estrogen promotes healthy blood vessels and may help delay the formation of plaque.
- HRT probably has no protective effect on women who begin the use of HRT later, in their mid-60s.
- HRT is potentially risky for women who begin taking it in their 60s, at least for the first year, especially if they have preexisting artery disease.

Overall, we share the conclusion of most cardiologists: there is no reason for women to take hormones primarily to help forestall or prevent cardiovascular disease, given that there are other effective ways of reducing heart-disease risk.

Before moving on, let us consider one other headline-grabbing, fear-inducing story from the WHI. In 2004, the WHI announced it was stopping the estrogen-only arm of the study because the use of estrogen increased the risk of nonfatal stroke by 12 per 10,000 women per year.179 However, the WHI investigators had an extremely broad definition of “stroke”—including transient, “subtle neurologic deficits” that resolved in a day or two. Some epidemiologists have argued that this small apparent increase was artificially introduced by a “detection bias”: the fact that women on HRT, having been made so sensitive to possible adverse effects of hormones, had become hyperalert to any symptoms. Indeed, when these critics reanalyzed the findings, controlling for detection bias, the increased risk of stroke vanished.179

**COGNITIVE FUNCTIONING AND ALZHEIMER’S DISEASE**

Laboratory studies with animals have suggested that estrogen can modify the structure of nerve cells in the brain and alter the way they communicate with each other, a process called neuroplasticity. The real-life applications of this research remain uncertain and controversial, but some evidence indicates that estrogen therapy administered after menopause may prevent, or at least delay, the onset of Alzheimer’s disease.180–196

The WHI researchers, however, remain unconvinced of this possibility. In their 2003 report, the WHI authors concluded that estrogen plus progestin increased the risk for dementia in women aged 65 and older, and did not prevent the development of mild cognitive impairment—further support, they said, for their conclusion that the risks of HRT outweighed any possible benefits.197 The increased incidence of dementia in the HRT group, compared with women on placebo, occurred as early as 12 months after the women started HRT. In contrast, there was no increased incidence of mild cognitive impairment between the 2 groups during the entire trial period. If HRT were really harmful to the brain, surely mild cognitive problems would emerge before full-blown dementia.

In 2004, a follow-up WHI article repeated the assertion that estrogen increased the risk for both dementia and mild cognitive impairment.198 However, when women who had mild cognitive impairment at the start of the study were excluded from analysis, the results were no longer statistically significant.199 Yes, we had to read that twice also. Estrogen is associated with cognitive impairments—but only among women who are already cognitively impaired.

**BENEFITS VERSUS RISKS OF HRT**

It is difficult to resist the conclusion that the WHI investigators have been doing everything they could to wring the bleakest possible interpretation from their recalcitrant data. They do not even acknowledge the single greatest benefit of HRT: its relief of menopausal symptoms. On the contrary, they have concluded that “in postmenopausal women, estrogen plus progestin did not have a clinically meaningful effect on health related quality of life,” even after taking HRT for 3 years.199 Because it takes less than a week for most symptomatic menopausal women to feel better after starting HRT, many readers of the WHI article may be forgiven for asking: what were these investigators thinking?

To be sure, the WHI was not interested in the effect of hormones on menopausal symptoms; they were investigating the big problems—breast cancer, heart disease, and cognitive impairment. That is a legitimate goal, of course, but then why publish an article on the menopausal symptoms they did not study? The article notes that women who reported moderate or severe menopausal symptoms “were discouraged from participating in the study” and, perhaps as a result, “Moderate or severe vasomotor symptoms at baseline were present in only 12.7% of study patients.” Not surprisingly, among those 12.7% with distressing symptoms, those randomized to take hormones reported significant relief compared with the women on placebo. The women who never had symptoms reported no relief of symptoms!

We have no way of knowing why the investigators associated with the WHI have been so determined and persistent in claiming...
that HRT is dangerous for most women, increasing the risks of breast cancer, heart disease, stroke, dementia, and cognitive impairment, while not even alleviating menopausal symptoms. These allegations run contrary to their own published data. In 2007, when it was reported that breast cancer rates in 2003 had declined, some investigators attributed it to the fall in HRT use following the 2002 WHI publication.208 Yet overall rates of breast cancer began to drop in 1999, and a decreasing death rate from breast cancer can be traced back to 1990, long before the initial publication of the WHI findings.201,202 Although we do not yet know all the complex factors required for a malignant cell to develop into a clinically detectable breast cancer, we know it takes more than 6 months—estimates range from 2 to 26 years, with an average of about 8 years.203–205 It is difficult to understand how a decrease in HRT use would be reflected in a decrease in breast cancer rates within a year. Finally, if the reported decreased incidence of breast cancer were due to a decrease in stimulation of subclinical (estrogen–induced or estrogen–stimulated) tumors, as proposed by the investigators, the decreased incidence should be confined to small early breast cancers. It is not.200

SUMMARY

What does all of this mean for women’s health? Concerns about HRT are valid, but HRT is not the clear and present danger that the WHI and much of the media have made it out to be. If women are going to stop taking HRT solely to avoid breast cancer, then, on the basis of the studies to date, they should also stop eating fish, consuming grapefruit, taking antibiotics, using electric blankets, or serving as flight attendants on Scandinavian airlines—all of which have been reported to have stronger associations with breast cancer than does HRT. On the other hand, cardiovascular concerns may be warranted, although largely among women who are at an elevated risk of heart disease or who begin HRT in their mid-60s.

The WHI concludes that the risks of HRT far outweigh the benefits, and even tried to hold HRT responsible for “increased deaths from all causes” in their 2008 report. None of these associations were statistically significant.206 Other investigators, though, feel just as strongly that the potential health benefits of postmenopausal estrogen replacement, as measured by decreased morbidity and increased life expectancy—by nearly 4 years, in 1 assessment—far exceed the risks.207 Even the WHI confirmed previously published reports of decreased risks of osteoporotic fractures and colon cancer for women on HRT.208,209

For us, the weight of the evidence is clear: women in menopause who have symptoms that seriously affect the quality of their lives should feel secure in taking HRT at the start of menopause and for as many years after as they must to control those symptoms. Any woman worried about her health and longevity should quit smoking before she quits hormones, and have screening mammograms and colonoscopies while she is at it. Years ago, Allen L. Hammond,210 then editor-in-chief of the American Academy of Science’s popular magazine, Science 80, described the challenge their journal faced: “Conveying the way science really works—the interplay of persistence and luck, the painstaking accumulation of evidence, the clash of proponent and critic, the gradual dawning of conviction—demands a look behind the headlines.” In an era when alarmist headlines get everyone’s attention, it is all the more important to read the fine print. Sometimes there is even good news there.

ACKNOWLEDGMENTS

This article was written at the suggestion of the American Council of Science and Health, whose panel of reviewers provided many helpful suggestions.

REFERENCES

146. Thomas DB. Do hormones cause breast cancer?


