

# A critique of Women's Health Initiative Studies (2002-2006)

James H. Clark 

 Corresponding Author: clarksciconsl@sbcglobal.net

Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas, USA

**The Women's Health Initiative Studies (WHI) were designed to examine the effects of estrogen and progestin (E+P; Prempro) and estrogen alone (Premarin) in post-menopausal women. The authors of the WHI studies and the National Heart Lung and Blood Institute (NHLBI) concluded that E+P treatment increased the risks of coronary heart disease, invasive breast cancer, stroke and venous thromboembolism. The following paper contains a reevaluation of these studies based on the graphic analysis of their tabulated data. In contrast to the conclusions reached by the WHI and the NHLBI, I conclude that treatment of post-menopausal women with estrogen and progestin (Prempro) does not increase the risks of cardiovascular disease, invasive breast cancer, stroke or venous thromboembolism. I also disagree with the claim that an increased risk of stroke existed in women treated with estrogen alone.**

Received July 14th, 2006; Accepted October 13th, 2006; Published October 30th, 2006 | **Abbreviations:** 95% CI: 95% confidence interval; CHD: coronary heart disease; E: estrogen; HR: hazard ratio; IBC: invasive breast cancer; NHLBI: National Heart, Lung and Blood Institute; P: progestin; VTE: venous thromboembolism; WHI studies: Women's Health Initiative studies | Copyright © 2006, James H. Clark. This is an open-access article distributed under the terms of the Creative Commons Non-Commercial Attribution License, which permits unrestricted non-commercial use distribution and reproduction in any medium, provided the original work is properly cited.

**Cite this article:** Nuclear Receptor Signaling (2006) 4, e023

## Introduction

The purpose of this paper is to point out problems with the data analysis and conclusions drawn by the authors of the Women's Health Initiative (WHI) studies published over the time period 2002-2006. The WHI was a randomized placebo controlled clinical trial to examine the effects of estrogen plus progestin on various health benefits and risks in post-menopausal women [Rossouw et al., 2002]. The trial was designed to last for 8.5 years but was stopped at 5.2 years because the National Heart, Lung and Blood Institute (NHLBI) accepted the study's Data and Safety Monitoring Board (DSMB) evaluation that the risks of breast cancer were increased and other health risks of hormone treatment were greater than the benefits [NHLBI, 2002]. Additional papers which extended the evaluation of these data were published on two subgroups of stroke [Wassertheil-Smoller et al., 2003], coronary heart disease [Manson et al., 2003] and breast cancer [Chlebowski et al., 2003].

The conclusions drawn from the first paper have been widely accepted as demonstrating that combined estrogen plus progestin (E+P: Prempro: 0.625 mg conjugated estrogens and 2.5 mg medroxyprogesterone acetate) replacement treatment in postmenopausal women increased the risk of invasive breast cancer (IBC), coronary heart disease (CHD), stroke and venous thromboembolism (VTE) [Fletcher and Colditz, 2002; Gann and Morrow, 2003; National Cancer Institute, 2002; NHLBI, 2002; Pedersen and Ottesen, 2003]. The conclusions in these papers were derived from Kaplan Meier estimates of cumulative hazards and Cox proportional hazard models [Cox, 1972]. These are the accepted methods for analyzing such data; however, I feel the conclusions drawn were incorrect and statistical methods used require modification.

The WHI also included a parallel study which examined the risk versus benefits of estrogen treatment alone (Premarin) in post-menopausal women. This arm of the WHI was continued until February 2004, which was approximately eight months earlier than the intended termination date of October 2004 to March 2005 [Anderson et al., 2004].

The authors concluded that use of conjugated equine estrogens increased the risk of stroke, decreased the risk of hip fracture and did not affect CHD and may reduce the risk of breast cancer. In the final paper on CHD it was concluded that there might be a protective effect in 50-59 year old women [Hsia et al., 2006]; that there was no increase in breast cancer [Stefanick et al., 2006] and that an early increase in venous thrombosis is associated with the use of estrogen alone [Curb et al., 2006]. As in the E+P studies cited previously, I believe these conclusions to be either incorrect or require important modifications.

## Relative risks versus absolute risks

One of the major problems with the interpretation and misunderstandings of this study was that it emphasized the use of relative risks rather than absolute risks. The latter provides a much more reasonable and understandable appraisal of the data. For example, the authors state that there was a 29% increase in CHD in the E+P group compared to the placebo group. This 29% value is a relative risk and is derived from the Hazard Ratio (1.29), which is calculated from the Cox Regression analysis used by these authors. Relative risk can be misleading and many people do not realize that the actual increases are very small and may be insignificant. The actual % increase is 0.07% which can be calculated from the following information:

For the E+P group there was a 0.37% incidence of CHD in a population of 10,000. Thus the number of individuals affected was 37. Since the incidence in the placebo group was 0.30% of 10,000, the number affected was 30. So the actual increase in patients with CHD was 7/10,000 or a 0.07% increase in absolute risk. Note that these increases were judged in the final analyses to be statistically insignificant. This is especially important when these results are presented by the media to the general public. The relationships between hazard ratios, relative risks and absolute risks are shown in Table 1. Some scientists considered such small absolute risks as statistically insignificant, and when all the results were published this supposition was correct.

This difference between relative and absolute risks has been pointed out by others, but has received little attention or emphasis [Goodman et al., 2003; Kelly, 2002; McDonough, 2002]. The Consolidated Standards of Reporting Trials (CONSORT) indicated that authors should state results as absolute, not just relative numbers [Moher et al., 2001]. In addition, many authors agree that unless a hazard ratio is 3.0 or higher it is of no consequence [Taubes, 1995]. Therefore, by this standard, all of the hazard ratios in this study would be considered insignificant.

## Coronary heart disease (CHD) in the 2002 paper

### Risk ratio

As discussed previously, the authors report a hazard ratio of 1.29 or a 29% relative risk for CHD in the E+P treated group. They state that this value reached nominal statistical significance; however, this statement is based on nominal (unadjusted) confidence intervals which should not be used in this type of study as indicated by them in their paper on page 325. If adjusted confidence intervals had been used, the HR of 1.29 would not have reached statistical significance.

Most of the WHI papers display data in tables as incidence of disease per year of hormone or placebo treatment. I have taken the incidence data for the two groups and calculated risk ratios (% incidence in hormone group/% incidence in the placebo group) and plotted these as a function of time in years (Figure 1). It is apparent that 4 of the total 6 values for CHD are equal to or very near to the no effect value of 1.0 (Figure 1A). The peak in risk ratio at year 5 is not due to a significant increased incidence of CHD in the E+P group, but to a decreased incidence in the placebo group (Figure 1B).

The final hazard ratio of 1.29 has an adjusted 95% confidence interval of 0.85-1.97. Such a broad confidence interval, which includes 1.0, and the fact that four of six risk ratio values are not different from the no effect level, demonstrate that there is no significant increase in absolute risk. As discussed below, the authors conclude in a later paper that the risks for CHD are not significant (Figure 2A and Figure 2B below; [Manson et al., 2003]).

## Incidence

The risk ratios in Figure 1A are derived from the erratic incidence values for CHD which increase and decrease between 0.2 and 0.4% as a function of time (Figure 1B). The values for both groups vary over a similar range and the differences are very small. Since the WHI study was done with older women, averaging 12 years post-menopausal, the elevated value for the risk ratio at year 1 may be due to the presence of asymptomatic atherosclerosis as a result of the loss and absence of endogenous estrogen following the onset of menopause [Grodstein et al., 2006; Harman et al., 2005; Turgeon et al., 2004]. Judging from variable incidence values in Figure 1B for the placebo group, the elevated risk at year 1 could be due to chance. Certainly the sudden peak in the risk ratio at year 5 is not due to a sudden increase in CHD, but instead, to an accompanying drop in values for the placebo group (Figure 1B). It seems obvious that no significant differences exist between incidence of CHD in the E+P and placebo groups when these data are viewed in a graphic format; however, such differences are not so obvious when the data are in table form.

## Coronary heart disease (CHD) in the 2003 paper

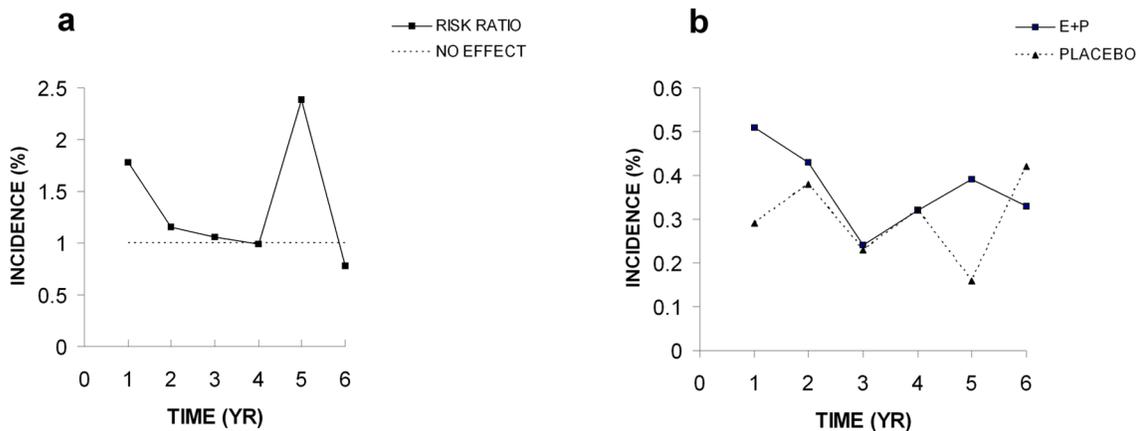
In a final report on CHD the WHI authors provide an analysis which does not support the original statement of a significant increased risk of CHD in women treated with E+P [Manson et al., 2003]. Instead, they conclude that E+P treatment may increase risks of CHD during the first year, but after that the risk decreases with time and this downward trend is statistically significant. The CHD hazard ratios for each year from that paper are shown in Figure 2A. The unadjusted 95% CI for these HR values are very wide and all but the first year value include the no effect level of 1.0. If adjusted 95% CIs had been used, all would have included 1.0. The peak at year 5 has decreased because the incidence of CHD in the placebo group has increased (compare Figure 2B with Figure 1B). As stated above, the incidence of CHD in the placebo group, not an increase in incidence of CHD in the E+P, accounted for the 5 year peak in risk in the 2002 paper. The final hazard ratio was 1.24 or an absolute increase of 0.06%, which is equal to 6 additional CHD events per 10,000 person years. They conclude that no significant increased risk of CHD is associated with E+P treatment. I believe this conclusion could have been reached in 2002 when the study was prematurely halted.

## Coronary heart disease in the estrogen only study

In this study the authors conclude that estrogen alone does not affect the risk of CHD in post-menopausal women (HR, 0.91; 95% CI, 0.75-1.12; [Anderson et al., 2004]). The final results of the estrogen alone study were divided into age groups of 50-59, 60-69 and 70-79 years [Hsia et al., 2006]. The conclusion was that estrogens provide no protection against CHD with the possible exception of those in the 50-59 age group (HR, 0.61; 95% nominal CI, 0.25-1.50). However, the incidence and risk

OUTCOME	HAZARD RATIO + 95% CI	RELATIVE RISK (%)*	ABSOLUTE RISK (%)*	NUMBER* /10,000**
CORONARY HEART DISEASE	1.29 0.85-1.97	+29	+0.07	7.0
INVASIVE BREAST CANCER	1.26 0.83-1.92	+26	+0.08	8.0
VENOUS THROMBOEMBOLISM	2.11 1.26-3.55	+111	+0.18	18.0
STROKE	1.41 0.86-2.31	+41	+0.08	8.0
ENDOMETRIAL CANCER	0.83 0.29-2.32	-17	-0.01	1.0
COLORECTAL CANCER	0.63 0.32-1.24	-37	-0.06	6.0
HIP FRACTURE	0.66 0.33-1.33	-34	-0.05	5.0

**Table 1. Relationship between hazard ratios, relative risks and absolute risks in the WHI study.** These data were taken from Writing group for the WHI, 2002. \* None of the numbers were judged to be statically significant when all data analysis had been completed; therefore, they do not constitute actual increases in risks. The only possible exception is the risk of hip fracture, which was judged by the WHI authors to be significantly reduced. \*\* The number of outcomes is expressed per 10,000 for ease in numerical comparison. The actual number of individuals per group was 8,506 and 8,102 in the E+P and placebo groups, respectively.



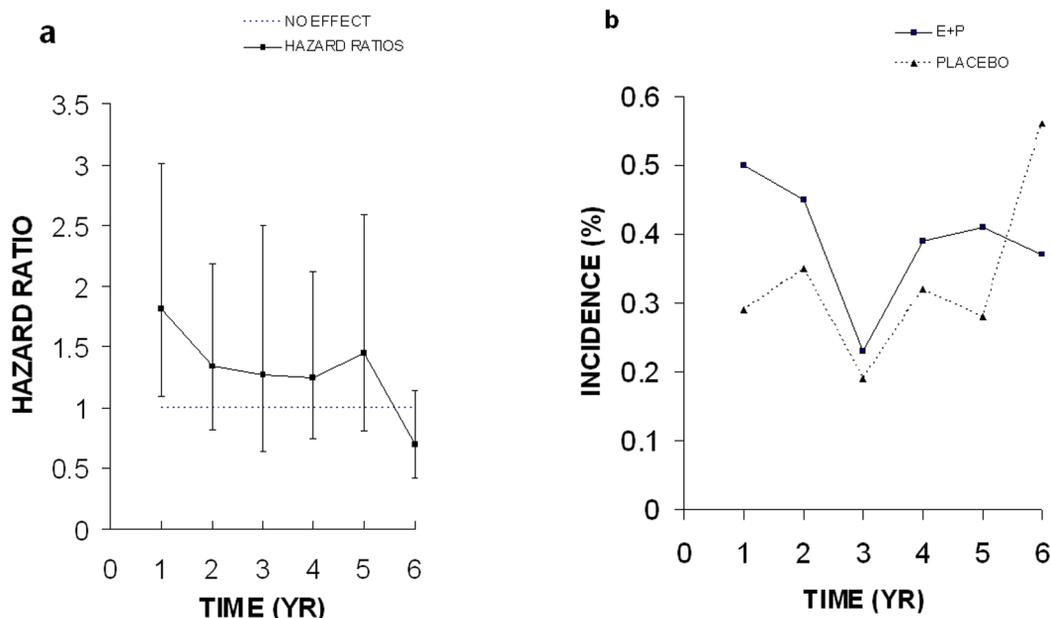
**Figure 1. Coronary heart disease (CHD) in 2002 paper.** A. Risk ratios calculated from incidence data. B. Incidence of CHD. All data were taken from Table 4 in the WHI 2002 publication.

ratio data for CHD in each of these groups is more erratic and variable than any of the data shown thus far. These results will be the subject of another paper and will not be discussed further here.

### Invasive breast cancer (IBC) 2002 paper Risk ratio

The value of 26% increase in the relative risk of invasive breast cancer in the E+P group has been cited over and

over by many people in the scientific and non-scientific media, even though the authors of the WHI paper acknowledge that it “almost reached nominal statistical significance”. Since “almost” is not statistical significance, the statement should have been: there was no significant difference in IBC risk between the placebo group and the E+P group. As in the analysis of CHD, if the authors had used adjusted confidence intervals there would be no doubt that risks were not increased.



**Figure 2. Coronary heart disease (CHD) 2003.** A. Risk ratios and unadjusted 95% confidence intervals. B. Incidence of coronary heart disease (CHD). All data were taken from Table 2 in the WHI 2003 publication.

The authors then indicate that “the weighted test statistic used for monitoring was highly significant”. This statistic would not have been elevated if the authors had examined the data more carefully. The apparent increase in risk ratios from years 2-5 is accompanied by a decline in the placebo groups (Figure 3A and Figure 3B). As discussed below, in the final analysis of these data this upward trend is not statistically significant (Figure 4A and Figure 4B). The final hazard ratio of 1.26 has an adjusted 95% confidence interval of 0.83-1.92, and the absolute risk increase is 0.08% or 8/10,000 person years. Such a broad confidence interval which includes 1.0 indicates there is no significant increase in risks due to hormone use. In addition, mere inspection of the data in Figure 3A clearly shows that four of the six values are not different from the no effect level, thus making it very unlikely that any real differences in risk existed.

### Incidence (detection)

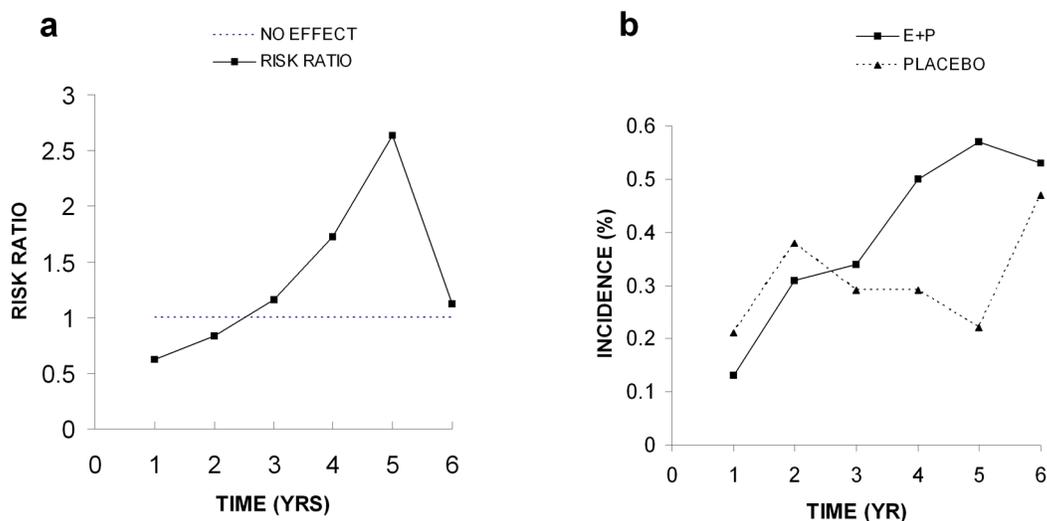
The percent incidence or detection, a more valid term, of breast cancer was similar up to year 3 with the detection level for the E+P group below that of the placebo group in years 1-2 (Figure 3B). By years 4 and 5 the two groups appear to diverge with the E+P group being greater than the placebo group. However, by year 6 the two groups came back together and were not different from one another. This divergence and convergence of detection at these times accounts for the spike in the risk ratios seen in Figure 3A. As discussed above, it is difficult to see how such data could be considered to be predictive of risk since 4 of the 6 risk ratio values are not different from the no effect level of 1.0 and during the last year the detection level in the placebo group increased dramatically.

The term detection is used instead of incidence because incidence implies cause and it is very unlikely that any breast cancer detected in this study was initiated by the hormone therapy. Other investigators have pointed out that the time between initiation of breast cancer and the time that cancer can be detected varies between 10 to 20 years [Dietel et al., 2005; Koscielny et al., 1985; von Fournier et al., 1980].

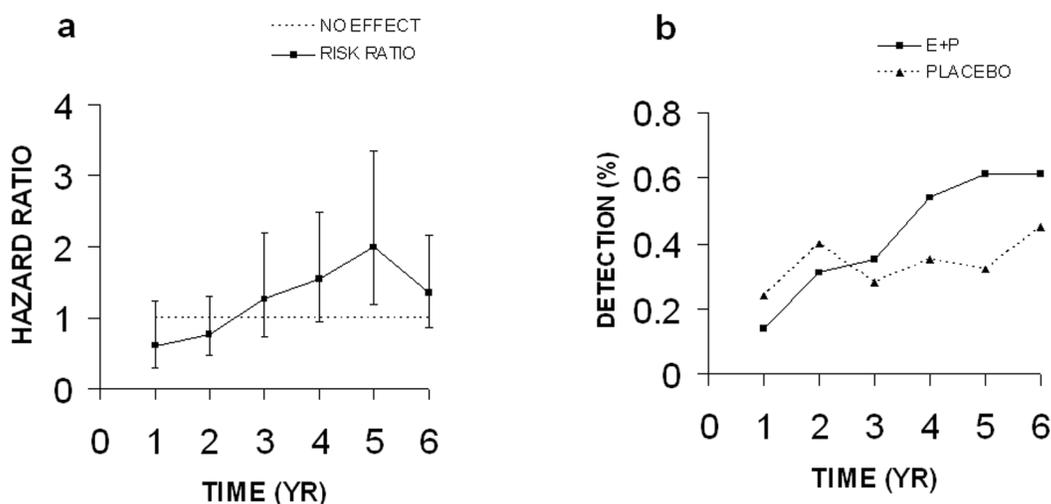
### Invasive breast cancer (IBC) 2003 paper

In a subsequent more detailed paper on the effects of E+P treatment on IBC the authors provide hazard ratios and their respective non-adjusted 95% CI intervals in table form (Figure 4A; [Chlebowski et al., 2003]). When these data are compared to those in Figure 3A we can see the hazard ratios for years 4 and 5 have been reduced from 1.73 to 1.54 (CI, 0.95-2.49) for year 4, and 2.64 to 1.99 (CI, 1.18-3.35) for year 5. The value for year 5 appears to be significant; however, the 95% CI is not adjusted, therefore it is not possible to judge with certainty. If an adjusted 95% CI had been used it probably would have included 1.0.

These lower HR values and broad non-adjusted 95% confidence intervals provide even more reasons to doubt the conclusions drawn by the WHI authors and the NHLBI that a significant increase in risks for IBC occurred in this study. Such doubt is amplified by the data on hazard ratios for IBC in women who had used or had not used post-menopausal hormones before the study (Figure 5A and Figure 5B). The hazard ratios for women with prior use are very variable and the unadjusted 95% confidence intervals are extremely large. Therefore, it seems unlikely that any meaningful conclusion can be drawn from this



**Figure 3. Invasive breast cancer (IBC) 2002.** A. Risk ratios. B. Detection of invasive breast cancer (IBC). All data were taken from Table 4 in the WHI 2002 publication.



**Figure 4. Invasive breast cancer (IBC) 2003.** A. Hazard ratios. B. Detection of invasive breast cancer (IBC) in the all women regardless of prior hormone use. All data were taken from Table 2 in the WHI 2003 publication.

group. In the group without prior use of hormones the hazard ratios look almost identical to those in Figure 2A, however the 95% CIs are larger. Four of the six HR values are below or near the no effect level and all 95% CIs overlap it. If adjusted CIs had been used, all would have included 1.0. Therefore, it seems impossible to draw any meaningful conclusions from such data.

### Invasive breast cancer in the estrogen only studies

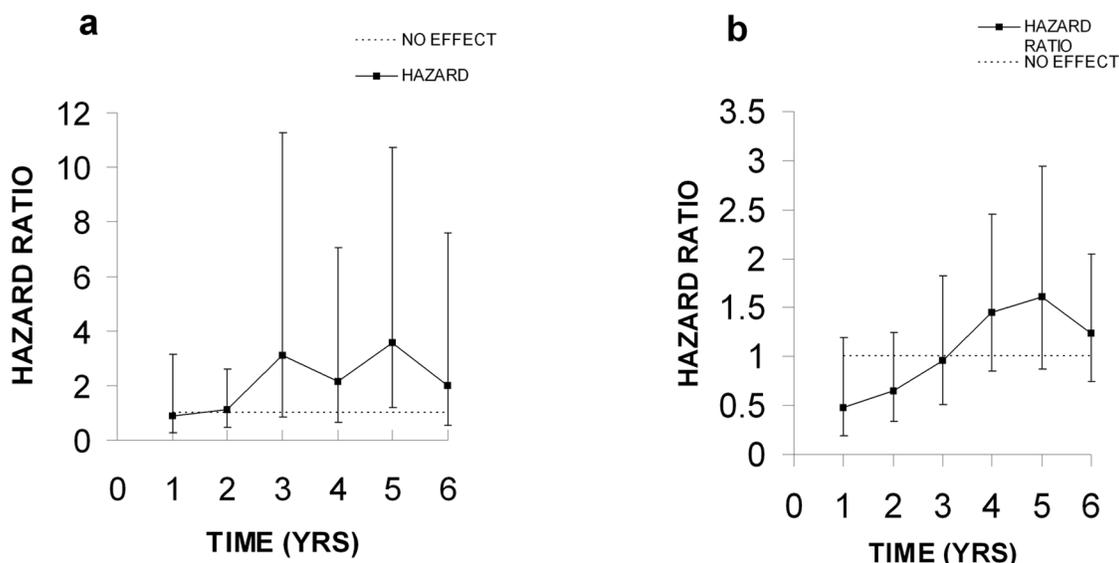
In the estrogen only arm of the WHI study invasive breast cancer was decreased by estrogen treatment [Anderson et al., 2004]. The hazard ratio was not statistically significant: 0.77 (CI, 0.59-1.01). A protective effect may be likely since the number of risk ratios which were near

or below the no effect level were greater than those above this level (data not shown). In the final report on this aspect of the study similar data and conclusions were reached [Stefanick et al., 2006].

### Venous thromboembolism (VTE) 2002 paper

#### Risk ratio

Similar criticisms to those above can be made of the VTE data. The authors indicate that there was a 2 fold increased risk of VTE in the E+P group (Figure 6A). This 2 fold increase constitutes an absolute risk of 0.18% for VTE in the E+P group compared to the placebo group. The final hazard ratio was reported as 2.22 with an adjusted 95% confidence interval of 1.26-3.55. The



**Figure 5. Invasive breast cancer (IBC) in women with prior or no prior use of hormones (2003).** A. Risk ratios for invasive breast cancer (IBC) in women with prior use of hormones. B. Detection of IBC in women with no prior use of hormones. All data were taken from Table 2 in the WHI 2003 publication.

authors do not indicate whether they believe this value is significant. However, as indicated previously, such broad confidence intervals in association with an overall decrease in risk ratios over time make it unlikely that these data demonstrate any actual risk.

### Incidence

The downward trend in risk ratios discussed previously is the result of a relatively higher incidence of VTE during the first year in the E+P group (Figure 6B). These higher values decline with time and are not different from the placebo group in year 6. These data show an overall downward occurrence of VTE in the E+P group. Certainly not an increased risk as claimed in the WHI paper in the NHLBI news release. As discussed above, these early occurrences may be due to the presence of asymptomatic atherosclerosis or clotting defects in these older women.

### Venous thromboembolism in the estrogen only study

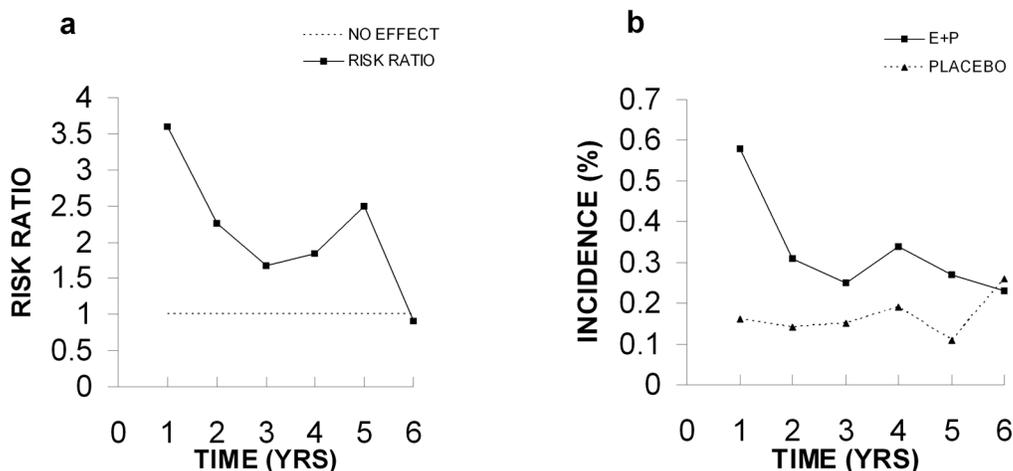
No data on a yearly basis were published for VTE in the estrogen only study; therefore, it was not possible to graph risk ratios or percent incidence as a function of time [Anderson et al., 2004]. The authors indicate the final hazard ratio was 1.33 (95% CI, 0.86-2.08) and that this was not significant. However, they say that the risk for the subgroup, deep vein thrombosis (DVT), is significant (HR 1.47; CI, 0.87-2.47). Since no yearly data were provided for DVT it was not possible to draw a graph; however, the authors did provide yearly data for pulmonary thrombosis (PE) which show the same erratic risk ratio and incidence values as in most of their other data (data not shown). Therefore, it is likely that the data for deep vein thrombosis shows similar, if not greater variation. This likelihood, plus the small absolute increase (0.06%) and the broad confidence intervals which cross 1.0, make it difficult to accept these values as significant.

This expectation of a high degree of variability and uncertainty was borne out by the data in the final paper from the WHI studies on venous thrombosis [Curb et al., 2006]. In this paper the authors provide hazard ratios and non-adjusted 95% confidence intervals for DVT, PE and venous thrombosis, VT (Figure 7). If adjusted 95% CIs had been used, all values would have included 1.0 and would have been judged insignificant. It is clear why the authors of this paper make no statement concerning statistical significance. Instead, they state that VT risk is associated with the use of estrogen during the first two years of exposure. It is clear that the very wide non-adjusted confidence intervals associated with the 0-2 year span for all three groups make it impossible to conclude anything concerning this period. The later time periods show no increased risk due to hormone treatment. It is puzzling why the authors in the 2004 paper conclude that the HR for DVT is significant and yet in the 2006 paper they conclude the HR is not significant, yet the data are virtually identical.

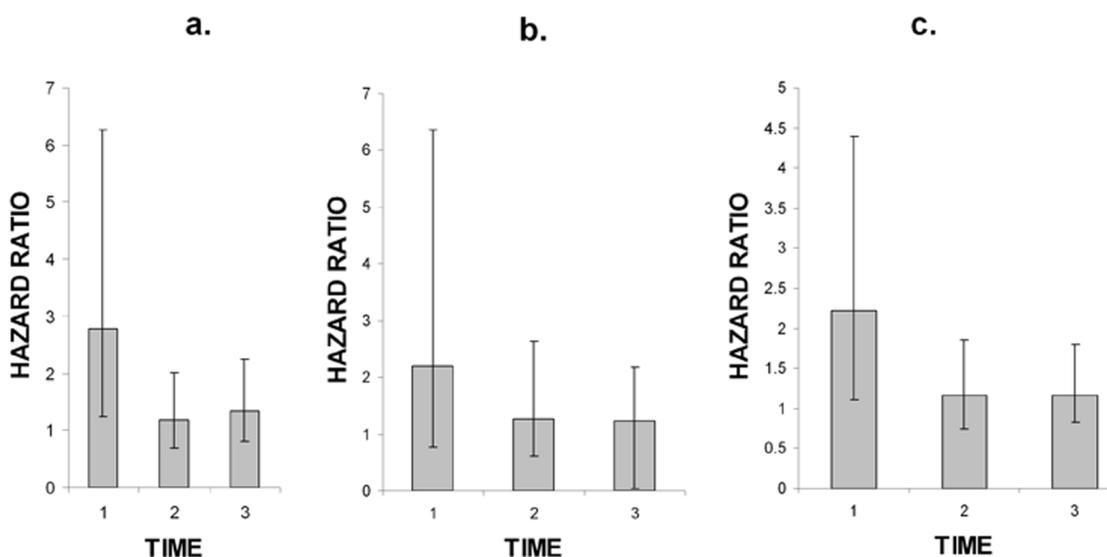
### Stroke 2002 paper

#### Risk ratio

The risk ratios for stroke in the E+P group increased from no effect at the first year to 1.72 by the second year and remain at about this level for the next four years (95% confidence intervals for each year were not published). Then by year 6 the risk ratio is below the no effect level (0.66; Figure 8A). The final hazard ratio determined by the WHI authors of 1.41 has an adjusted 95% confidence interval of 0.86-2.31. Such broad confidence intervals which include 1.0 are not significant. Once again the absolute risk is extremely low (0.08%), making it unlikely that these data demonstrate any clinically significant risk. The authors of the 2002 paper make no judgment as to the statistical significance of these data; however, the



**Figure 6. Venous thromboembolism (VTE).** A. Risk ratios, and B. Incidence of venous thromboembolism (VTE). All data were taken from Table 4 in the WHI 2002 publication.



**Figure 7. Hazard ratios for deep vein thrombosis [A], pulmonary embolism [B] and venous thrombosis [C].** Error bars show the 95% CIs for the hazard ratios. Data taken from Table 6 in Curb et al 2006. Time periods: 1 = 0-2 yrs, 2 = 2-5 yrs, 3 = 5 or greater yrs.

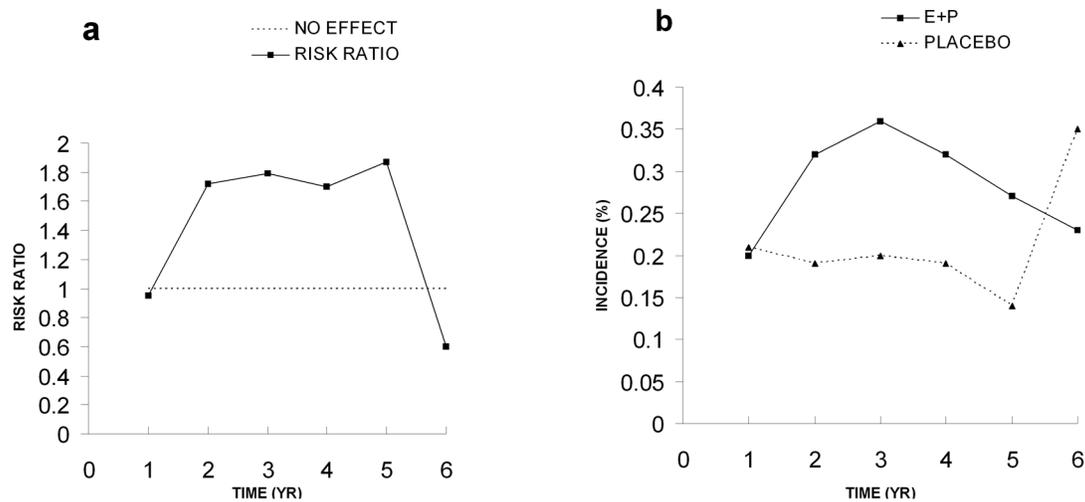
NHLBI assumed the increase of 41% was significant and indicated this in their news release.

In a subsequent paper, the effects of E+P on ischemic and hemorrhagic stroke were examined [Wassertheil-Smoller et al., 2003]. Since no data for risk ratios or incidence for each year of the study were provided, no graphs similar to those in Figure 8 above could be drawn. The final hazard ratio for total strokes was 1.31 (95% CI, 1.02-1.68), which the authors claim is significant. However, the adjusted 95% CI was 0.93-1.84, which makes this value statistically insignificant. The hazard ratios for hemorrhagic and ischemic stroke were judged to be statistically insignificant.

## Stroke in the estrogen only study

### Risk ratios and incidence

The risk ratios for stroke in this study are low but generally above the no effect level (Figure 9A). These minimal ratio values are the result of very variable incidence levels, which indicate that these groups were not different from one another for the first five years of the study (Figure 9B). This period is followed by a small increase in the estrogen group at year 5, which is followed by a steady decline to low levels equal to those of the placebo group. Such declines in the risk ratio and incidence in the estrogen group suggest a beneficial effect of estrogen treatment. The authors indicate that the final hazard ratio



**Figure 8. Stroke.** A. Risk ratios. B. Incidence of stroke. All data were taken from Table 4 in the WHI 2002 publication.

of 1.39 was significant; however, this was based on non-adjusted 95% CI. When the adjusted 95% CI is used (0.97-1.99), the HR becomes statistically insignificant.

### Colorectal and endometrial cancer

In the E+P and E-alone studies the number of women is small and the incidence levels are in the 0.1% range (data not shown). The final hazard ratio for the E+P studies was 0.63 (95% CI, 0.32-1.24) and was considered by the authors to be significant. In contrast, the risk for colorectal cancer in the estrogen only group was judged to be insignificant. The authors conclude that the incidence of endometrial cancer in the E+P group was not affected since the final hazard ratio was 0.83 (95% CI, 0.29-2.32).

### Hip fracture

The incidence of hip fracture in the E+P and E-alone studies are low and the risk ratios are all below 1.0, and the authors indicate these are significant reductions.

### Global index

The authors state that the global index in the E+P study was significantly increased; however, the HR of 1.15 and an adjusted 95% CI of 0.95-1.39, indicates that it is not significant. This 15% increase is actually an increase of 19/10,000 or 0.19% in absolute risk. When the global index ratio is viewed on an annual basis four values are very near or below the no effect level of 1.0 and the other two are low (not shown), and it is difficult to see how such data could have any significance. In the E-only study the global index was 1.01 (CI: 0.89-1.14) and was judged by the authors to be insignificant.

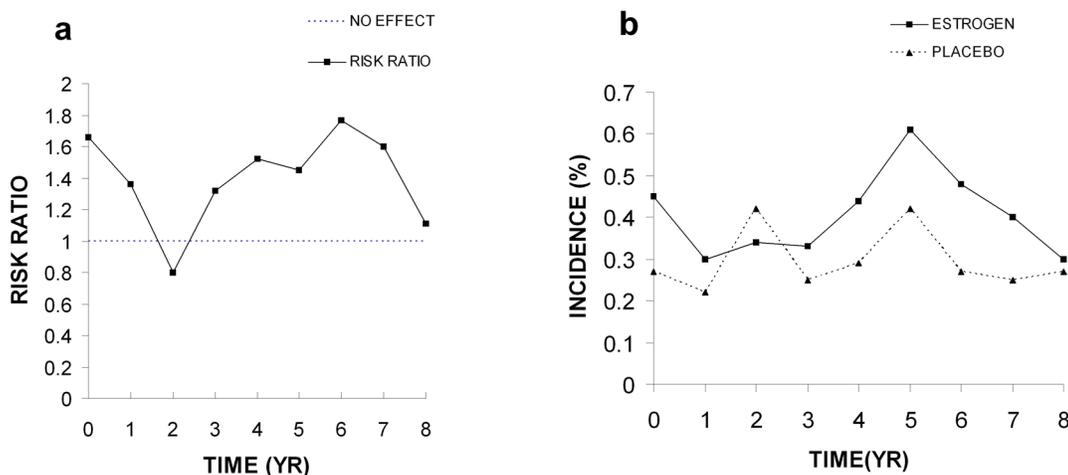
### Conclusions

The WHI study of E+P has received much acclaim and praise [Fletcher and Colditz, 2002; National Cancer Institute, 2002]. Although the methods are generally accepted as being appropriate for such clinical studies, I believe the conclusions drawn were incorrect as a result

of the use of unadjusted 95% CIs and failure to closely examine incidence data. It is difficult to believe that any elevated risks exist when graphs of risk ratios and disease incidence show no differences from control or from the no effect levels.

The 5 year peak in CHD, IBC and VTE may be due to factors other than increased risk from E+P. Possible confounders at the 5 year time point include: (A) decrease in event in the placebo group; and (B) a drop out rate that was six times greater than in the preceding year [Machens and Schmidt-Gollwitzer, 2003]. If such statistical artifacts and circumstances had been taken into account the overall evaluation of the WHI study would have turned out differently and the NHLBI would not have stopped the study prematurely. It has also been pointed out that the decision to terminate the E+P arm of the WHI study was based on unadjusted risk hazards for breast cancer and cardiovascular disease [Machens and Schmidt-Gollwitzer, 2003]. When confidence intervals are adjusted for multiple testing of the hazard ratios for these two diseases they are not statistically elevated.

In addition to my objections to the conclusions drawn by the estrogen plus progestin study, Dietel et al. [Dietel et al., 2005] have summarized the views of several investigators who have indicated that the E+P WHI study does not even qualify as a randomized placebo-controlled study [Braendle and Kuhl, 2003; Goodman et al., 2003; Machens and Schmidt-Gollwitzer, 2003; McDonough, 2002; Shapiro, 2003]. The reasons for this statement are: 1. After randomization the women were free to decide whether to continue their assigned treatment or whether to undergo diagnostic procedures. 2. The rate of unblinding in the E+P group was 45%, which means that almost half of the women were aware of their treatment. 3. Several warnings were sent to the participants about the detection of increased risks of myocardial infarction, stroke and pulmonary embolism during the study. These problems make the WHI study no better than any observational study with all of their limitations.



**Figure 9. Stroke in estrogen alone study.** A. Risk Ratios. All data for risk ratios were calculated from the data in Figure 3 of the WHI 2004 publication. B. Incidence of stroke in estrogen or placebo treated groups.

In addition to all of these problems, the women in these WHI studies were 12-15 years past the onset of menopause. Thus, these women were without their pre-menopausal levels of estrogen and progesterone long enough to bring about changes in various bodily functions which are the precursors of disease or of undiagnosed disease. For instance, ovarian hormones are important for maintaining normal structure and function of the blood vascular system [Karas and Clarkson, 2003]. Once vascular disease has begun, hormone treatment is not likely to reverse the effects [Naftolin et al., 2004]. Proper bone strength is maintained by estrogen and when estrogen is no longer secreted at menopause bones begin to lose calcium and the first stages of osteoporosis begin. Most reproductive scientists believe that post-menopausal hormones should be used as preventive, not corrective therapy; therefore, treatment should begin during the menopausal transition.

In summary, the findings of the E+P study should have been: no significant risks were found for cardiovascular disease, invasive breast cancer, stroke and venous thromboembolism. Instead, the WHI authors and the NHLBI concluded that post-menopausal hormone treatment increased the risks for all of these diseases. When this was released to the news media it resulted in considerable confusion among patients and doctors alike, and caused an untold number of women to go without potentially beneficial hormone therapy. The final consequences of these incorrect conclusions are yet to be determined, but it is likely that an untold number of women will suffer from diseases which post-menopausal hormone treatment could have prevented.

## References

Anderson, G. L., Limacher, M., Assaf, A. R., Bassford, T., Beresford, S. A., Black, H., Bonds, D., Brunner, R., Brzyski, R., Caan, B., Chlebowski, R., Curb, D., Gass, M., Hays, J., Heiss, G., Hendrix, S., Howard, B. V., Hsia, J., Hubbell, A., Jackson, R., Johnson, K. C., Judd, H., Kotchen, J. M., Kuller, L., LaCroix, A. Z., Lane, D., Langer, R. D., Lasser, N., Lewis, C. E., Manson, J., Margolis, K., Ockene, J., O'Sullivan, M. J., Phillips, L., Prentice, R. L., Ritenbaugh, C., Robbins, J., Rossouw, J. E., Sarto, G., Stefanick, M. L., Van Horn, L., Wactawski-Wende, J., Wallace, R. and

Wassertheil-Smoller, S. (2004) Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial *Jama* **291**, 1701-12.

Braendle, W and Kuhl, H (2003) Stellungnahme zur Millionen-Frauenstudie und Brustkrebs. *J Menopause* **3**, 3-4.

Chlebowski, R. T., Hendrix, S. L., Langer, R. D., Stefanick, M. L., Gass, M., Lane, D., Rodabough, R. J., Gilligan, M. A., Cyr, M. G., Thomson, C. A., Khandekar, J., Petrovitch, H. and McTiernan, A. (2003) Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial *Jama* **289**, 3243-53.

Cox, DR (1972) Regression analysis and life tables. *J R Stat. Soc.* **34**, 187-200.

Curb, J. D., Prentice, R. L., Bray, P. F., Langer, R. D., Van Horn, L., Barnabei, V. M., Bloch, M. J., Cyr, M. G., Gass, M., Lepine, L., Rodabough, R. J., Sidney, S., Uwaifo, G. I. and Rosendaal, F. R. (2006) Venous thrombosis and conjugated equine estrogen in women without a uterus *Arch Intern Med* **166**, 772-80.

Dietel, M., Lewis, M. A. and Shapiro, S. (2005) Hormone replacement therapy: pathobiological aspects of hormone-sensitive cancers in women relevant to epidemiological studies on HRT: a mini-review *Hum Reprod* **20**, 2052-60.

Fletcher, S. W. and Colditz, G. A. (2002) Failure of estrogen plus progestin therapy for prevention *Jama* **288**, 366-8.

Gann, P. H. and Morrow, M. (2003) Combined hormone therapy and breast cancer: a single-edged sword *Jama* **289**, 3304-6.

Goodman, N, Goldzieher, J and Ayala, C (2003) Critique of the report from the writing group of the WHI. *Menopausal Med* **10**, 1-4.

Grodstein, F., Manson, J. E. and Stampfer, M. J. (2006) Hormone therapy and coronary heart disease: the role of time since menopause and age at hormone initiation *J Womens Health (Larchmt)* **15**, 35-44.

Harman, S. M., Naftolin, F., Brinton, E. A. and Judelson, D. R. (2005) Is the estrogen controversy over? Deconstructing the Women's Health Initiative study: a critical evaluation of the evidence *Ann N Y Acad Sci* **1052**, 43-56.

Hsia, J., Langer, R. D., Manson, J. E., Kuller, L., Johnson, K. C., Hendrix, S. L., Pettinger, M., Heckbert, S. R., Greep, N., Crawford, S., Eaton, C. B., Kostis, J. B., Caralis, P. and Prentice, R. (2006) Conjugated equine estrogens and coronary heart disease: the Women's Health Initiative *Arch Intern Med* **166**, 357-65.

Karas, RH and Clarkson, TB (2003) Considerations in interoperating the cardiovascular effects of hormone replacement therapy observed in the WHI: timing is everything. *Menopausal Med* **10**, 8-10.

Kelly, P (2002) Recent HRT studies: the findings in perspective. *San Francisco Medicine Magazine*

Koscielny, S., Tubiana, M. and Valleron, A. J. (1985) A simulation model of the natural history of human breast cancer *Br J Cancer* **52**, 515-24.

Machens, K. and Schmidt-Gollwitzer, K. (2003) Issues to debate on the Women's Health Initiative (WHI) study. Hormone replacement therapy: an epidemiological dilemma? *Hum Reprod* **18**, 1992-9.

Manson, J. E., Hsia, J., Johnson, K. C., Rossouw, J. E., Assaf, A. R., Lasser, N. L., Trevisan, M., Black, H. R., Heckbert, S. R., Detrano, R., Strickland, O. L., Wong, N. D., Crouse, J. R., Stein, E. and Cushman, M. (2003) Estrogen plus progestin and the risk of coronary heart disease *N Engl J Med* **349**, 523-34.

McDonough, P. G. (2002) The randomized world is not without its imperfections: reflections on the Women's Health Initiative Study *Fertil Steril* **78**, 951-6.

Moher, D., Schulz, K. F. and Altman, D. (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials *Jama* **285**, 1987-91.

Naftolin, F., Taylor, H. S., Karas, R., Brinton, E., Newman, I., Clarkson, T. B., Mendelsohn, M., Lobo, R. A., Judelson, D. R., Nachtigall, L. E., Heward, C. B., Hecht, H., Jaff, M. R. and Harman, S. M. (2004) The Women's Health Initiative could not have detected cardioprotective effects of starting hormone therapy during the menopausal transition *Fertil Steril* **81**, 1498-501.

National Cancer Institute. (2002) Experts weigh in on hormone therapy. [www.cancer.gov](http://www.cancer.gov)

National Heart, Lung, and Blood Institute. (2002) NIH news release from the National Heart, Lung, and Blood Institute. <http://www.nhlbi.nih.gov/new/press/02/07/09.htm>

Pedersen, A. T. and Ottesen, B. (2003) Issues to debate on the Women's Health Initiative (WHI) study. Epidemiology or randomized clinical trials--time out for hormone replacement therapy studies? *Hum Reprod* **18**, 2241-4.

Rossouw, J. E., Anderson, G. L., Prentice, R. L., LaCroix, A. Z., Kooperberg, C., Stefanick, M. L., Jackson, R. D., Beresford, S. A., Howard, B. V., Johnson, K. C., Kotchen, J. M. and Ockene, J. (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial *Jama* **288**, 321-33.

Shapiro, S. (2003) Risks of estrogen plus progestin therapy: a sensitivity analysis of findings in the Women's Health Initiative randomized controlled trial *Climacteric* **6**, 302-10; discussion 310-3.

Stefanick, M. L., Anderson, G. L., Margolis, K. L., Hendrix, S. L., Rodabough, R. J., Paskett, E. D., Lane, D. S., Hubbell, F. A., Assaf, A. R., Sarto, G. E., Schenken, R. S., Yasmeen, S., Lessin, L. and Chlebowski, R. T. (2006) Effects of conjugated equine estrogens on breast cancer and mammography screening in postmenopausal women with hysterectomy *Jama* **295**, 1647-57.

Taubes, G. (1995) Epidemiology faces its limits *Science* **269**, 164-9.

Turgeon, J. L., McDonnell, D. P., Martin, K. A. and Wise, P. M. (2004) Hormone therapy: physiological complexity belies therapeutic simplicity *Science* **304**, 1269-73.

von Fournier, D., Weber, E., Hoeffken, W., Bauer, M., Kubli, F. and Barth, V. (1980) Growth rate of 147 mammary carcinomas *Cancer* **45**, 2198-2207.

Wassertheil-Smoller, S., Hendrix, S. L., Limacher, M., Heiss, G., Kooperberg, C., Baird, A., Kotchen, T., Curb, J. D., Black, H., Rossouw, J. E., Aragaki, A., Safford, M., Stein, E., Laowattana, S. and Mysiow, W.

J. (2003) Effect of estrogen plus progestin on stroke in postmenopausal women: the Women's Health Initiative: a randomized trial *Jama* **289**, 2673-84.