



Published in final edited form as:

Arthritis Rheum. 2008 March 15; 59(3): 302–310. doi:10.1002/art.23325.

Effects of Postmenopausal Hormone Therapy on Rheumatoid Arthritis: The Women's Health Initiative Randomized Controlled Trials

BRIAN WALITT, MD, MPH¹, MARY PETTINGER, MS², ARTHUR WEINSTEIN, MD¹, JAMES KATZ, MD³, JAMES TORNER, PhD⁴, MARY CHESTER WASKO, MD, MSc⁵, BARBARA V. HOWARD, PhD⁶, and for the WOMEN'S HEALTH INITIATIVE INVESTIGATORS

¹Washington Hospital Center, Washington, DC ²Fred Hutchinson Cancer Research Center, Seattle, Washington ³George Washington University, Washington, DC ⁴University of Iowa, Iowa City ⁵University of Pittsburgh, Pittsburgh, Pennsylvania ⁶MedStar Research Institute, Hyattsville, Maryland.

Abstract

Objective—To study the effects of postmenopausal hormone therapy (PHT) on the incidence and severity of rheumatoid arthritis (RA).

Methods—The Women's Health Initiative randomized controlled trials evaluated the effects of unopposed estrogen (E-alone) and estrogen plus progestin (E+P) compared with placebo on a diverse set of health outcomes over 7.1 and 5.6 years, respectively. RA cases were identified using historical and medication data. The hazard of developing RA was estimated using Cox proportional hazards regression models. Disease symptom severity was estimated using the Short Form 36 (SF-36) and self-reported joint pain scores at baseline and after 1 year. Mean changes in severity were compared using linear regression models.

Results—Of the 27,347 participants, 63 prevalent cases and 105 incident cases of RA were identified. A nonsignificant reduction in the risk of developing RA (hazard ratio 0.74; 95% confidence interval [95% CI] 0.51, 1.10) was noted with PHT use. PHT use led to improved SF-36 scores in unadjusted analyses (percent change 12.5%; 95% CI –24.45, –0.57) but not after adjustment for relevant covariates ($P = 0.33$). Nonsignificant improvements in joint pain scores were seen with PHT use (odds ratio [OR] 4.10; 95% CI 0.83, 20.20). PHT did not improve swelling (OR 1.27; 95% CI 0.08, 19.63) or prevent new joint pains (OR 0.72; 95% CI 0.11, 4.68) in RA participants.

Conclusion—There were no statistically significant differences in the risk of developing RA or the severity of RA between the PHT and placebo groups.

Address correspondence to Brian Walitt, MD, MPH, Washington Hospital Center, 110 Irving Street, Washington, DC 20010. E-mail: Brian.t.walitt@medstar.net.

AUTHOR CONTRIBUTIONS

Dr. Walitt had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study design. Walitt, Weinstein, Katz, Torner, Wasko, Howard.

Acquisition of data. Walitt, Torner.

Analysis and interpretation of data. Walitt, Pettinger, Weinstein, Katz, Wasko.

Manuscript preparation. Walitt, Pettinger, Weinstein, Katz, Torner, Wasko, Howard.

Statistical analysis. Walitt, Pettinger.

ClinicalTrials.gov identifier: NCT00000611.

Dr. Howard received consultant fees (less than \$10,000 each) from Merck and Egg Nutrition Council and speaking fees (less than \$10,000) from Schering-Plough.

INTRODUCTION

Although rheumatoid arthritis (RA) is recognized as a female-predominant disease, the reasons for sex discrepancy in disease susceptibility and expression are unclear (1). Studies have demonstrated that antibody response to antigenic stimulation is greater in women than in men (2). Estradiol levels decrease lymphocyte counts (3), increase mature helper thymocytes (4), modulate proinflammatory cytokine release through the modulation of CD16 expression (5), and deplete suppressor and cytotoxic thymocytes (6). Disease severity is also influenced by both pregnancy (7) and the menstrual cycle (8). This evidence suggests that sex hormones influence the pathogenesis or clinical expression of RA (9).

The effect of exogenous estrogen on RA has been the subject of many conflicting reports. Early case-control studies have demonstrated that exogenous estrogen use exerts a protective effect against the development of RA (10-12). These findings are disputed by later case-control studies (13-17) and cohort studies (18-20). The reason for this disparity is not clear and is mostly attributed to methodologic differences. Findings on the effects of estrogen on RA symptom severity have also been conflicting. Studies have demonstrated disease-modifying effects (21,22), marginal effects on RA severity (23,24), and no effect (25). It is still unclear whether exogenous estrogen has an effect on RA incidence or symptom severity.

The Women's Health Initiative (WHI) clinical trials and observational study, initiated in 1993, were designed to study the major causes of morbidity and mortality in postmenopausal women. The WHI clinical trials included trials of the effects of postmenopausal hormone therapy (PHT) on coronary heart disease, hip fracture, and breast cancer, using separate trials of estrogen plus progestin and estrogen alone (26). The WHI estrogen plus progestin (E+P) trial was stopped prematurely in July 2002 because the overall health risks of E+P exceeded the benefits in otherwise healthy women (27). In March 2004, the National Institutes of Health also ended the intervention phase of the estrogen alone (E-alone) trial for similar reasons (28).

Although the WHI was not designed to address the effects of PHT on the incidence or severity of RA, it did collect self-reported diagnostic information, medication data, and serial measures of general well-being and musculoskeletal symptoms on all participating women. With these data, it is possible to perform the first post hoc analysis of a prospective, double-blind, placebo-controlled trial of the effects of PHT therapy on the incidence and severity of RA. Even though the interpretation of these results cannot be generalized easily to groups other than postmenopausal women, the WHI offers a means to evaluate the effects of estrogens on RA with a blinded, controlled design. The primary goal of the present analysis was to determine if PHT has a significant effect on the initial occurrence of RA. Our secondary goal was to determine if PHT has significant effects on self-reported outcome measures in patients with RA.

PATIENTS AND METHODS

Setting and patients

Postmenopausal women 50–79 years of age were recruited from 40 clinical centers across the US to participate in the WHI (29). In brief, women were considered postmenopausal if they had no vaginal bleeding for the previous 6 months (12 months for women age 50–54) or had ever received PHT. Women were excluded if they had medical conditions associated with competing risk, safety concerns, or adherence and retention concerns. Relevant to this analysis is that women receiving oral corticosteroids at screening were excluded from the study. Subsequent corticosteroid therapy was recorded and did not exclude women from ongoing participation.

Randomization and interventions

The randomized, controlled hormone trials (CT) were designed to evaluate the effects of PHT on specific disease outcomes (26). The E+P trial randomized 16,608 women with an intact uterus to receive either 0.625 mg of conjugated equine estrogens (CEE) plus 2.5 mg of medroxyprogesterone or placebo and prospectively monitored the women for an average of 5.6 years. The E-alone trial randomized 10,739 women without an intact uterus to receive either 0.625 mg of CEE or placebo and prospectively monitored the women for an average of 7.1 years. The WHI was performed in compliance with the Declaration of Helsinki. Further details of the CT trial design, recruitment, screening, randomization, and eligibility criteria are described elsewhere (26,27,29,30).

Assessments and followup

Routine followup of all women in the CT consisted of contacting participants semiannually to collect self-reported outcomes and assess participant safety and more extensive annual visits to perform physical measurements, gynecologic examinations, and to obtain mammogram reports. Self-reported diagnostic information was collected annually about the presence or absence of arthritis, RA, and osteoarthritis/other arthritis. Information on medication use was collected at baseline and at study years 1, 3, 6, and 9 by trained research staff. Women were instructed to bring all of the medication they were currently taking, including over-the-counter and herbal medications. The Short Form 36 (SF-36), which has been used as a surrogate measurement of disease severity in RA (31), was collected at baseline and at 1 year after randomization. The SF-36 consists of 2 component scores, the Physical Component Score (PCS) and the Mental Component Score (MCS). The PCS score is more sensitive than the MCS in arthritic conditions (31). Measures of self-reported joint pain/stiffness and swelling in the hands/feet were also collected at baseline and 1 year after randomization. Participants were asked to rate their symptom severity on a 0–3 scale, with 0 signifying that the symptom was not present and 3 indicating severe symptoms.

In the CT, 2,541 participants reported RA at some point over the course of the study, which grossly overestimated the prevalence density that would be expected in the general population. To adjust for presumed overreporting, a participant was considered to have RA if 1) the participant reported having RA and 2) the participant was taking antirheumatic medications at the time of the self-report. Antirheumatic medications considered in the definition included hydroxychloroquine, sulfasalazine, minocycline, methotrexate, leflunomide, cyclosporine, azathioprine, gold compounds, cyclophosphamide, anti-tumor necrosis factor medications, and oral steroids. The use of nonsteroidal antiinflammatory drugs was not considered in the case definition. No patients received D-penicillamine, tacrolimus, or anakinra during the entire study. In an RA disease validation study in the WHI, coupling the self-reported diagnosis with medications was shown to have a positive predictive value of 62.2%, a negative predictive value of 93.9%, and a kappa of 0.53 (32). The definition was more accurate with the inclusion of oral steroids than without.

Statistical analysis

All statistical analyses were performed based on the intent-to-treat principle using SAS software, version 9.1 (SAS Institute, Cary, NC). The primary analysis focused on determining the effects of PHT on the initial development of RA using failure-time methods. The number of days from randomization to the first self-report of diagnosis was used as the time of event. Followup time for participants without an event was censored at the date each of the 2 trials was stopped, the date of last contact, or death, whichever occurred first. After excluding prevalent cases, comparisons between treatment groups were reported as hazard ratios (HRs) and 95% confidence intervals (95% CIs) estimated from Cox proportional hazards regression models; stratified by hormone trial, age group at screening, and randomization assignment in

the diet modification trial; and adjusted for smoking history (current and past) and prior PHT and oral contraceptive pill use. Separate models were also run by trial, race/ethnicity, and age decade. Tests for interaction of these factors with treatment assignment were performed in Cox models that included main effects for the factor and treatment assignment, and their interaction. The Kaplan-Meier method was used to estimate and plot cumulative hazard rates.

Secondary analyses focused on determining the effects of PHT on overall well-being and musculoskeletal symptom severity in women with RA. The primary end point of the secondary analysis was mean change in PCS score from baseline between the PHT and placebo groups at 1 year. Data were collected after 1 year for only a small subset of participants, precluding our ability to look at further longitudinal outcomes. Secondary end points also included mean change in MCS scores, joint pain assessment, and hand or foot swelling assessment. RA was evaluated in 2 different ways for these analyses. In the first set of analyses, only prevalent cases of RA were considered. A second set of analyses were performed that included both prevalent cases and incident cases of RA at 1 year. No substantial differences in the results were noted between the 2 approaches, thus our results included both prevalent and 1-year incident cases.

To estimate the mean change in SF-36 scores among prevalent RA cases and make comparisons between groups, linear regression models were performed with the change in score as the dependent variable and treatment assignment as the independent variable. Analyses were performed both unadjusted and adjusted for hormone trial, age, comorbidity as measured by the Charlson Index (33), years since menopause, total years of menstruation, current smoking, smoking history, prior use of hormone therapy, prior use of oral contraceptives, antirheumatic medication potency, and current use of oral steroids. Antirheumatic potency was categorized into mild agents (hydroxychloroquine, sulfasalazine, minocycline) and potent agents (methotrexate, leflunomide, cyclosporine, azathioprine, gold compounds, cyclophosphamide, and anti-tumor necrosis factor medications). Additional models also analyzed the E-alone and E+P trials separately. To determine if the effect of PHT on the change in scores differed between participants with and without prevalent RA, interaction terms were included in models run on all women.

The analysis of joint pain and hand or foot swelling split the participants into symptomatic and asymptomatic groups based on their baseline scores. Joint pain and hand or foot swelling were analyzed independently. The self-reporting of symptoms as moderate or severe (a score of 2 or 3, respectively) was defined as symptomatic, whereas the reporting of mild or none (a score of 0 or 1, respectively) was defined as asymptomatic. Logistic regression models were used to estimate the odds of relief/improvement or onset at year 1 resulting from treatment with PHT. Analyses were run separately for participants with and without prevalent RA.

The aforementioned statistical methods were all post hoc analyses of data from a randomized clinical trial that was not designed to determine the effects of PHT on RA. For this reason, power analyses were not performed because their utility is in study design and would not provide any new information in this study setting (34).

RESULTS

Of the 16,608 participants with an intact uterus, 8,506 were randomly assigned to the E+P group and 8,102 were assigned to the placebo group. Of the 10,739 participants without an intact uterus, 5,310 were randomly assigned to the E-alone group and 5,429 were assigned to the placebo group. Baseline demographic information for the women in the trials has been previously published (30). There were differences in the demographics between the E-alone group and the E+P group. Women in the E-alone group were more often African American (15.06% versus 6.77%), had less education and total family income, had a higher body mass

index (30.1 kg/m² versus 28.5 kg/m²), had fewer years of menstruation (29.6 years versus 38.0 years), were more likely to have previously used PHT (48.4% versus 25.93%), and had lower PCS scores (46.5 versus 49.4) compared with women in the E+P group. The E-alone and E+P groups were comparable on other demographic variables, including substance use, joint surgeries, and gynecologic variables.

There were 105 incident cases and 63 prevalent cases of RA. Seventeen of the incident cases occurred by year 1. Age and racial composition of the women with RA were similar to the study population (Table 1). Women with RA had higher comorbidity scores and worse PCS, joint pain, and hand/foot swelling scores compared with the study population. Disease and nondisease groups were similar in terms of substance use (coffee consumption, smoking, and alcohol use), gynecologic variables (live births, induced abortions, breastfeeding, oral contraceptive use, and benign breast disease), joint surgeries, and length of followup.

There was an average of 5.6 years and 7.1 years of followup of the E+P and E-alone trials, respectively. There were no statistically significant differences in incident cases of RA between the PHT and placebo groups (Table 2). There was a nonsignificant reduction in incident cases of RA that persisted after adjusting for multiple covariates (HR 0.74; 95% CI 0.51, 1.10). E-alone (HR 0.69; 95% CI 0.41, 1.14) appeared to contribute more to this reduction than E+P (HR 0.83; 95% CI 0.46, 1.51) but no statistical difference was noted (interaction $P = 0.65$). Graphically, this effect appeared to increase over time (Figure 1); however, no significant interaction effects with time were noted (interaction P values were 0.26 for combined trials, 0.67 for the E-alone trial, and 0.27 for the E+P trial).

The effects of PHT use on SF-36 scores are shown in Table 3. Participants with RA who were receiving PHT had a 7% improvement in their scores whereas those receiving placebo had a 5.5% deterioration in their scores. This 12.5% difference in PCS scores between the 2 treatment arms was statistically significant (95% CI -24.45, -0.57; $P = 0.04$). However, the difference diminished after adjusting for age, comorbidity, antirheumatic medication potency, steroid use, current smoking, years since menopause, and total years of menstruation ($P = 0.33$). E-alone demonstrated a significant protective effect on PCS scores (change of 22.36%; 95% CI -38.56, -6.16; $P = 0.009$) that persisted as a statistical trend after making relevant adjustments ($P = 0.055$). The E+P arm revealed no significant difference between the groups ($P = 0.83$). No differences in mental well-being, as measured by the MCS, were noted between the PHT and placebo groups (0.39% change; 95% CI -12.71, 13.49). PHT had a greater effect on PCS scores in the RA cases when compared with noncases even after adjustment for covariates (12.5% versus 1.0%; interaction $P = 0.10$).

The effects of PHT use on joint pain scores are shown in Table 4. Participants with RA had more relief of joint pain with PHT use (odds ratio [OR] 4.10; 95% CI 0.83, 20.20), which was not statistically significant. No protective effects were noted among asymptomatic women with RA with PHT use at 1 year (OR 0.72; 95% CI 0.11, 4.68). PHT use did not lead to significant improvements in swelling of the hands and feet in participants with RA (OR 1.27; 95% CI 0.08, 19.63). PHT did not prevent new swelling in the asymptomatic RA population (OR 0.42; 95% CI 0.11, 1.57). Because of the small numbers of cases, we were unable to perform analyses of swelling stratified by type of PHT therapy.

DISCUSSION

The effects of estrogen on the development and severity of RA remains controversial despite multiple studies. Our study is the only placebo-controlled trial to evaluate the effect of PHT on the initial development of RA and the sixth to evaluate the effects of PHT on perceived disease severity. Methodologic issues such as self-selection bias (15), selection biases due to

the use of site-specific cohorts (17), and the inherent inability of case-control studies to directly ascertain incidence rates could be addressed for the first time in a double-blind, randomized trial in a nationwide cohort. We were unable to demonstrate any statistically significant effects on either RA incident cases over an average of 5–6 years or RA symptom severity after 1 year.

Although we were unable to demonstrate statistically significant changes in the number of incident RA cases, our results do demonstrate trends suggesting that PHT may protect against developing RA. In our Kaplan-Meier plots, we note that there is separation between the placebo and the combined PHT arms at 2 years and what appears to be divergence at 4 years. This trend is more apparent with E-alone use. However, no significant interaction effects were noted, making it difficult to speculate that these trends become more prominent over time. It is possible that if more cases were noted after year 4 or if the study had a longer followup, a statistical difference would have been seen; however, our data are unable to support that PHT therapy has any effect on RA initiation.

Our results also do not demonstrate that PHT use leads to statistically significant differences in RA severity. However, participants with RA taking PHT appeared to have improved SF-36 and joint pain scores after 1 year of PHT use. These instruments seem to measure the same aspects of health because they were found to be highly correlated ($P < 0.0001$) (data not shown). Considering that a clinically significant change on the PCS scale is 4.4 points (35), clinical improvement was typically small (0.77 points) but important when compared with the degradation seen in the placebo group (–3.65 points). A significant effect might have been seen with a larger sample size, longer intervention period, or more specific outcome measures. Although the SF-36 and visual analog scales are well-validated tools and have been used in RA (30), they do not have the disease specificity of the American College of Rheumatology criteria for 20% improvement in disease activity, the Disease Activity Score in 28 joints, the Health Assessment Questionnaire, or measures of radiographic progression in measuring response to treatment. Such measures may have been more effective at capturing the nuisances of RA that are missed with the general measures we used. In contrast, the effect may simply reflect the small improvement in joint pain/stiffness with PHT use reported previously in women without RA (36). Regardless, our data are unable to support that 1 year of PHT therapy has a clinically significant effect on RA symptom severity.

Our study represents the eighth to investigate the relationship between PHT and the risk of developing RA and the sixth to investigate the relationship between PHT and RA severity (Table 5). The initial study of PHT effects on incident RA cases reported a protective effect (12), with the early followup studies finding trends toward a protective effect in past PHT users (14,18). Later studies reported that PHT had no effect on incident RA cases (15–17,19,20). Unlike our post hoc analysis of an interventional study, all of these aforementioned studies were observational and therefore subject to unknown confounding. Our study supports the more current literature demonstrating no significant protective benefit from PHT use in preventing RA.

All but one of the trials investigating the relationship between PHT use and RA severity reported salutary effects of PHT use (21–25). Those studies were blinded, placebo controlled, and of short duration. However, they used various RA severity measures and some of the studies had larger sample sizes than were available in the WHI. Our findings do not support the previous literature findings of a beneficial clinical effect of PHT use on RA symptoms. While our case numbers and followup are not as large as some of the prior studies, our study suggests that any true clinical effect not captured by our statistical testing is likely to be mild.

The prevalence of RA in our study was approximately half of the 0.5–1% that is expected in the general population (37,38). This finding may be related to the exclusion of any potential

participants taking prednisone at the time of screening, the tendency for clinical trials to recruit healthier patients, and the possibility that potential prevalent RA participants may have already been taking PHT for osteoporosis prevention. Prior studies have estimated the incidence of RA to be between 25 and 50 cases per 100,000 women between the ages of 30 and 50 (39,40). When calculated in our study, the RA incidence proportion was 57 cases per 100,000 women. This is arguably a small overestimate of the incidence rate expected to be seen in older women. However, any ascertainment bias related to the disease definition should be equal in both the placebo and PHT arms due to the randomized nature of the study. The RA incidence also suggests that our case definition is not the cause for the low prevalence rate noted in this study.

Although the methodology of the WHI has many advantages over prior studies, its small number of cases and reliance on the self-reporting of disease states are important limitations. Despite a relatively large number of enrolled patients, the sample size was much less than that which is required to observe an effect size of PHT on the rare event of RA initiation. Thus, this study was not powered for this end point. However, it is unlikely that there will be more statistically powerful prospective trials of PHT due to the health risks of treatment. The baseline RA sample size was also not as robust as would be ideal. It was large enough to confidently capture large effects on symptoms but was too small to rule out the possibility of small but significant effects. Our reliance on self-reported information for our case definition offers its own problems. By defining cases as women who both reported RA and were concomitantly taking appropriate medications, we tried to minimize overreporting and response biases. However, this approach is not as sensitive as performing chart reviews or physical examinations. Some of the diagnostic inaccuracies inherent in the diagnosis are offset by the randomization of the trials, but inaccuracies in our case definition need to be considered when interpreting our data.

In summary, the design of the WHI provided a unique opportunity to examine the effects of PHT on RA. Despite the participation of 27,347 women, there was no statistically significant evidence of a difference in the hazard of RA incidence or a difference in RA symptom severity between the PHT and placebo groups.

Acknowledgements

Supported by the American College of Rheumatology Clinical Investigator Fellowship award. The study drugs used in the Women's Health Initiative were provided by Pfizer, Merck, and Schering-Plough.

REFERENCES

1. Firestein, G. Kelly's textbook of rheumatology. Vol. 6th ed.. Saunders; Richmond (VA): 2001. Etiology and pathogenesis of rheumatoid arthritis; p. 924
2. Ansar Ahmed S, Penhale WJ, Talal N. Sex hormones, immune responses and autoimmune diseases: mechanisms of sex hormone action. *Am J Pathol* 1985;121:531–51. [PubMed: 3907369]
3. Mathur S, Mathur RS, Goust JM, Williamson HO, Fudenberg HH. Cyclic variations in white cell subpopulations in the human menstrual cycle: correlations with progesterone and estradiol. *Clin Immunol Immunopathol* 1979;13:246–53. [PubMed: 313307]
4. Novotny EA, Raveche ES, Sharrow S, Ottinger M, Steinberg AD. Analysis of thymocyte subpopulations following treatment with sex hormones. *Clin Immunol Immunopathol* 1983;28:205–17. [PubMed: 6603317]
5. Kramer PR, Kramer SF, Guan G. 17β -estradiol regulates cytokine release through modulation of CD16 expression in monocytes and monocyte-derived macrophages. *Arthritis Rheum* 2004;50:1967–75. [PubMed: 15188374]
6. Ansar Ahmed S, Talal N. Sex hormones and autoimmune rheumatic disorders. *Scand J Rheumatol* 1989;18:69–76. [PubMed: 2786646]

7. Persellin RH. The effect of pregnancy on rheumatoid arthritis. *Bull Rheum Dis* 1977;27:922–7. [PubMed: 799947]
8. Rudge SR, Kowanko IC, Drury PL. Menstrual cyclicality of finger joint size and grip strength in patients with rheumatoid arthritis. *Ann Rheum Dis* 1983;42:425–30. [PubMed: 6882039]
9. Cutolo M, Lahita R. Estrogens and arthritis. *Rheum Dis Clin North Am* 2005;31:19–27. [PubMed: 15639053]
10. Wingrave SJ, Kay CR. Reduction in incidence of rheumatoid arthritis associated with oral contraception. *Lancet* 1978;1:569–71. [PubMed: 76118]
11. Vandenbroucke JP, Valkenburg HA, Boersma JW, Cats A, Festen JJ, Huber-Bruning O, et al. Oral contraceptives and rheumatoid arthritis: further evidence for a preventive effect. *Lancet* 1982;2:839–42. [PubMed: 6126710]
12. Vandenbroucke JP, Witteman J, Valkenburg H, Boersma JW, Cats A, Festen JJ, et al. Noncontraceptive hormones and rheumatoid arthritis in perimenopausal and postmenopausal women. *JAMA* 1986;255:1299–303. [PubMed: 3944948]
13. Linos A, Worthington JW, O'Fallon WM, Kurland LT. Case-control study of rheumatoid arthritis and prior use of oral contraceptives. *Lancet* 1983;1:1299–300. [PubMed: 6134094]
14. Carette S, Marcoux S, Gingras S. Postmenopausal hormones and the incidence of rheumatoid arthritis. *J Rheumatol* 1989;16:911–3. [PubMed: 2769663]
15. Spector TD, Brennan P, Harris P, Studd JW, Silman AJ. Does estrogen replacement therapy protect against rheumatoid arthritis? *J Rheumatol* 1991;18:1473–6. [PubMed: 1765970]
16. Koepsell TD, Dugowson CE, Nelson JL, Voigt LF, Daling JR. Non-contraceptive hormones and the risk of rheumatoid arthritis in menopausal women. *Int J Epidemiol* 1994;23:1248–55. [PubMed: 7721528]
17. Doran MF, Crowson CS, O'Fallon WM, Gabriel SE. The effect of oral contraceptives and estrogen replacement therapy on the risk of rheumatoid arthritis: a population based study. *J Rheumatol* 2004;31:207–13. [PubMed: 14760786]
18. Hernandez-Avila M, Liang MH, Willet WC, Stampfer MJ, Colditz GA, Rosner B, et al. Exogenous sex hormones and the risk of rheumatoid arthritis. *Arthritis Rheum* 1990;33:947–53. [PubMed: 2369431]
19. Merlino LA, Cerhan JR, Criswell LA, Mikuls TR, Saag KG. Estrogen and other female reproductive risk factors are not strongly associated with the development of rheumatoid arthritis in elderly women. *Semin Arthritis Rheum* 2003;33:72–82. [PubMed: 14625816]
20. Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. *Arthritis Rheum* 2004;50:3458–67. [PubMed: 15529351]
21. Bijlsma JW, Huber-Bruning O, Thijssen JH. Effect of oestrogen treatment on clinical and laboratory manifestations of rheumatoid arthritis. *Ann Rheum Dis* 1987;46:777–9. [PubMed: 3318724]
22. D'Elia HF, Larsen A, Mattsson LA, Waltbrand E, Kvist G, Mellstrom D, et al. Influence of hormone replacement therapy on disease progression and bone mineral density in rheumatoid arthritis. *J Rheumatol* 2003;30:1456–63. [PubMed: 12858441]
23. Hall GM, Daniels M, Huskisson EC, Spector TD. A randomized controlled trial of the effect of hormone replacement therapy on disease activity in postmenopausal rheumatoid arthritis. *Ann Rheum Dis* 1994;53:112–6. [PubMed: 8129455]
24. MacDonald A, Murphy E, Capell H, Bankowska U, Ralston S. Effects of hormone replacement therapy in rheumatoid arthritis: a double blind placebo-controlled study. *Ann Rheum Dis* 1994;53:54–7. [PubMed: 8311557]
25. Van den Brink HR, van Everdingen AA, van Wijk MJ, Jacobs JW, Bijlsma JW. Adjuvant oestrogen therapy does not improve disease activity in postmenopausal patients with rheumatoid arthritis. *Ann Rheum Dis* 1993;52:862–5. [PubMed: 8311536]
26. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* 1998;19:61–109. [PubMed: 9492970]
27. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. the Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin

- in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321–33. [PubMed: 12117397]
28. Anderson GL, Limacher M, Assaf AR, Bassford T, Beresford SA, Black H, et al. the Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in post-menopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA* 2004;291:1701–12. [PubMed: 15082697]
 29. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 2003;13(9 Suppl):S18–77. [PubMed: 14575939]
 30. Stefanick ML, Cochrane BB, Hsia J, Barad DH, Liu JH, Johnson SR. The Women's Health Initiative postmenopausal hormone trials: overview and baseline characteristics of participants. *Ann Epidemiol* 2003;13(9 Suppl):S78–86. [PubMed: 14575940]
 31. Wiles NJ, Scott DG, Barrett EM, Merry P, Arie E, Gaffney K, et al. Benchmarking: the five year outcome of rheumatoid arthritis assessed using a pain score, the Health Assessment Questionnaire, and the Short Form-36 (SF-36) in a community and a clinic based sample. *Ann Rheum Dis* 2001;60:956–61. [PubMed: 11557653]
 32. Walitt B, Constantinescu F, Katz J, Weinstein A, Wang H, Khorana RG, et al. Validation of the self-report of rheumatoid arthritis and systemic lupus erythematosus: the Women's Health Initiative. *J Rheumatol*. In press
 33. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83. [PubMed: 3558716]
 34. Thomas L. Retrospective power analysis. *Conserv Biol* 1997;11:276–80.
 35. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 2000;43:1478–87. [PubMed: 10902749]
 36. Hays J, Ockene JK, Brunner RL, Kotchen JM, Manson JE, Patterson RE, et al. the Women's Health Initiative Investigators. Effects of estrogen plus progestin on health-related quality of life. *N Engl J Med* 2003;348:1839–54. [PubMed: 12642637]
 37. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini EH, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998;41:778–99. [PubMed: 9588729]
 38. Kvien TK, Glennas A, Knudrod OG, Smedstad LM, Mowinckel P, Forre O. The prevalence and severity of rheumatoid arthritis in Oslo: results from a county register and population survey. *Scand J Rheumatol* 1997;26:412–8. [PubMed: 9433400]
 39. Drosos AA, Alamanos I, Voulgari PV, Psychos DN, Katsaraki A, Papadopoulos I, et al. Epidemiology of adult rheumatoid arthritis in northwest Greece 1987–1995. *J Rheumatol* 1997;24:2129–33. [PubMed: 9375871]
 40. Uhlig T, Kvien TK, Glennas A, Smedstad LM, Forre O. The incidence and severity of rheumatoid arthritis, results from a county register in Oslo, Norway. *J Rheumatol* 1998;25:1078–84. [PubMed: 9632067]

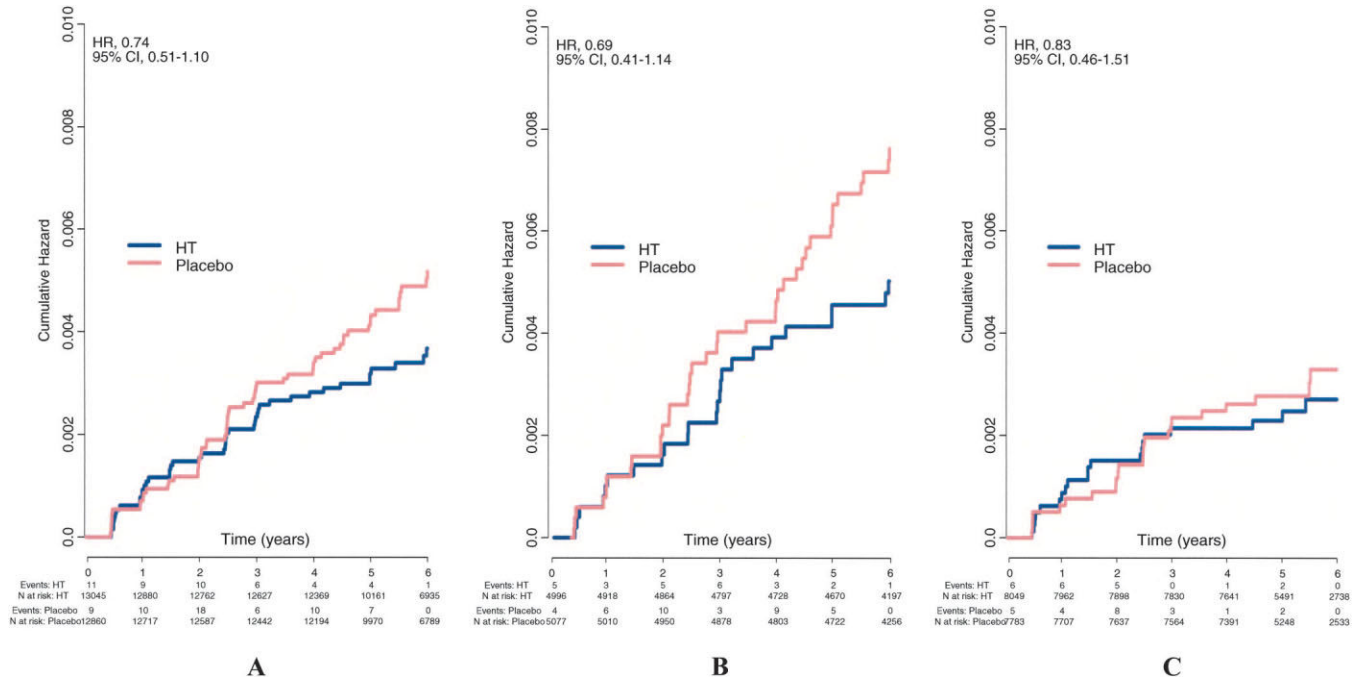


Figure 1. Cumulative hazard of developing rheumatoid arthritis by treatment arm. **A**, All participants. **B**, Participants receiving estrogen alone. **C**, Participants receiving estrogen plus progestin. HR = hazard ratio; 95% CI = 95% confidence interval; HT = hormone therapy.

Table 1

Study demographics*

Variable	Prevalent RA (n = 63)	Incident RA (n = 105)	All other women (n = 27,161)
Age, years	64.1 ± 7.5	63.6 ± 7.4	63.4 ± 7.2
Race/ethnicity, %			
White	84.1	78.1	80.6
African American	7.9	11.4	10.0
Hispanic	4.8	8.6	5.6
Education, %			
Less than HS diploma	4.8	6.7	8.0
HS diploma	25.4	22.1	20.8
Post-HS	49.2	46.2	40.6
College degree or higher	20.6	25.0	30.6
Income, %			
<\$20,000	22.6	24.0	23.5
\$20,000–<\$35,000	27.4	26.0	28.4
\$35,000–<\$50,000	16.1	30.0	20.4
\$50,000–<\$75,000	21.0	12.0	16.4
≥\$75,000	12.9	8.0	11.4
Smoking status, %			
Never smoked	57.1	50.0	50.3
Past smoker	34.9	38.5	39.2
Current smoker	7.9	11.5	10.5
Comorbidity (Charlson score), %			
0	64.5	59.6	68.3
1	21.0	26.3	19.0
≥2	14.5	14.1	12.7
Years of menstruation	34.5 ± 6.1	33.6 ± 7.4	34.8 ± 7.1
Years of menopause	17.4 ± 10.4	15.8 ± 10.0	15.6 ± 9.3
Baseline PCS score	39.6 ± 12.1	44.9 ± 9.5	48.3 ± 9.6
Baseline MCS score	54.0 ± 10.4	52.8 ± 9.5	53.2 ± 8.3
Baseline joint pain score, %			
No symptoms	1.6	6.9	27.3
Mild	28.6	47.5	46.8
Moderate	44.4	31.4	20.0
Severe	25.4	13.9	6.0
Baseline swelling score, %			
No symptoms	42.9	53.9	66.4
Mild	30.2	31.4	25.8
Moderate	22.2	12.7	6.4
Severe	4.8	2.0	1.3
Years of followup	5.7 ± 1.5	6.7 ± 1.2	6.2 ± 1.6

* Values are the mean ± SD unless otherwise indicated. RA = rheumatoid arthritis; HS = high school; PCS = Physical Component Score; MCS = Mental Component Score.

Table 2

Incidence of rheumatoid arthritis in the controlled trial*

	Incident cases randomized to PHT	Incident cases randomized to placebo	HR	95% CI	P
E-alone study arm	25	37	0.69	0.41, 1.14	0.149
E+P study arm	20	23	0.83	0.46, 1.51	0.543
Combined results	45	60	0.74	0.51, 1.10	0.134

* PHT = postmenopausal hormone therapy; HR = hazard ratio; 95% CI = 95% confidence interval; E-alone = estrogen alone; E+P = estrogen plus progestin.

Table 3
 Comparison of the effects of PHT versus placebo on PCS scores over 1 year in prevalent and incident cases of rheumatoid arthritis*

	PCS with PHT			PCS with placebo			Difference		
	N	Mean ± SD		N	Mean ± SD		Mean ± SE	P (unadjusted)	P (adjusted)
All PHT									
Baseline	38	40.92 ± 11.72		40	39.02 ± 11.79		-1.90 ± 2.66	0.478	0.089
Year 1	37	41.92 ± 10.82		36	36.64 ± 11.27		-5.28 ± 2.59	0.045	0.005
Year 1: baseline (absolute change)	37	0.77 ± 10.25		35	-2.89 ± 6.38		-3.65 ± 2.03	0.076	0.488
Year 1: baseline (% change)	37	7.00 ± 30.46		35	-5.51 ± 18.54		-12.51 ± 5.98	0.040	0.353
Estrogen alone									
Baseline	11	42.43 ± 11.80		16	35.79 ± 11.20		-6.64 ± 4.48	0.151	0.552
Year 1	10	44.07 ± 10.37		14	30.38 ± 11.24		-13.69 ± 4.51	0.006	0.105
Year 1: baseline (absolute change)	10	0.64 ± 7.43		14	-6.01 ± 5.66		-6.64 ± 2.67	0.021	0.086
Year 1: baseline (% change)	10	5.66 ± 23.52		14	-16.70 ± 14.82		-22.36 ± 7.81	0.009	0.055

* PHT = postmenopausal hormone therapy; PCS = Physical Component Score.

Table 4
Effects of PHT on joint pain/stiffness at year 1 in women with rheumatoid arthritis*

Effect at year 1	PHT	Placebo	OR	95% CI	P
No. asymptomatic at baseline (% with relief)					
Relief of joint pain/stiffness (n = 52)	23 (30.4)	29 (10.3)	4.10	0.83, 20.20	0.08
Relief of swelling of hands or feet (n = 20)	11 (81.8)	9 (55.6)	1.27	0.08, 19.63	0.86
No. asymptomatic at baseline (% with onset)					
Onset of joint pain/stiffness (n = 23)	14 (28.6)	9 (44.4)	0.72	0.11, 4.68	0.73
Onset of swelling of hands or feet (n = 56)	26 (15.4)	30 (30.0)	0.42	0.11, 1.57	0.20

* Includes both prevalent and incident cases of rheumatoid arthritis at year 1 (n = 80). Five cases had no change in joint pain/stiffness scores, 4 cases had no change in swelling scores. PHT = postmenopausal hormone therapy; OR = odds ratio; 95% CI = 95% confidence interval.

Table 5

Prior studies investigating the relationship between PHT and RA incidence and severity*

Author (ref.), year	Trial design	PHT stratification	Results
RA incidence			
Vandenbroucke (12), 1986	Case-control	Substitution hormones/substitution estrogens	Protective: OR 0.32 (95% CI 0.16, 0.64)
Carette (14), 1989	Case-control	Never users/past users/current users	Protective: past users OR 0.95 (95% CI 0.56, 1.60) Protective: current users OR 0.89 (95% CI 0.49, 1.63)
Hernandez-Avila (18), 1990	Prospective cohort	Never users/past users/current users	Protective: past users OR 0.7 (95% CI 0.5, 1.2) Not protective: current users OR 1.3 (95% CI 0.9, 2.0)
Spector (15), 1991	Case-control	ERT per person-years	Not protective: OR 1.08 (95% CI 0.30, 6.75)
Koepsell (16), 1994	Case-control	E-alone/E+P	Not protective: E-alone RR 0.97 (95% CI 0.62, 1.53) Not protective: E+P RR 0.81 (95% CI 0.45, 1.45)
Merlino (19), 2003	Cohort	Never users/past users/current users	Not protective: past users RR 1.47 (95% CI 1.04, 2.06) Not protective: current users RR 1.02 (95% CI 0.61, 1.72)
Doran (17), 2004	Case-control	Never users/any users/current users	Not protective: any users OR 1.26 (95% CI 0.81, 1.96) Not protective: current users OR 1.67 (95% CI 0.81, 3.41)
Karlson (20), 2004	Prospective cohort	Never users/past users/current users	Not protective: past users RR 1.3 (95% CI 1.0, 1.6) Not protective: current users RR 1.0 (95% CI 0.8, 1.3)
RA severity			
Bijlsma (21), 1987	12-week double-blind crossover	E-alone vs. placebo	Improved 30-meter walking time ($P = 0.03$)
Van den Brink (25), 1993	52-week double-blind placebo-controlled	E+P vs. placebo	Worse pain VAS at 12 months ($P = 0.01$)
Hall (23), 1994	6-month single-blind placebo-controlled	E+P vs. placebo	No difference in noncompliers; improvement in articular index ($P < 0.001$), pain score ($P < 0.05$), and morning stiffness ($P < 0.001$) in compliers
MacDonald (24), 1994	48-week double-blind placebo-controlled	Estrogen patches \pm progestin vs. placebo	Improvement in Nottingham Health Care Profile ($P < 0.01$)
D'Elia (22), 2003	2-year single-blind placebo-controlled	E+P (2 types) or E-alone vs. placebo	Improvement in ESR ($P = 0.025$), DAS28 ($P = 0.036$), Larsen score ($P = 0.026$)

* PHT = postmenopausal hormone therapy; RA = rheumatoid arthritis; OR = odds ratio; 95% CI = 95% confidence interval; ERT = estrogen replacement therapy; E-alone = estrogen alone; E+P = estrogen plus progestin; RR = relative risk; VAS = visual analog scale; ESR = erythrocyte sedimentation rate; DAS28 = Disease Activity Score in 28 joints.