Serum Sex Steroids in Premenopausal Women and Breast Cancer Risk Within the European Prospective Investigation into Cancer and Nutrition (EPIC)


Background: Contrasting etiologic hypotheses about the role of endogenous sex steroids in breast cancer development among premenopausal women implicate ovarian androgen excess and progesterone deficiency, estrogen excess, estrogen and progesterone excess, and both an excess or lack of adrenal androgens (dehydroepiandrosterone [DHEA] or its sulfate [DHEAS]) as risk factors. We conducted a case–control study nested within the European Prospective Investigation into Cancer and Nutrition cohort to examine associations among premenopausal serum concentrations of sex steroids and subsequent breast cancer risk. Methods: Levels of DHEAS, (Δ4)-androstenedione, testosterone, and sex hormone binding globulin (SHBG) were measured in single prediagnostic serum samples from 370 premenopausal women who subsequently developed breast cancer (case patients) and from 726 matched cancer-free control subjects. Levels of progesterone, estrone, and estradiol were also measured for the 285 case patients and 555 matched control subjects who had provided information about the day of menstrual cycle at blood donation. Conditional logistic regression models were used to estimate relative risks of breast cancer by quartiles of hormone concentrations. All statistical tests were two-sided. Results: Increased risks of breast cancer were associated with elevated serum concentrations of testosterone (odds ratio [OR] for highest versus lowest quartile = 1.73, 95% confidence interval [CI] = 1.16 to 2.57; \( P \) trend = .01), androstenedione (OR for highest versus lowest quartile = 1.56, 95% CI = 1.05 to 2.32; \( P \) trend = .01), and DHEAS (OR for highest versus lowest quartile = 1.48, 95% CI = 1.02 to 2.14; \( P \) trend = .10) but not SHBG. Elevated serum progesterone concentrations were associated with a statistically significant reduction in breast cancer risk (OR for highest versus lowest quartile = 0.61, 95% CI = 0.38 to 0.98; \( P \) trend = .06). The absolute risk of breast cancer for women younger than 40 followed up for 10 years was estimated at 2.6% for those in the highest quartile of serum testosterone versus 1.5% for those in the lowest quartile; for the highest and lowest quartiles of progesterone, these estimates were 1.7% and 2.6%, respectively. Breast cancer risk was not statistically significantly associated with serum levels of the other hormones. Conclusions: Our results support the hypothesis that elevated blood concentrations of androgens are associated with an increased risk of breast cancer in premenopausal women. [J Natl Cancer Inst 2005;97:755–65]

Introduction

Endogenous sex steroids have long been implicated in the development of breast cancer (1). Breast cancer risk is increased, relative to general population levels before and after menopause.

Affiliations of authors: International Agency for Research on Cancer (IARC-WHO), Lyon, France (RK, SR, LD, CB, TN, PF, NS, ER), Epidemiology Unit, Cancer Research UK, University of Oxford, Oxford, UK (TK, AR), National Cancer Institute, Milan, Italy (FB, GS, VK), Julius Centre for Health Sciences and Primary Care, University Medical Centre, Utrecht, Netherlands (PP, CHvG, NCO-M), MRC Dunn Human Nutrition Unit, Welcome Trust/MRC Building, Cambridge, UK (SB), Clinical Gerontology Unit, Addenbrooke’s Hospital, Cambridge, UK (KTK), Centre for Nutrition and Health, National Institute for Public Health and the Environment, Bilthoven, Netherlands (HBBdM), German Institute of Human Nutrition, Department of Epidemiology, Potsdam-Rehbrücke, Germany (HB, PHL), Division of Clinical Epidemiology, German Cancer Research Centre, Heidelberg, Germany (JCC, GN), INSERM, Institut Gustave Roussy, Villejuif, France (FCC, AF, A. Thiebaut), Department of Epidemiology, Catalan Institute of Oncology, Barcelona, Spain (CAG), The Danish Cancer Society, Institute of Cancer Epidemiology, Copenhagen, Denmark (A. Tjønneland, AO), Department of Epidemiology and Social Medicine, Aarhus University, Aarhus, Denmark (KO), Sección Información Sanitaria, Consejería de Salud y Servicios Sanitarios de Asturias, Asturias, Spain (JRGQ), Epidemiology Department, Murcia Health Council, Murcia, Spain (MDC), Public Health Institute of Navarra, Spain (ABG), Public Health Division of Gipuzkoa, Health Department of the Basque Country, Donostia-San Sebastian, Spain (PA), Molecular and Nutritional Epidemiology Unit, CISP, Scientific Institute of Tuscany, Florence, Italy (DP), Department for Clinical and Experimental Medicine, Federico II University Naples, Italy (SP), Cancer Registry, Azienda Ospedaliera “Civile M.P. Arezzo,” Ragus, Italy (RT); Imperial College London, UK and University of Torino, Italy (PV); University of Athens Medical School, Athens, Greece (EC, A. Trichopoulou, DT).

Correspondence to: Dr. Rudolf Kaaks, Nutrition and Hormones Group, International Agency for Research on Cancer (IARC-WHO), 150 cours Albert-Thomas, 69372 Lyon cedex 08, France (e-mail: kaaks@iarc.fr).

See “Notes” following “References.”

DOI: 10.1093/jnci/dji132
Journal of the National Cancer Institute, Vol. 97, No. 10, © Oxford University Press 2005, all rights reserved.
among women who had an early menarche and is reduced among women who experience early natural menopause or who have bilateral ovariectomy before age 45 years. These observations implicate both the early onset and the total lifetime duration of ovarian steroidogenic activity—particularly the synthesis of progesterone and/or estrogens—in breast cancer development.

Prospective cohort studies have shown that postmenopausal women who have elevated blood levels of either androgens of adrenal and/or ovarian origin (i.e., dehydroepiandrosterone [DHEA] or its sulfate [DHEAS], androstenedione, and testosterone) or estrogens (i.e., estrone and estradiol) have an increased risk of breast cancer risk (2, 3; Kaaks R, Rinaldi S, Key TJ, Berro F, Peeters PH, Biessy C, et al.: unpublished observations). Epidemiologic studies that have examined associations among breast cancer risk and premenopausal plasma levels of sex steroids have produced unclear results. This type of study is particularly complicated for estradiol and progesterone because the circulating levels of these hormones vary greatly during the menstrual cycle. In addition, previous prospective cohort studies have been limited by the low numbers of incident cases of breast cancer among premenopausal women and by the relatively few women of premenopausal age included in the cohorts.

To our knowledge, only five prospective studies have addressed the relationship between breast cancer risk and premenopausal blood levels of estradiol (4–8). Collectively, these studies, which included a total of 225 women who developed breast cancer subsequent to blood donation, showed that the women who developed breast cancer had somewhat higher pre-diagnostic blood levels of estradiol than women who did not develop breast cancer; however, this difference was not statistically significant, either in any of the individual studies or in combined analysis of data from all five studies (9). With regard to the androgens, early data from analyses carried out between 1961 and 1986 in a prospective cohort of women living on the island of Guernsey in the British Isles (the Guernsey cohort) indicated a possible inverse relationship between levels of the adrenal androgen DHEAS and its metabolites in blood (and urine) and breast cancer risk in young women (10, 11). However, these results were based on a relatively small number of observations and have not been confirmed by subsequent prospective studies (12). In the Guernsey cohort, there was also no statistically significant association between testosterone levels and breast cancer risk, although testosterone levels were slightly higher among breast cancer case patients than among control subjects (7). However, a prospective study in an Italian cohort (13) in which all female participants provided blood samples during days 20–24 after the start of their last menstrual cycle showed a statistically significant increase in breast cancer risk for women who had elevated levels of testosterone and comparatively low levels of progesterone.

We conducted a case–control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) to examine relationships among premenopausal serum concentrations of sex steroids and subsequent breast cancer risk. Our study included 370 incident breast cancer case patients and 726 matched control subjects and thus was approximately twice the size of all previously published cohort studies on breast cancer risk in relation to premenopausal serum levels of testosterone, androstenedione, or estradiol, and equal in size to all studies on DHEAS combined.

**Subjects and Methods**

**Baseline Collection of Data and Blood Samples and Inclusion of Research Centers**

The EPIC project is a European network of prospective cohorts that was set up to examine relationships of cancer risk with nutrition and metabolic risk factors (14). The EPIC cohorts include a total of approximately 250,000 women and 150,000 men, recruited through 23 regional and national research centers located in 10 western European countries: Norway (Tromsø), Sweden (Umeå and Malmö), Denmark (Copenhagen and Århus), England (Oxford and Cambridge), The Netherlands (Utrecht and Bilthoven), Germany (Potsdam and Heidelberg), France (Paris), Spain (Oviedo, San Sebastian, Pamplona, Murcia, and Granada), Italy (Turin, Milan, Florence, Naples, and Ragusa) and Greece (Athens). Extensive standardized questionnaire data on diet and nondietary variables, anthropometric measurements, and blood samples were collected from all study subjects between 1992 and 1998. Questionnaires included detailed questions about menstrual and reproductive history and current and past use of oral contraceptives (OCs) (women only) as well as questions about history of previous illness and surgical operations, lifetime history of tobacco smoking and consumption of alcoholic beverages, usual diet, and physical activity (men and women). Questionnaires were highly standardized across all countries contributing to the present study. For all cohorts except Oxford, the height and weight of all female participants were measured to the nearest centimeter and kilogram, respectively, according to standardized protocols, during a visit to a recruitment center. For 90% of the Oxford cohort, however, height and weight were self-reported. For all EPIC study centers, questionnaire and anthropometric data (as well as data on follow-up for cancer incidence and vital status; see below) were checked for coding errors and internal consistency of information and were entered into a central database at the International Agency for Research on Cancer (IARC; Lyon, France).

The present study includes case patients whose breast cancers were diagnosed after blood collection and matched control subjects from 19 of the 23 recruitment centers located in eight of the 10 participating countries: Denmark, France, Germany, Greece, Italy, The Netherlands, Spain, and the United Kingdom. Norway was not included in the present study because blood samples have been collected only recently from a subsample of cohort participants and only very few cases of breast cancer had been diagnosed after blood donation. Sweden was not included because, at the time of blood donation, we had limited questionnaire information about menopausal status, past and current OC use, and phase of menstrual cycle.

Nonfasting blood samples (at least 30 mL) were drawn from participants in the eight countries contributing to the present study. All samples except for those from the Oxford center were stored at 5–10°C and protected from light from the time of collection through their transfer to local laboratories in each of the 23 recruitment centers, where they were further processed and separated into aliquots. For study subjects recruited through the Oxford center, blood samples were collected by a network of general practitioners in the United Kingdom and transported to a central laboratory in Oxford by mail; they were protected from light but were exposed to ambient temperature [the stability of serum sex steroid levels measured in blood transported at ambient...
temperatures or kept for at ambient temperatures for 24 or 48 hours has been documented previously (15–17)]. For participants in seven of the eight countries included in the present study (i.e., all except Denmark), 0.5-mL aliquots of the blood fractions (serum, plasma, red cells, and buffy coat for DNA extraction) were placed in 28 plastic straws (12 aliquots of plasma, eight aliquots of serum, four aliquots of erythrocytes, and four aliquots of buffy coat), which were heat-sealed and stored in liquid nitrogen (−196°C); half of the aliquots (i.e., six aliquots of plasma, four aliquots of serum, two aliquots of erythrocytes, two aliquots of buffy coat) were stored in each of the local research centers, and the other half were stored centrally at the IARC. In Denmark, 1-mL aliquots of the blood fractions were placed into Nunc tubes and stored in the vapor phase of liquid nitrogen containers (−150°C).

Assessment of Menopausal Status and Phase of Menstrual Cycle at Blood Donation

Women were considered premenopausal at the time of blood donation if they reported having had at least nine menstrual periods over the previous 12 months. Women who had missing or incomplete questionnaire data or menstrual periods or who had had a hysterectomy were considered premenopausal if they were younger than 42 years because among the female EPIC participants who had complete questionnaire data, 99.5% of those younger than age 42 years were premenopausal. Women aged 42 years or older whose questionnaire data on menopausal status were incomplete or equivocal (i.e., because they had had a hysterectomy or used exogenous hormones) were excluded from the study.

Two different methods were used to determine the phase of a woman’s menstrual cycle at blood donation: “forward dating” counted forward from the woman’s reported date of the start of her last menses, and “backward dating” counted backward from the date of the start of her next menses after blood donation, which the woman reported on a prepaid postcard that she sent back to the recruitment center after her visit to donate a blood sample. Both dating methods were used for some participants; in such cases we used the information from the backward dating method because the length of the second half of the cycle (luteal phase) is generally more constant among women than that of the first half (follicular phase). With forward dating, the first day of a woman’s last period was set to 0 days, and all subsequently days were counted sequentially, up to a maximum of 39 days. With backward dating, the first day of the woman’s next period after blood donation was set to 28 days, and the date on which a blood sample had been provided was counted backwards from this date to a minimum acceptable value of −11 days. Thus, the first day of a woman’s next reported menses, using backward dating, corresponded to day 28 for woman whose cycle day was determined by forward dating, assuming an average length of cycle of 29 days for all women (the average length observed among premenopausal women in EPIC). When the date of blood donation differed by more than 40 days from a woman’s most recent (either last or next) reported date of her menstrual period, menstrual cycle phase was considered undetermined. In France, The Netherlands, Greece, and Germany, data were available only for forward dating of the phase of menstrual cycle at blood donation, whereas for the vast majority of cohort participants in Italy, Spain, and Oxford, data were also available for backward dating.

Follow-up for Cancer Incidence and Vital Status

In Denmark, The Netherlands, the United Kingdom, Spain, and Italy, incident breast cancer cases were identified through record linkage with regional cancer registries. In France, Germany, and Greece, incident breast cancers were ascertained by active follow-up through a national health insurance company (France) or through direct contacts with study subjects or their next of kin (France, Germany, and Greece). In all study centers, a breast cancer diagnosis was confirmed by a comprehensive review of pathology reports that followed a detailed protocol for the collection and standardization of clinical and pathology data. In most EPIC study centers, data on vital status were collected from regional or national mortality registries; in Greece and Germany, data on vital status were collected by active follow-up. For each EPIC study center, the closure date for the study period was defined as the latest dates of complete follow-up for both cancer incidence and vital status (closure dates varied among centers and ranged from June 1988 to December 2000 at the time, i.e., November 2002, when we started planning the present case-control study).

Selection of Study Subjects

Study subjects (case patients and control subjects) were selected from among the women in each cohort who, at blood donation, reported that they did not use exogenous hormones for contraception or medical purposes and who had not been previously diagnosed with cancer (except for nonmelanoma skin cancer). Case patients were women who developed breast cancer after they were recruited into the EPIC study and donated blood and before the closure date of the study period.

The eight countries that were included in the present study contributed 61,943 women who were premenopausal at the time of blood donation. By November 2002, 491 women within this subcohort were reported to have had a diagnosis of breast cancer; among these women, 445 used no exogenous hormones at the time of blood donation, 424 had serum samples that were sufficient for subsequent analyses, and 419 were matched to control subjects. Among the 416 case patients for whom we had complete hormone assays using methods described below, 46 had carcinoma in situ and 370 had an invasive tumor. Results of the present study are for women diagnosed with invasive cancers only.

For each breast cancer case patient, two control subjects were chosen at random among all cohort members who were alive and free of cancer (except nonmelanoma skin cancer) at the time of diagnosis of the case patient. We used an incidence density sampling protocol to choose control subjects; i.e., control subjects could include women who might become case patients at a later point in time, and each control subject could be sampled more than once. Matching characteristics included the study center where the subjects were enrolled in the cohort and gave their blood sample and age (±6 months), the time of the day (±1 hour), fasting status, and phase of menstrual cycle at blood donation. We used five matching categories for phase of menstrual cycle: “early follicular” (days 0–7 of the cycle), “late follicular” (days 8–11), “periovulatory” (days 12–16), “midluteal” (days 20–24), and “other luteal” (days 17–19 or days 25–40). In total, 285 breast cancer case patients and 555 control subjects could be matched on menstrual cycle phase: Case patients and control subjects were matched by whether their menstrual cycle phase
was determined by forward dating (154 case patients and 305 control subjects) or backward dating (131 case patients and 250 control subjects). The number of breast cancer case patients included in our analysis, by country, were seven from Germany, nine from Greece, 30 from France, 30 from Denmark, 43 from the Netherlands, 58 from the United Kingdom (of which 44 were from Oxford), 88 from Spain, and 105 from Italy.

**Hormone Assays**

Serum concentrations of testosterone and DHEAS were measured by radioimmunoassay (ImmunoTech, Marseilles, France). Serum androstenedione concentrations were measured by a radioimmunoassay that used a double-antibody system for the separation of free and bound antigen (Diagnostic Systems Laboratories, Webster, TX). Serum concentrations of SHBG were measured by a solid phase “sandwich” immunoradiometric assay (Cis-Bio International, Gif-sur-Yvette, France). Serum concentrations of DHEAS were measured for 370 breast cancer case patients and 725 control subjects, serum concentrations of androstenedione for 370 case patients and 724 control subjects, serum concentrations of testosterone for 369 case patients and 715 control subjects, and serum concentrations of SHBG for 369 case patients and 724 control subjects.

Measurements of serum concentrations of progesterone, estradiol, and estrone were attempted only for the 285 breast cancer case patients and 555 matched control subjects for whom the phase of menstrual cycle at blood donation could be determined. Progesterone measurements were completed for 278 case patients and 535 control subjects, using an assay from Immunotech, estradiol measurements were completed for 284 case patients and 551 control subjects, using an assay from DiaSorin (Saluggia, Italy), and estrone measurements were completed for 284 case patients and 552 control subjects, using an assay from Diagnostic Systems Laboratories. All hormones except androstenedione were measured in duplicate.

All assays were performed by study personnel at the Nutrition and Hormones Group Laboratory at IARC who were blinded as to the case–control status of the study subjects. Samples pertaining to matched study subjects, with each matched set containing samples from one case patient and two matched control subjects, were always analyzed in the same analytical batch (i.e., a single immunoassay kit could assay 72–76 serum samples and 24–28 standards and quality control samples). Based on results obtained for the quality control samples, the mean intrabatch coefficients of variation were estimated to be 6.1% for DHEAS, 8.4% for testosterone, 5.8% for androstenedione, 6.7% for SHBG, 4.9% for estrone, 3.2% for estradiol, and 7.4% for progesterone. For the hormones that were measured in duplicate within the same batch, the within-batch intraclass coefficients of correlation were 0.94 for testosterone and greater than 0.96 for the other hormones. Interbatch coefficients of variation were 7.2% for estradiol, 9.4% for progesterone, 15.3% for testosterone, 18.9% for androstenedione, 16.5% for SHBG, 12.4% for DHEAS, and 12.6% for estrone.

**Statistical Analyses**

We used analysis of variance, adjusting for age and case–control status, to examine study (recruitment) center as a determinant of measured hormone levels. Statistical significance of case–control differences in mean hormone levels and in continuous baseline covariates were evaluated by paired t tests of the case patient value versus the average of the values for the two matched control subjects in each case–control set (18). For binary variables, the statistical significance of case–control differences was tested using a chi-square test. Pearson’s partial correlation coefficients, adjusted for age, laboratory batch, and case–control status, were computed to examine the strength of linear associations among the various hormones and between each hormone and body mass index (BMI). Progesterone, estradiol, and estrone showed systematic variation over the estimated phases of the menstrual cycle; no such variation was observed for DHEAS, androstenedione, and testosterone. We used spline regression methods (the nonlinear regression [NLR]) procedure of the Statistical Analysis System, Version 8 [SAS Institute; Cary, NC] to model the systematic variations in the levels of the former three hormones during the menstrual cycle for case patients and control subjects combined, and we calculated residuals of the spline regression models to describe how the hormone levels for individuals deviated from the predicted average hormone levels at a given day of the cycle. By using the residuals of these regression models as the exposure variables of interest, we were able to use the same set of standardized cut points (i.e., quartiles) for different categories of serum levels of estradiol and progesterone, irrespective of the menstrual cycle phase. We also performed analyses on estradiol and progesterone levels without correcting by spline modeling; however, we analyzed the levels separately for the first (i.e., follicular) phase of the cycle (day 12 of cycle or earlier) and the second (i.e., luteal) phase (after day 12 of the cycle).

Relative risks (odds ratios [ORs]) for breast cancer in relation to different serum hormone levels were calculated by conditional logistic regression, using the PHREG procedure of the SAS software package (Version 8). For DHEAS, androstenedione, and testosterone, all analyses were performed on the absolute serum concentrations, and for progesterone, estradiol, and estrone, odds ratios were estimated with respect to the residuals of the spline regression models of systematic variations of hormone concentrations during the menstrual cycle. The serum levels of the various sex steroids (residuals of the spline models for progesterone, estradiol, and estrone) and of SHBG were examined by quartile categories, using quartile cut points that were based on the distributions among the control subjects from all EPIC centers combined. We used likelihood ratio tests to assess linear trends in odds ratios with increasing exposure level as a continuous variable or with assigned quantitative scores equal to the medians for the four quartile categories. We computed 95% confidence intervals (CIs) from the standard errors of the pertinent regression coefficients.

The effects of additional potential confounders (other than the matching criteria, which were controlled for by design) were examined by including additional regression terms into the logistic regression models. Potential confounders included BMI as a continuous variable, age at menarche (younger than 12 years, 12 years, 13 years, 14 years, older than 14 years, missing), past use of exogenous hormones (i.e., oral contraceptives) as a binary variable (yes, no, or missing), age at first full-term pregnancy (0 [no full-term pregnancies], younger that 23 years, 24–25 years, 26–28 years, 29 years or older, or missing), number of full-term pregnancies (0, 1, 2, 3, 4 or more, or missing), and history of breast feeding (yes, no, or missing).
Tests for heterogeneity of associations among hormone levels and breast cancer risk (e.g., by study center, age at breast cancer diagnosis, time between blood donation and cancer diagnosis, and phase of menstrual cycle) were performed using chi-square tests. All statistical tests were two-sided.

Absolute risks of developing breast cancer for a 10-year follow-up were calculated using the method described by Gail et al. (19). These calculations were performed for women in four different age categories (<40, 40–44, 45–49, and ≥50 years). The age-specific hazards of dying from other causes than breast cancer were set at $9 \times 10^{-4}$, $1.6 \times 10^{-4}$, $2.1 \times 10^{-4}$, and $4.8 \times 10^{-4}$, respectively, for each of these four age groups, based on average national breast cancer mortality rates in Western industrialized countries. Similar figures were used by Gail et al. (19).

**RESULTS**

The mean age of the study subjects at blood donation was 45.6 years (5th–95th percentile range = 37–53 years) (Table 1). Cancer was diagnosed, on average, 2.8 years after blood donation (5th–95th percentile range = 0.2–5.8 years), and the vast majority of case patients (95%) were diagnosed with breast cancer before age 55 years. Compared with control subjects, case patients were older at their first full-term pregnancy, although this difference was not statistically significant ($P = .07$). Case and control subjects did not show differences in previous use of oral contraceptives, number of full-term pregnancies, age at menarche, or mean BMI (Table 1).

Spline models revealed that serum levels of progesterone varied during the subjects’ menstrual cycles, with a strong increase in levels occurring after day 12 of the cycle, which corresponds to the luteal phase of the cycle, during which progesterone is normally produced (Fig. 1). Spline models also revealed a clearly cyclic pattern for serum estradiol levels, with the first peak occurring during the late follicular phase and a second peak for days of the cycle corresponding to the midcycle, preovulatory, and luteal phases, a pattern identical to that described by Rosenberg et al. for the prospective cohort of the New York University Women’s Health Study (5). Serum estrone levels showed the least variation throughout the menstrual cycle; only a slight increase in serum levels was observed during the middle part of the cycle (Fig. 1).

To examine the degree of between-center variations in average serum hormone levels relative to total between-subject variation, we performed an analysis of variance, with age at blood donation, study center, and case–control status as explanatory variables. This analysis revealed statistically significant variations among the 19 study centers with respect to the serum concentrations of SHBG and of all sex steroids except estrone. However, the magnitude of this variation was small when expressed as a percentage of total between-subject variance (4.1% for DHEAS, 10.1% for androstenedione, 8.4% for testosterone, 6.4% for progesterone, 5.1% for estradiol, and 10.3% for SHBG). Further adjustment for BMI did not alter the results of these analyses.

Partial correlation coefficients adjusted for assay batch, age at blood donation, and case–control status indicated that serum concentrations of the three androgens (DHEAS, androstenedione, and testosterone) were directly correlated with each other ($r$ range = .40–.62) (Table 2). For estrone, estradiol, and progesterone, the spline-adjusted residuals correlated highly with the original, nonadjusted variables of estradiol and progesterone ($r$ range = .79–.95), reflecting the fact that the spline models explained only a modest proportion of the overall variation in serum concentrations of these hormones (the spline model explained only 10% of the variation in serum estrone concentrations and 38% of the variation in serum progesterone concentrations; see Fig. 1). All androgen levels correlated weakly with spline-adjusted residual levels of the estrogen ($r$ range = .06–.19), whereas SHBG levels were inversely correlated with BMI ($r = -0.28$) and DHEAS ($r = -0.21$) and were weakly correlated with estradiol ($r = .19$ for the spline-adjusted residuals).

Table 1. Baseline characteristics of breast cancer case patients and control subjects*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Case patients</th>
<th>Control subjects</th>
<th>$P$ for difference†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of subjects</td>
<td>370</td>
<td>726</td>
<td>—</td>
</tr>
<tr>
<td>Number of subjects matched</td>
<td>285</td>
<td>555</td>
<td>—</td>
</tr>
<tr>
<td>on menstrual cycle phase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at recruitment, y (5th – 95th percentiles)</td>
<td>45.6 (37.7–52.9)</td>
<td>45.6 (37.3–52.7)</td>
<td>.16</td>
</tr>
<tr>
<td>Mean age at blood donation, y (5th – 95th percentiles)</td>
<td>45.9 (37.7–52.9)</td>
<td>45.9 (37.6–52.7)</td>
<td>.32</td>
</tr>
<tr>
<td>Mean age at diagnosis, y (5th – 95th percentiles)</td>
<td>48.2 (40.0–55.0)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Mean no. of years between blood</td>
<td>2.8 (0.2–5.8)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>donation and diagnosis (5th–95th percentiles)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at first full-term pregnancy, y (5th–95th percentiles)</td>
<td>25.6 (19.0–33.0)</td>
<td>24.9 (19.0–32.0)</td>
<td>.07</td>
</tr>
<tr>
<td>Mean no. of full-term pregnancies (5th–95th percentiles)</td>
<td>2.2 (1.0–4.0)</td>
<td>2.2 (1.0–4.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Mean age at menarche, y (5th – 95th percentiles)</td>
<td>12.8 (11.0–15.0)</td>
<td>12.9 (11.0–15.0)</td>
<td>.12</td>
</tr>
<tr>
<td>Mean body mass index, kg/m² (5th – 95th percentiles)</td>
<td>25.1 (19.8–33.0)</td>
<td>25.3 (19.9–33.7)</td>
<td>.45</td>
</tr>
<tr>
<td>Past use of oral contraceptive, %</td>
<td>62.7</td>
<td>61.2</td>
<td>.64</td>
</tr>
<tr>
<td>Ever had a full-term pregnancy, %</td>
<td>82.4</td>
<td>84.8</td>
<td>.29</td>
</tr>
</tbody>
</table>

*Not applicable.
†$P$ values for two-sided $t$ test (all variables except “past use of oral contraceptive” and “ever had a full-term pregnancy”) or for two-sided chi-square test (“past use of oral contraceptive” and “ever had a full-term pregnancy”).
absolute risk of breast cancer for women younger than 40 years followed up for 10 years was estimated at 2.6% for the highest quartile of serum testosterone versus 1.5% in the lowest quartile.

Likewise, in conditional logistic regression analyses, women in the highest quartile of mean serum progesterone concentration had a statistically significantly lower risk of breast cancer than women in the lowest quartile of mean serum progesterone concentration (OR for the highest versus the lowest quartile = 0.61, 95% CI = 0.38 to 0.98; \( P_{\text{trend}} = .06 \)). This inverse relationship was driven mostly by the observations for women who were in the second (luteal) phase of their menstrual cycle when they provided their blood samples. When we restricted this analysis to women who were in the luteal phase of their menstrual cycles when they donated blood and expressed progesterone levels without adjustment by spline models, we still observed an inverse association between progesterone levels and breast cancer risk, but the association was no longer statistically significant (OR for the highest versus the lowest quartile = 0.66, 95% CI = 0.33 to 1.31; \( P_{\text{trend}} = .11 \)).

Serum levels of SHBG, estradiol, and estrone were not associated with breast cancer risk, either overall (Fig. 2) or by menstrual cycle phase, irrespective of the method of analysis chosen for the estrogens (spline-adjusted residuals or unadjusted values on the original scale) (data not shown). For the highest versus lowest quartiles of progesterone, absolute risks of breast cancer for a 10-year follow-up were estimated at 1.7% and 2.6%, respectively, for women younger than 40 years.

We examined the effects of various confounders on these relative risk estimates by performing subgroup analyses. Relative risk estimates remained virtually unchanged after adjustments for BMI, age at first full-term pregnancy, age at menarche, number of full-term pregnancies, and past use of OCs (results not shown). Furthermore, when center-specific quartile cut points were used for hormone levels, relative risk estimates were very similar to the relative risk estimates from analyses that used EPIC-wide cut points. Finally, estimates of relative risks of breast cancer associated with hormone levels expressed as quartiles or on a continuous scale showed no statistically significant heterogeneity among study centers (results not shown).

The inverse association between progesterone level and breast cancer risk was present only in the subgroup of case patients and control subjects who were matched for phase of menstrual cycle at blood donation (OR for the highest versus the lowest quartile = 0.48, 95% CI = 0.26 to 0.88; \( P_{\text{trend}} = .08 \)), and not in the subgroup of case patients and control subjects who were matched by backward dating (OR for the highest versus the lowest quartile = 0.90, 95% CI = 0.42 to 1.96; \( P_{\text{trend}} = .97 \)); however, there was no association between estrone or estradiol levels and breast cancer risk in either subgroup (results not shown).

Among the 340 women for whom cycle length could be calculated from dates of both their last and their next menses, we found no statistically significant association between cycle length and progesterone levels, either globally (\( r = -.09, P = .10 \)) or by phase of the menstrual cycle (\( r = -.00, P = .99 \) in the luteal phase; \( r = -.02, P = .81 \) in the follicular phase), or between cycle length and breast cancer risk (data not shown).

The relative risk estimates changed only slightly after we excluded women who were diagnosed with breast cancer within either the first 3 or the first 6 months after blood donation. There was also no statistically significant heterogeneity between relative risk estimates for women who were diagnosed with breast

For all study subjects combined, women who developed breast cancer had statistically significantly higher mean levels of testosterone, androstenedione, and DHEAS, nonstatistically significantly higher mean levels of the estrogens, and statistically significantly lower mean levels of progesterone than the control subjects (Table 3). In conditional logistic regression analyses, these differences reflected the increasing relative risks of breast cancer associated with increasing prediagnostic serum levels of testosterone (OR for the highest versus the lowest quartile = 1.73, 95% CI = 1.16 to 2.57; \( P_{\text{trend}} = .01 \)), androstenedione (OR for the highest versus the lowest quartile = 1.56, 95% CI = 1.05 to 2.32; \( P_{\text{trend}} = .01 \)), and DHEAS (OR for the highest versus the lowest quartile = 1.48, 95% CI = 1.02 to 2.14; \( P_{\text{trend}} = .10 \)) (Fig. 2). The

Fig. 1. Serum levels of progesterone, estradiol, and estrone by day of menstrual cycle at blood donation. Dotted lines = mean serum hormone concentrations as estimated from 3-day averages, estimated over successive 3-day intervals. Solid lines = mean serum hormone concentrations predicted by spline models.
cancer less than 2 years after blood donation and those for women who were diagnosed with breast cancer more than 2 years after blood donation. However, when we restricted the analyses to the 174 women who were younger than 49 years (the median age at diagnosis for all case patients, which is just below the median age at menopause [50 years] in the EPIC cohort) at breast cancer diagnosis, the associations among breast cancer risk and hormone levels were slightly stronger than those in the overall analysis (for DHEAS: OR = 2.00 [95% CI = 1.14 to 3.51], P_trend = .02; for testosterone: OR = 2.22 [95% CI = 1.18 to 4.18], P_trend = .02; and for androstenedione: OR = 1.92 [95% CI = 1.06 to 3.45], P_trend = .04), as was the inverse relationship between breast cancer risk and progesterone levels based on residuals of the spline models (OR = 0.52, 95% CI = 0.28 to 0.95; P_trend = .04). For women who were 49 years or older at breast cancer diagnosis (n = 196), these associations were less pronounced and not statistically significant. However, there was no statistically significant heterogeneity in relative risk estimates by age at breast cancer diagnosis for any of the hormones.

A subgroup analysis dichotomized by median age at blood donation (46.2 years) produced results that were similar to those of the subgroup analysis dichotomized by age at tumor diagnosis (results not shown); this result was not surprising, given that follow-up times (i.e., the time between blood donation and breast cancer diagnosis) were relatively short, so that there was a high correlation (r = .92) between age at blood donation and age at tumor diagnosis.

**DISCUSSION**

In this large multicenter cohort study, we found that premenopausal women who had elevated serum levels of testosterone or androstenedione or low levels of progesterone had an increased risk of breast cancer. Our data showed no clear relationship between breast cancer risk and premenopausal serum levels of estrone or estradiol.

Our study has several strengths. In contrast with traditional case–control studies, our study used a prospective design, which avoided inverse causation bias that may occur when alterations in endogenous hormone levels are induced by metabolic effects of a tumor or by antitumor treatments, psychological stress, or lifestyle changes after cancer diagnosis. Although our study included subjects from 19 recruitment centers in eight European countries, recruitment and blood collection protocols and questionnaire data were standardized across study centers. Furthermore, all hormone assays were performed in one laboratory, and samples from case patients and matched control subjects were always assayed at the same time. Our finding that women who had elevated premenopausal testosterone levels had an increased risk of breast cancer is consistent with observations made in another recent, carefully designed prospective study (13) in the Italian ORDET cohort (see discussion below) and is compatible with findings from the Guernsey cohort, in which women who developed breast cancer also had moderately (but nonstatistically significantly) higher levels of testosterone than control subjects (7).

In the 1960s, Grattarola (20) hypothesized that breast cancer risk is increased among women who have an ovarian androgen excess, chronic anovulation, and an associated reduction of luteal-phase progesterone production [later called the ovarian hyperandrogenism/luteal inadequacy hypothesis (21)]. This hypothesis was based on the observation that breast cancer patients often show hyperplasia of the endometrium—a pathognomonic symptom of ovarian androgen excess, chronic anovulation, and progesterone deficiency—and received some initial confirmation from a series of traditional case–control studies that showed that breast cancer patients have higher plasma or urinary concentrations of testosterone or its urinary metabolites (21,22) than cancer-free control subjects (20,23,24). However, there is conflicting experimental evidence regarding the possible role of androgens in breast cancer development: androgens have been found to have both stimulatory and inhibitory effects on the proliferation

### Table 2. Pearson’s partial correlation coefficients (P values) for associations among endogenous hormone levels and BMI*

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Testosterone</th>
<th>Androstenedione</th>
<th>DHEAS</th>
<th>SHBG</th>
<th>Estrone residuals</th>
<th>Estradiol residuals</th>
<th>Progesterone residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrone, pmol/L</td>
<td>0.19 (&lt;.001)</td>
<td>0.15 (&lt;.001)</td>
<td>-0.21 (&lt;.001)</td>
<td>0.08 (&lt;.001)</td>
<td>0.02 (.68)</td>
<td>0.04 (.26)</td>
<td>0.01 (.71)</td>
</tr>
<tr>
<td>Estradiol, pmol/L</td>
<td>0.08 (.03)</td>
<td>0.12 (.001)</td>
<td>-0.06 (.08)</td>
<td>0.19 (&lt;.001)</td>
<td>0.68 (&lt;.001)</td>
<td>0.04 (.28)</td>
<td>-0.1 (.003)</td>
</tr>
<tr>
<td>Progesterone</td>
<td>-0.02 (.62)</td>
<td>0.06 (.10)</td>
<td>-0.01 (.68)</td>
<td>0.07 (.71)</td>
<td>0.04 (.26)</td>
<td>0.01 (.71)</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>0.13 (&lt;.001)</td>
<td>0.02 (.42)</td>
<td>0.15 (&lt;.001)</td>
<td>-0.28 (&lt;.001)</td>
<td>-0.15 (&lt;.001)</td>
<td>-0.16 (.08)</td>
<td>-0.06 (.08)</td>
</tr>
</tbody>
</table>

*Analyses were adjusted for age at breast donation, laboratory batch, and case–control status. Residuals of estrone, estradiol, and progesterone were calculated as deviations of estrone, estradiol, and progesterone levels from mean levels predicted by from spline regression models, to adjust for menstrual cycle variations. BMI = body mass index; DHEAS = dehydroepiandrosterone sulfate; SHBG = sex hormone binding globulin.

**Table 3. Mean hormone levels (95% confidence intervals) for breast cancer case patients and control subjects**

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Case patients</th>
<th>Control subjects</th>
<th>P for difference†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testosterone, nmol/L</td>
<td>1.80 (1.71 to 1.88)</td>
<td>1.66 (1.60 to 1.72)</td>
<td>.01</td>
</tr>
<tr>
<td>Androstenedione, nmol/L</td>
<td>5.58 (5.58 to 5.58)</td>
<td>5.15 (5.15 to 5.15)</td>
<td>.01</td>
</tr>
<tr>
<td>DHEAS, μmol/L</td>
<td>3.93 (3.93 to 3.93)</td>
<td>3.64 (3.50 to 3.79)</td>
<td>.03</td>
</tr>
<tr>
<td>SHBG, nmol/L</td>
<td>51.1 (51.1 to 51.1)</td>
<td>53.1 (50.7 to 55.5)</td>
<td>.21</td>
</tr>
<tr>
<td>Estrone, pmol/L</td>
<td>436.0 (436.0 to 436.0)</td>
<td>409.9 (383.4 to 436.4)</td>
<td>.23</td>
</tr>
<tr>
<td>Estradiol, pmol/L</td>
<td>400.2 (364.3 to 436.2)</td>
<td>381.8 (356 to 407.6)</td>
<td>.32</td>
</tr>
<tr>
<td>Progesterone, nmol/L</td>
<td>9.46 (7.55 to 11.37)</td>
<td>11.96 (10.59 to 13.34)</td>
<td>.01</td>
</tr>
</tbody>
</table>

*DHEAS = dehydroepiandrosterone sulfate; SHBG = sex hormone binding globulin.
†Two-sided P values; paired t test.
of mammary epithelial and cancer cells in vitro and on the growth of experimentally induced mammary tumors in animals (25–28). In addition to having direct androgenic effects on breast tissue, elevated serum androgen levels may lead indirectly to increased estrogenic exposures of breast tissue because all steroidogenic enzymes necessary for the formation of estrogens from androgenic precursor molecules are present in normal mammary tissues and breast tumor specimens (29).

According to a second major hypothesis associating breast cancer risk with hormone levels, breast cancer risk would be increased among women with an elevated combined exposure to both estrogens and progesterone. This “estrogen-plus-progesterone” hypothesis was motivated largely by observations of increased proliferation rates of breast epithelium during the luteal phase of the menstrual cycle, when the ovaries produce both estradiol and progesterone (30). A more recent observation that can be interpreted as providing support for the estrogen-plus-progesterone hypothesis is that postmenopausal estrogen-plus-progesterone replacement therapy increases breast cancer risk to a greater extent than replacement therapies that contain estrogen alone (31–36). However, the interpretation of this second observation with respect to the possible actions of natural progesterone is complicated by the fact that many synthetic progestins are derived from androgenic precursor molecules and also have some androgenic activity. Experimental studies have provided conflicting evidence regarding the possible role of progesterone in breast cancer development. Some animal studies have shown that progesterone is a critical determinant for the initiation of mammary tumors in response to carcinogens (37–39); however, other studies have shown that progesterone reduces estrogen-induced proliferation of breast epithelial cells (40–42).

Our findings of an increased risk of breast cancer among women with elevated premenopausal serum testosterone levels and comparatively low serum levels of progesterone—also reported in the Italian ORDET cohort (13)—do not support the

---

**Fig. 2.** Relative risk of breast cancer by sex steroid and sex hormone binding globulin (SHGB) levels in quartiles. Solid squares = odds ratios (ORs), estimated by conditional logistic regression, for quartiles of serum hormone concentrations (quartile cut points based on the distribution of control subjects); horizontal bars = 95% confidence intervals (CIs); \( P_{\text{trend}} = P \) value for a test of linear trend (two-sided test), using median hormone values for the four quartiles as quantitative scores of exposure. Analysis based on 370 case patients and 726 control subjects. Case patients and control subjects were matched on EPIC (European Prospective Investigation into Cancer and Nutrition) recruitment center, age at blood donation, phase of menstrual cycle, time of day at blood donation, and fasting status. DHEAS = dehydroepiandrosterone sulfate.
estrogen-plus-progestin hypothesis but do provide support for the ovarian hyperandrogenism hypothesis. However, although our findings were relatively robust for testosterone and androstenedione, our results for progesterone should be interpreted with caution. The information that women provided about the phase of menstrual cycle during which they gave their blood samples clearly allowed some prediction of variations in serum progesterone levels during the menstrual cycle, and this prediction (through spline regression models) helped to account for such systematic variation in case–control comparisons (Fig. 1). Nevertheless, it is clear that only part of the cyclic variation in progesterone levels may have been predicted accurately and that not all women may have been correctly classified by their menstrual cycle phase, especially if they had menstrual cycles that were considerably longer or shorter than the assumed average cycle length of 28 days. However, our data showed no evidence for any correlation between menstrual cycle length and serum progesterone levels or breast cancer risk.

Epidemiologic findings have led to opposing hypotheses about the relationship between breast cancer risk and the predominantly adrenal androgen DHEA or its sulfate, DHEAS (levels of which are usually highly correlated with DHEA levels in blood). In some studies, especially those within the Guernsey cohort, low serum or urine levels of DHEA, DHEAS, or their metabolites were associated with increased breast cancer risk among premenopausal women (1,9,10). By contrast, elevated blood levels of DHEA or DHEAS have been consistently associated with an increase in risk of breast cancer in postmenopausal women in all large prospective studies, including the EPIC cohort (2,3; Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, et al.: unpublished observations). To reconcile these contrasting observations, it has been proposed that, in a highly estrogenic milieu, such as the one that exists in premenopausal women, DHEA might act as an estrogen antagonist through competitive binding of its metabolite 5-androstene-3ß,17ß-diol (ADIOL) to the estrogen receptor; by contrast, in postmenopausal women, who have much lower serum estradiol levels, ADIOL would be more likely to act as a moderate estrogen agonist (43). In contrast to the Guernsey cohort studies, our data did not show a reduction in breast cancer risk among women who had elevated serum DHEAS, but did show a statistically significant direct association between serum levels of DHEAS and breast cancer risk that was even more pronounced when the analysis was limited to women who were younger than 49 years (the median age at diagnosis) when their breast cancers were diagnosed. Another recent prospective study within the Nurses’ Health Study cohort showed no clear association between breast cancer risk and premenopausal serum levels of DHEA or DHEAS, irrespective of age at subsequent breast cancer diagnosis (12).

A third hypothesis regarding the relationships between ovarian sex steroids and breast cancer risk stipulates that estrogens stimulate tumor development independently of other hormones (“estrogen-alone” hypothesis). This hypothesis has gained considerable support because of consistent reports from prospective cohort studies and the EPIC cohort (2,7; Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, et al.: unpublished observations) that postmenopausal women who have elevated serum concentrations of estrone and estradiol have an increased breast cancer risk. Experimental studies have also provided strong and consistent evidence that estrogens can promote breast tumor development and growth (44,45). However, as was the case for previous cohort studies on premenopausal women (4–8), we observed no clear relationship between serum estrogen levels and breast cancer risk. It is conceivable that breast cancer risk is related in a nonlinear fashion to circulating estrogen levels, such that a clear, direct relationship exists only within the lower, postmenopausal range of endogenous estrogen levels, but not at higher, premenopausal estrogen concentrations. However, we assayed only one blood sample for each woman, and it is also possible that, despite our efforts to match case patients and control subjects by the subphase of their menstrual cycles and our further adjustment for cyclic variations in serum estradiol levels by spline regression, the measured estradiol levels may not have accurately represented the average, long-term estradiol levels in the women in our study. Several studies that have addressed the reproducibility of serum estradiol measurements made over time in premenopausal women found that a single measurement of estradiol does not accurately reflect a woman’s long-term average blood concentration, unlike single measurements of androgens (46–48); however, this result may have been due, in part, to the lack of standardization of phase of menstrual cycle in some of these methodologic studies (47).

A limitation of our study was the absence of information about the menopausal status of the case patients at breast cancer diagnosis. However, given that the average time between blood donation and breast cancer diagnosis was only 2.8 years and that 95% of the case patients had their tumors detected within 6 years after they donated blood and before they reached age 55 years, most of the women who had gone through menopause between the times of blood donation and breast cancer diagnosis had been menopausal for a very short time. Although it is possible that the relationship between hormone levels and breast cancer risk could change abruptly after a woman’s transition through menopause, we believe that such an abrupt change is unlikely. Our main finding—that elevated serum testosterone and androstenedione levels are associated with increased risk of breast cancer—appeared to be equally applicable to women who were younger or older than 49 years (the median age at diagnosis) when they were diagnosed with breast cancer (although not statistically significant in either subgroup alone). It is interesting that similar observations were made in prospective studies of postmenopausal women, including the EPIC study. Thus, our data indicate that breast cancer risk may be associated with elevated serum levels of testosterone and androstenedione in both pre- and postmenopausal women.

In conclusion, this prospective cohort study provides strong evidence that the risk of breast cancer among premenopausal women is directly related to circulating levels of testosterone and androstenedione. Further studies are needed, however, to establish more clearly the physiologic origins—adrenal or ovarian—of the circulating androgens in women who are at an increased risk of breast cancer and to understand whether changes in lifestyle or other interventions can contribute to decreasing the risk of breast cancer by lowering circulating levels of androgens.

REFERENCES


NOTES

Results of this nested case-control study were obtained with financial support from the National Cancer Institute (USA) grant 1U01CA98216-01. The EPIC study was funded by the “Europe Against Cancer” Programme of the European Commission (SANCO); Ligue contre le Cancer (France); Société 3M (France); Mutuelle Générale de l’Éducation Nationale; Institut National de la Santé et de la Recherche Médicale (INSERM); German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund (FIS) of the Spanish Ministry of Health; the participating regional governments and institutions of Spain; USCIII, Red de Centros RCESP, C03/09; Cancer Research UK; Medical Research Council, UK; the Stroke Association, UK; British Heart Foundation; Department of Health, UK; Food Standards Agency, UK; the Wellcome Trust, UK; Greek Ministry of Health; Greek Ministry of Education; Italian Association for Research on Cancer; Italian National Research Council; Dutch Ministry of Public Health, Welfare and Sports; Dutch Ministry of Health; Dutch Prevention Funds; LK Research Funds; Dutch ZON (Zorg Onderzoek Nederland); World Cancer Research Fund (WCRF); Swedish Cancer Society; Swedish Scientific Council; Regional Government of Skane, Sweden; and Norwegian Cancer Society.

Manuscript received September 16, 2004; revised March 31, 2005; accepted April 4, 2005.