Endocrine, *vol. 24, no. 3, 187–193, August 2004* 0969–711X/04/24:187–193/$25.00

The Million Women Study

A Critique

Malcolm Whitehead[1] and Richard Farmer[2]

[1]Consultant Gynaecologist, King's College Hospital NHS Trust, Denmark Hill, London SE5 9RS, UK; and [2]Professor of Epidemiology, Post Graduate Medical School, University of Surrey, UK

The regulatory authority in the UK, the Committee on Safety of Medicines, issued advice to health professionals on the day the results of the Million Women Study (MWS) were published. This course of action was undertaken before review of the study by the international medical and scientific community. We have reviewed the methodology of the MWS. It is our belief that the flaws in the design render the results largely uninterpretable because built in biases have affected risk estimates.

Key Words: HRT; breast cancer; Million Women Study.

Background

The results of the Million Women Study (MWS) were published in the *Lancet* in August 2003 *(1)*. On the same day UK Committee on Safety of Medicines (CSM) issued a press release and a letter to health professionals. The press release stated:

> A large research project, the "Million Women Study," provides important information on the risk of breast cancer in association with using hormone replacement therapy (HRT). The …CSM and its expert working group on HRT has considered these new facts and is issuing advice to practitioners and women.

It continues:

> the risk increases the longer it (HRT) is taken and starts to become apparent within 1–2 years of initiating treatment…. This study confirmed that …. this increase in risk begins to decline when HRT is stopped and, by 5 years returns to the same level as in women who have never taken HRT… (http://www.mca.gov.uk).

These same conclusions were then published as recommendations at the end of September 2003 in the official factsheet of the CSM "Current Problems in Pharmaco-Vigilance." All this activity took place before the methodology and interpretation of the MWS were critically reviewed in the correspondence section of the *Lancet* (18 October issue) and subsequently in many specialist journals.

The membership of the CSM's expert working group on HRT included one of the principal investigators for the MWS and the chair of the Steering Committee for the study; it did not include a breast oncologist or a tumor biologist (http://www.mca.gov.uk).

It is a matter for concern that the regulatory authority decided on a course of action before the publication of a scientific report and thus without the benefit of comments from the wider medical and scientific community. The CSMs advice undermined confidence in HRT and adversely affected the lives of hundreds of thousands of women, both in the UK and elsewhere. In this article we review the methodology of the MWS in an attempt to understand its relevance to clinical practice.

The Population and Methodology

In the UK every woman is invited to attend a local screening clinic for a mammogram every 3 yr between the ages of 50 and 65. The program is organized through the National Health Service Breast Cancer Screening Programme (NHSBSP). The MWS investigators sent questionnaires to approx 2 million women who were due to attend for screening between May 1996 and March 2001. They recruited 1,084,110, just over 50% of those eligible. The completed questionnaires were collected at the screening clinic and were then forwarded to the investigators' headquarters at Oxford where they were entered into the research database *(2)*. Thus, all information was self-reported. Small validation studies of certain aspects of the questionnaire had previously been performed by comparing self-reported data with information in the family physicians records *(3)*. The questionnaire focused on the most recent HRT product used and the life-time use of any HRT. Importantly, there was no provision for collection of data about prior use of other forms of HRT.

Each woman was "flagged" with the National Health Service Central Registers, which notified the investigators when the woman was registered as having developed invasive breast cancer (cancer registry) or died from invasive

Received August 2, 2004; Revised August 5, 2004; Accepted August 10, 2004.

Author to whom all correspondence and reprint requests should be addressed: Malcolm Whitehead. Consultant Gynaecologist, King's College Hospital NHS Trust, Denmark Hill, London SE5 9RS, UK. E-mail: malcolm. whitehead@btconnect.com

breast cancer (death register). For two thirds of the areas in the UK involved in the study, the last date for incidence "follow-up" was 31 December 2001. In the remaining third the last "follow-up" was 31 December 2000. Thus, for some of the women recruited in one third of the areas there was no follow-up because the women were recruited after the last update of the relevant cancer registers. The death registers were last updated on 31 December 2002.

No investigation was carried out on the completeness, accuracy, or timing of cancer registrations. No validation was undertaken on the completeness or accuracy of death registrations. From the published paper it appears that the incidence study (registrations) and mortality (death) study were kept separate, that is, a woman who was identified as having died from breast cancer would not be added to the incidence study unless she had been registered with cancer before death.

No attempt appears to have been made to establish the pattern of HRT use in the interval between the screening mammogram and the registration of invasive breast cancer or death from cancer. The investigators assumed that the type of HRT used at screening was the determinant of any breast cancer.

Results

The average age at recruitment was 55.9 yr. The average period of follow-up for cancer incidence was 2.6 yr and 4.1 yr for the analyses of mortality. On average, the breast cancers were diagnosed 1.2 yr after recruitment. The average time between diagnosis and death was 1.7 yr. At screening, women were classified into one of three categories: "current" users, "past" users, and "never" users.

The main analyses were restricted to 828,923 post-menopausal women; pre- and peri-menopausal women were excluded to reduce confounding due to spontaneous ovarian function. For ease of comprehension we have reproduced the findings set out in Figs. 1–4 of the original manuscript as Tables 1–4. The risk of breast cancer was 1.66 for "current" users of any HRT product. The risk among "past" users was not elevated (Table 1). Relative to "never" users, the risk of breast cancer was found to be elevated for "current" users of any type of product; the highest risk was associated with combined estrogen and progestagen preparations (Table 2). For all types of products the risk among "current" users was shown to increase with duration of use (Table 3). There was no dose response relationship among users of estrogen-only HRT and no significant difference between products (Table 4).

It should be noted that in their Table 4 the authors show the RRs for two estrogens (conjugated equine estrogens and ethinylestradiol); ethinylestradiol was not used as HRT in the UK during the period of the study. It is likely that this category actually refers to estradiol; however, no erratum has been published in clarification.

Table 1
Relative Risk of Incident Invasive Breast Cancer in Relation to Recency of Use of HRT

HRT use at baseline	Cases	Population	Relative risk	95% floating confidence intervals
Never user	2894	392,757	1.00	0.97, 1.04
Current user	3202	285,987	1.66	1.60, 1.72
Last use <5 yr previously	579	81,875	1.04	0.95, 1.12
Last use 5–9 yr previously	207	29,395	1.01	0.88, 1.16
Last use >9 yr previously	79	12,568	0.90	0.72, 1.12

Table 2
Relative Risk of Incident Invasive Breast Cancer in Relation to Recency and Type of HRT Use

HRT use at baseline	Cases	Population	Relative risk	95% floating confidence intervals
Never users	2894	392,757	1.00	0.96, 1.04
Past users	1044	150,179	1.01	0.95, 1.08
Current users				
Estrogen only	991	115,383	1.30	1.22, 1.38
Estrogen–progestagen	1934	142,870	2.00	1.91, 2.09
Tibolone	184	18,186	1.45	1.25, 1.67
Other/unknown	93	9,548	1.44	1.17, 1.76

Table 3
Relative Risk of Incident Invasive Breast Cancer in Relation to Recency, Total Duration of Use, and Type of HRT Used at Baseline

Total duration of use of HRT by type of HRT used at baseline	Cases	Population	Relative risk	95% floating confidence intervals
Never users	2894	392,757	1.00	0.96, 1.04
Past users				
<1 yr	311	47,606	0.94	0.84, 1.05
1–4 yr	384	55,823	1.01	0.92, 1.12
5–9 yr	230	29,614	1.14	1.00, 1.30
≥10 yr	80	11,654	1.05	0.84, 1.30
Current users of estrogen only				
<1 yr	25	4,452	0.81	0.55, 1.20
1–4 yr	251	29,582	1.25	1.10, 1.41
5–9 yr	416	47,310	1.32	1.20, 1.46
≥10 yr	277	31,862	1.37	1.22, 1.54
Current users of estrogen/progestagen				
<1 yr	97	9,771	1.45	1.19, 1.78
1–4 yr	582	49,240	1.74	1.60, 1.89
5–9 yr	850	56,912	2.17	2.03, 2.33
≥10 yr	362	23,673	2.31	2.08, 2.56
Current users of other/unknown				
<1 yr	19	1,728	1.63	1.04, 2.56
1–4 yr	83	8,794	1.34	1.08, 1.66
5–9 yr	102	10,342	1.42	1.17, 1.72
≥10 yr	59	4,739	1.93	1.50, 2.50

Table 4
Relative Risk of Incident Invasive Breast Cancer by Constituent, Dose, and Formulation of Estrogen-Only HRT Preparation Used at Baseline

Total duration of use of HRT by type HRT used at baseline	Cases	Population	Relative risk	95% floating confidence intervals
All estrogen-only HRT	991	115,383	1.30	1.21, 1.40
All equine estrogen	426	48,386	1.29	1.16, 1.43
≤0.625 mg equine estrogen	288	33,039	1.25	1.11, 1.41
>0.625 mg equine estrogen	135	15,181	1.36	1.14, 1.61
All ethinyloestradiol	454	56,322	1.24	1.12, 1.37
≤1mg ethinylestradiol	367	44,898	1.25	1.12, 1.40
>1mg ethinylestradiol	47	6,455	1.19	0.89, 1.58
By formulation				
Oral	606	68,351	1.32	1.21, 1.45
Transdermal	324	40,015	1.24	1.11, 1.39
Implanted	54	5,272	1.65	1.26, 2.16

The relative risk of fatal breast cancer was 1.22 (FCI 1.05–1.41); this was estimated from 191 deaths amongst 285,987 "current" users.

Design Issues

Some commentators were impressed with the size of the MWS and accepted the findings uncritically *(4–6)*. Others were more sceptical *(7–11)*. Van Leeuwen and Rookus *(10)* wrote: "The accrual of more than a million women into a study is near unprecedented and a major accomplishment. However, the mere size of a study population does not imply that its results should be taken at face value, particularly when they do not concur with other studies." We agree with this view. It is our opinion that there are major flaws in the design of the study and that these flaws explain in part why the risk estimates are not in agreement with those from other studies and appear to be inconsistent with the known biology of breast tumors.

The Study Population

In the UK, regional take-up rates of first invitations to clinics for screening mammography vary between 57% and 78% with a national average of 71% *(12)*. About 88% of women invited to attend for re-screening accepted the invitation. This study focused on women who were aged 50–64 yr between May 1996 and March 2001, i.e., women born between 1932 and 1951. Most of the women born between 1944 and 1951 would have been recruited at their first screening visit. Those born between 1932 and 1943 would have been recruited during the course of one of their periodic re-screening visits; it follows that 71% of eligible women born between 1944 and 1951 (aged 50–52 yr at recruitment) would have attended (although only 71% of these participated). Sixty-three percent of eligible women born between 1927 and 1943 would have been recruited at their re-screening visits; this figure is derived from the assumption that 71% attended for first screening and 88% of those attended for re-screening.

The rate of detection of breast cancer on the first screen was about 7%, whereas the rate on re-screening was about 6%. Most women in the study who were born between 1931 and 1946 would have been eligible for two screening visits during the study period. Those born after 1946 would be too young to qualify for more than one screen and those born before 1931 would be too old to qualify for more that one visit.

The number of screening visits will determine the probability of cancer being detected. The facts that the probability of a woman in the target age range being asked to participate varies according to year of birth and the probability of diagnosis also varies by year of birth will inevitably introduce bias. It is impossible to estimate the effect of this on the risk estimates from the material available in the published paper.

It is well recognized that take-up rates are higher in women who perceive themselves to be at an increase in risk for breast cancer and those who are health conscious. Women who use HRT tend to be more health conscious. In the MWS about half the women had used HRT; this is about 60% higher than in the general population of women in the UK *(13)*. The authors of the MWS accept that there was an excess of women using HRT included in their study, although they describe the excess as "slight." They also report that there was an excess of women from affluent backgrounds. Their view was that "this difference would not bias internal comparisons within the cohort." This is a matter of speculation rather than fact.

Exposure Categories

Women were classified as "current" users, "past" users, or "never" users at enrollment to the study and all the analyses were carried out on these "baseline" categories. A subsample analysis of 12,221 participants carried out part way

through the study revealed that 22% of "current" users at baseline subsequently ceased treatment, 19% of "past" users had started to take HRT again, and 11% of "never" users had started to take HRT. The authors argued that this would have little effect on their results in the following terms:

> Amongst current users at base line, the total duration of use of HRT at the time of diagnosis of breast cancer would be slightly longer than that recorded at baseline. However, any resultant underestimation of total duration of use of HRT would be counteracted, to some extent, by the fact that during follow-up some current users would have become past users and some never users would have become current users.

We do not agree. The investigators appear to have confused two issues. First, in at least 16% of the cases the use category (current, never, and past) at baseline was not the same as that when the cancer was diagnosed; 11% of the reference category (never users) were not "never users" at the time of the diagnosis of breast cancer, which means that the reference category is a heterogeneous group in terms of actual exposure. Thus, the true risk relative to this group is probably uninterpretable. The second issue is related to the duration of exposure of those classified as current users. We disagree with the author's assertion that "the total duration of use of HRT at the time of diagnosis …. would be slightly longer than at baseline." The total duration of use could be between 1 mo and 5 yr greater than that recorded at baseline depending on the year of recruitment and year of diagnosis. This could have the effect of moving women to different duration of exposure categories.

It is impossible to predict the effect of misclassification on the results. Clearly, because of the study design, the investigators have a problem. The only exposure data collected were at the time of the recruitment mammogram: no information on HRT was collected thereafter.

Incomplete Documentation of Hormone Use

The questionnaire used by the investigators is available on the web at www.millionwomenstudy.org.uk. Only limited data were collected on the life-time history of HRT use. Treated women were asked the name of their current HRT product, the total period that the product was used, and the total duration of life-time use. No data were collected on products other than that being used at the time of the mammogram. The analysis was limited to the current product but the duration of use categorization was based on total life-time use irrespective of any other products that could have been used. Thus, for example, a woman could have used tibolone for 3 mo prior to enrollment. Before that she could have used a combined sequential product for 5 yr. She would have been classified as a 5–9 yr user of tibolone because of the study design. Clearly the aggregation of all products in the determination of duration of use will distort the risk calculations.

In a sample of the study population, comparisons were made of the self-reported data with information from the family physician's records *(3)*. There was a 90% agreement regarding current use at recruitment, a 97% agreement regarding type of preparation (estrogen-alone as compared to estrogen/progestagen), and a 90% agreement for specific product and dose; 34% of current users at enrollment had previously used at least one other type of HRT preparation.

Because the authors collected data only on current use at the time of screening mammography, they can have no knowledge of HRT switching after mammography and prior to the diagnosis of invasive breast cancer or death. Indeed, the investigators have no means of validating that patients were still taking HRT at the time of diagnosis of invasive breast cancer or death from invasive breast cancer. Some women may well have stopped treatment soon after mammography. Switching of types of HRT is common, and women frequently stop for some months before re-starting often with a different therapy. Indeed, the General Practice Research Database (GPRD) from the UK suggests that switching between different HRT regimens occurs in up to 45% of women who have used HRT for 3 yr or more *(13)*.

Detection Bias

An additional source of bias would result from normal medical practice. Current users are more likely to be offered regular health checks in the primary care setting at the time their HRT prescription is renewed.

Effect on HRT on Mammographic Sensitivity

The authors clearly state that use of HRT reduces the sensitivity of mammographic screening and increases the diagnosis of interval cancers. Another critique *(14)* has challenged this belief. In reviewing seven studies it was observed that there were relatively few interval cancers in the HRT user groups. Nevertheless, six of the seven studies did report a decrease in mammographic sensitivity in hormone users with increases in interval cancers in users compared to non-users. However, one prospective study of screening mammography concluded that recall rates were essentially the same when hormone users were compared with non-users and that HRT rarely caused a diagnostic dilemma. One of the major problems with this entire topic is that most of the studies have only small numbers of interval cancers and it is uncertain how real or how large the effect of HRT is because of the difficulty in controlling for factors such as age, age at menopause, and time since menopause. If the HRT does decrease mammographic sensitivity to the extent that diagnosis is delayed, then this would result in an apparent decrease in incidence of breast cancer not, as the authors suggest, an increase.

Discussion and Interpretation of Results

The MWS demonstrated a statistically significant association between the use of HRT up to the time of routine

Table 5
Results from MWS and WHI Compared

	Million Women Study		Women's Health Initiative	
	Relative risk	95% floating confidence intervals	Relative risk	95% floating confidence intervals
Current users of estrogen/progestogen vs never users (MWS). Treated vs placebo (WHI)	2.00	1.22–2.09	1.24	1.01-1.54
Current users of estrogen only vs never users (MWS). Treated vs placebo (WHI)	1.30	1.22–1.38	0.77	0.57-1.06
Up to 1 yr of use of estrogen+ progestagen (MWS). Treated for 1 yr (WHI)	1.45	1.19–1.78	0.48	0.19-1.20
5–9 yr use of estrogen + progestagen (MWS) 5 yr after entry (WHI)	2.17	2.03–2.33	1.61	0.88-2.94

mammography and subsequent diagnosis of breast cancer. The authors extrapolated this observation and concluded that "Use of HRT by UK women aged 50-64 years in the past decade is estimated to have resulted in an extra 20,000 incident breast cancer cases, combined oestrogen and progestogen HRT accounting for 15,000 of these additional cancers." The design of the study has many limitations and it is necessary to ask whether the investigators estimate of the impact of HRT on breast cancer incidence is justified. The explanation for a statistical association between an exposure and a particular outcome is not always causal. A commonly used set of criteria that are used to distinguish between causal and non-causal associations were set out by Hill in 1965 *(15)*. These include strength of association, consistency, temporality and biological plausibility. Here we consider the results of MWS in the context of these criteria.

Strength of Association. Hill argued that a strong association is more likely to indicate cause than a weak one. A relative risk of around 2 in an observational study does not suggest a strong association.

Consistency. The authors of the MWS report argue that their findings are consistent with those of other investigators. In support of this assertion they cite 10 publications *(16,25)*. The re-analysis of 51 epidemiological studies *(16)* was carried out by the same team that was responsible for the MWS. Most of the studies included in the re-analysis were of estrogen-only therapy. A pooled relative risk of 1.14 (standard error 0.03) for ever-use of HRT was reported. Interestingly, the pooled estimated RR for prospective studies was 1.09 (not significant), for case-control studies with population controls it was 1.15 (not significant), and for case-control studies using hospital controls it was 1.27 (significant). This demonstrates the sensitivity of observational studies to the design. Within the plethora of relative risks in the paper few are significant.

Bush et al.'s paper *(26)* was not cited by the MWS authors. It does not support the view that HRT increases the risk of breast cancer. This review considered many of the papers included in the re-analysis *(16)*. It demonstrated that in the published literature there are as many studies reporting no increase in risk as there are reporting an increased risk of breast cancer associated with combined HRT and that most of the studies on oestrogen only HRT showed no increase in risk.

The authors of the MWS claim that their study results are consistent with those of the Women's Health Initiative Study *(16,17,27)*. Table 5 shows some comparative RR's from the two studies; the studies are not consistent with each other. The lack of consistency between studies weakens the case for inferring a causal association between HRT (particularly estrogen only) and breast cancer.

Temporality. Temporality is a relatively simple concept, exposure should pre-date the outcome. This is a straightforward issue when dealing with an outcome that immediately follows the exposure—for example, a car accident and a fracture. It is much more difficult when dealing with cancer where there may be a long interval between the exposure and start of the disease process (induction period) and there is a further interval between the start of the disease process and its detection (the latent period). The latent period can be reduced by early detection but the induction period will not be. Neither the induction nor the latent periods are precisely known for cancer of the breast but some inferences can be made from studies of tumour growth rates.

Kopernik and Shoham *(28)* cited two studies concerned with induction and latent periods for breast cancer. Heuser et al. *(29)* estimated that the interval between induction of a tumor and it developing to a size which permitted detection by mammography to be about 6.8 yr: the interval between mammographic detection and detection by palpation was calculated at approx 3.4 yr. Gullino *(30)* estimated an interval of 10 yr between a cell transformation and a clinically detectable tumor. The MWS reported that women had a significant increase in risk of breast cancer within a year of starting combined HRT. If the combined induction and

latent period is 9–10 yr, then the start of the disease must have predated any exposure to HRT. Therefore, it follows cause cannot be inferred in the MWS. The apparent increase in breast cancer risk observed in MWS could be explained by an acceleration of growth of an existing tumor rather than the induction of a new tumor. Were this to be the case, then there would be a reduction in the latent period rather than the induction period. Heuser et al. *(29)* found tumor volume doubling times varied widely (between 109 and 944 d with a mean of 325). Even taking the shortest doubling time, it is unlikely that a growth rate could be achieved that would transform undetectable tumors to detectable tumors within 12 mo. It follows that the hypothesized growth induction would have started before exposure to the HRT for those cases that occurred within a year of initiation of HRT. We conclude that the elevated risk found among short-term users cannot be consistent with there being a causal association.

Biological Plausibility. The issue of biological plausibility is closely related to that of temporality. To sustain a causal hypothesis it is necessary to postulate an induction period far shorter than has been demonstrated in studies of tumor biology. If the hypothesis of growth acceleration is to be sustained, then it is necessary to revise our current knowledge of tumor growth.

The observation in MWS that cessation of HRT results in the risk of breast cancer returning to the same level as that seen among women who have never been exposed is not in agreement with our knowledge of tumor biology. There is no precedent for breast tumors regressing after the causal agent has been removed. If a tumor has been induced, then it is likely to continue to grow whether or not the causal agent is present. If the hypothesis is that the HRT accelerates tumor growth, then it could be postulated that removal of the HRT might result in a return to the "normal" growth rate. The effect would be to reduce the apparent incidence to lower than that among women continuing to use HRT but not for it to revert to that of women who had never used HRT, unless it is proposed that the post-HRT growth rate is significantly less than that among women not exposed to HRT.

Conclusions

The Million Women Study reports a weak association between HRT use and breast cancer. The results are not consistent with those from many other studies: in particular, there are few studies that show an increase associated with estrogen-only therapy. The results are neither consistent with a hypothesis that tumor induction followed exposure nor with the hypothesis that HRT accelerated tumor growth because the temporal relationship between the outcome and the event (which must have preceded its detection) is wrong. The disappearance of risk after cessation of therapy is not biologically plausible. We conclude that the study should not be interpreted as indicating that HRT either causes or accelerates the growth of breast tumors.

The design of the study was unusual. Normally it would be expected that the design of a study should be driven by a hypothesis: in this study it appears that the investigators were simply looking for associations between exposure to HRT and breast cancer incidence. Were they to have been investigating the possibility of tumor induction, then it would be expected that data on the exposure up to the time of the hypothetical date of induction should have been collected rather than exposures that would clearly postdate induction. It would also be expected that account would be taken of the differences in lead time between mammographically detected tumors and clinically detected tumors. In the event crude exposure data were collected up to the date of a screening mammogram, which was carried out at a point in time that had no fixed relationship to the woman's date of menopause, HRT use, or occurrence of breast cancer. The design of the study is such that the results are uninterpretable as it is impossible to predict how the effect of the built-in biases in the study have affected the risk estimates.

References

1 Beral, V. and Million Women SC. (2003). *Lancet* **362,** 419–427.
2. The Million Women Study Collaborative Group. (1999). *Breast Cancer Research* **1,** 73–80.
3 Banks, E., Beral, V., Cameron, R., et al. (2001). *J. Epidemiol. Biostat.* **6,** 357–363.
4 Vandenbroucke, J. P., van Leeuwen, F. E., and Helmerhorst, F. M. (2003). *Nederlands Tijdschrift voor Geneeskunde* **147,** 1829–1834.
5 Doren, M. (2003). *Maturitas* **46,** 93–94.
6 Lagro-Janssen, T., Rosser, W. W., and van Weel, C. (2003). *Lancet* **362,** 414–415.
7. Garton, M. (2003). *Lancet* **362,** 1328, (author's reply) 1330–1331.
8 Bliss, J. M. and Gray, R. (2003). *Lancet* **362,** 1328–1329.
9. Bundred, N. J. and Morris, J. (2003). *Lancet* **362,** 1329, (author's reply) 1330–1331.
10 Van Leeuwen, F. E. and Rookus, M. A. (2003). *Lancet* **362,** 1330–1331.
11 Braendle, W. (2003). *Maturitas* **46,** 101–102.
12. Patnick, J. (2003). *Breast Screening Programme, England, 2002–2003*.
13 Bromley, S. E., de Vries, C. S., and Farmer, R. D. (2004). *BJOG* **111,** 369–376.
14 Speroff, L. (2003). *Maturitas* **46,** 1–6.
15 Hill, A. B. (1965). *Proc. R. Soc. Med.* **58,** 295–300.
16. Collaborative Group on Hormonal Factors in Breast Cancer. (1997). *Lancet* **350,** 1047–1059.
17 Rossouw, J. E., Anderson, G. L., Prentice, R. L., et al., Writing Group for the Women's Health Initiative Investigators. (2002). *JAMA* **288,** 321–333.
18 Chlebowski, R. T., Hendrix, S. L., Langer, R. D., et al., Investigators WHI. (2003). *JAMA* **289,** 3243–3253.
19 Beral, V., Banks, E., and Reeves, G. (2002). *Lancet* **360,** 942–944.
20 Beral, V., Banks, E., Reeves, G., and Appleby, P. (1999). *J. Epidemiol. Biostat.* **4,** 191–210.
21 Bergkvist, L., Adami, H. O., Persson, I., Hoover, R., and Schairer, C. (1989). *N. Engl. J. Med.* **321,** 293–297.
22 Magnusson, C., Baron, J. A., Correia, N., Bergstrom, R., Adami, H. O., and Persson, I. (1999). *Int. J. Cancer* **81,** 339–344.

23 Schairer, C., Lubin, J., Troisi, R., Sturgeon, S., Brinton, L., and Hoover, R. (2000). *JAMA* **283,** 485–491.
24 Ross, R. K., Paganini-Hill, A., Wan, P. C., and Pike, M. C. (2000). *J. Natl. Cancer Inst.* **92,** 328–332.
25 Li, C. I., Malone, K. E., Porter, P. L., et al. (2003). *JAMA* **289,** 3254–3263.
26 Bush, T. L., Whiteman, M., and Flaws, J. A. (2001). *Obstet. Gynecol.* **98,** 498–508.
27 Anderson, G. L., Limacher, M., Assaf, A. R., et al., Women's Health Initiative Steering Committee. (2004). *JAMA* **291,** 1701–1712.
28 Kopernik, G. and Shoham, Z. (2004). *Fertil. Steril.* **81,** 1458–1477..
29 Heuser, L., Spratt, J. S., and Polk, H. C. Jr. (1979). *Cancer* **43,** 1888–1894.
30. Gullino, P. M. (1977). *Cancer* **39,** Suppl.-703.