Treg-mediated immune tolerance and the risk of solid cancers: Findings from the EPIC-Heidelberg study

Sebastian Barth[1], Janika Schulze[2], Tilman Kühn[1], Eva Raschke[2], Theron Johnson[1], Rudolf Kaaks[1] and Sven Olek[2]

[1]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[2]Ivana Türbachova Labor für Epigenetik, Epiontis GmbH, Berlin, Germany

Corresponding Authors:

Sven Olek, Epiontis GmbH, Tel: +49-30-63923475; Fax: +49-30-63923476; Email: sven.olek@epiontis.com

Rudolf Kaaks, DKFZ, Tel: 06221 422219; Fax: 06221 422203; Email: r.kaaks@dkfz.de

Running Title: Peripheral immune tolerance and cancer risk

Key Words: Foxp3$^+$ regulatory T-Lymphocytes (Tregs), ImmunoCRIT, immune tolerance, prospective study, cancer risk

Word count: 2.956

Abstract

Background: Laboratory-based mechanistic and prognosis studies suggest that a shift from anti-tumor immunity towards tumor-immune tolerance plays a major role in carcinogenesis. However, prospective epidemiological studies analyzing the association between inter-individual variation of tolerance levels in healthy individuals and cancer risk are missing.

Methods: A case-cohort study embedded in EPIC-Heidelberg was conducted comprising incident cases of breast (n=399), colorectal (n=185), lung (n=149), and prostate (n=378) cancer, which occurred during 6.6 years of follow-up, and a sub-cohort (n=807). Foxp3+ Regulatory T-lymphocytes and CD3+ T-lymphocytes were measured by qPCR-based DNA methylation analysis in pre-diagnostic leukocyte samples. Hazard ratios (HR) for associations between the ratio of both parameters, the "cellular ratio of immune tolerance" (ImmunoCRIT), and cancer risk were estimated using Cox regression models. All statistical tests were two-sided.

Results: ImmunoCRIT values were positively associated with the risk of lung (highest vs. lowest tertile; HR: 1.98, 95% confidence interval: 1.06-3.69; ptrend = 0.03) and colorectal cancer (HR: 1.59, 95% CI: 0.99-2.54; ptrend = 0.007) after multivariable adjustment, but not with prostate cancer risk. Regarding breast cancer significant heterogeneity by estrogen receptor (ER) status was observed (pheterogeneity = 0.02), and the ImmunoCRIT was associated with the risk of ER-negative breast cancer (HR: 3.34, 95% CI: 1.52-7.35; ptrend = < 0.001), but not ER-positive breast cancer.

Conclusions: The present study indicates that increased peripheral immune tolerance may be an independent risk factor for lung, colorectal and ER-negative breast cancer, whereas its role on the development of prostate and ER-positive breast tumors remains uncertain.

Introduction

The ability of tumor cells to evade immune surveillance by suppression of the immune system and alteration of tumor-cell characteristics is a hallmark of carcinogenesis [1, 2]. Evidence in support of the existence of immunological defense mechanisms against cancer in humans is based on the observation that cancer progression and prognosis are related to the functional status of the immune system [3]. Moreover, it has been shown that reconstitution of tumor-specific immune responses are induced by both conventional and targeted anti-cancer therapies and it has therefore been suggested that an intact immune system protects against malignancies [4, 5].

In a healthy person, adaptive immune responses are controlled by a balance between total CD3+ T-lymphocytes (tTLs) - which mainly consist of effector cells driving the elimination of abnormal cells - and suppressor cells, in particular $Foxp3^+$ regulatory T-lymphocytes (Tregs) - which modulate the aggressiveness of the response [6, 7]. Thus, an increased ratio between Tregs and tTLs may facilitate cancer development [8]. Actually, intratumoral accumulation of Tregs frequently correlates with greater tumor aggressiveness in patients affected by various cancer types [9-12]. The notion that the

balance between Tregs and tTLs determines immunity against tumors is further supported by clinical studies on interleukin 2 (IL-2) therapies. Both Tregs and activated effector T-lymphocytes express high levels of Interleukin-2 receptor α chain (IL-2Rα). Hence, IL-2 administration – as used for salvage therapy for patients with refractory malignant melanoma [13] and renal cell carcinoma [14] - leads to expansion of tTLs and Tregs [15]. This parallel expansion serves as explanation for the limited clinical response observed in the majority of patients [16]. Taken together, these findings suggest that Treg accumulation is a dominant mechanism of tumor evasion owing to suppression of tumor-specific effector T lymphocyte responses and development of immune tolerance to malignant cells.

It is still under debate whether a more tolerant microenvironment facilitates early tumor development, if it is simply the consequence of tumor-mediated local enrichment, or both [17]. Uncertainty further exists as to whether immune tolerance is mainly a localized phenomenon, even though increased numbers of Tregs have been found in peripheral blood of patients with several types of cancer, including pancreas, breast, hepatocellular, prostate and lung carcinomas [18-24].

To our knowledge, the concept of Treg-mediated tolerance as a significant barrier to antitumor immunity has not been studied so far in observational studies of initially healthy populations. Here, we report the findings from a case-cohort study within the EPIC-Heidelberg cohort, including incident cases of the four most frequent cancer types (breast, colorectal, lung, and prostate cancer). Epigenetic assays using DNA based qPCR approaches were applied to quantify the ratio of Tregs [25] and tTLs [26] referred

to as the cellular ratio of immune tolerance (“ImmunoCRIT”) after long-term storage [27]. Our prior hypothesis was that an elevated ratio of Tregs-to-tTL in the blood of initially healthy subjects might be associated with an increased cancer risk.

Methods

Study Population

The European Prospective Investigation into Cancer and Nutrition (EPIC) - Heidelberg study was initiated as part of the Europe-wide EPIC project and includes 11929 male and 13611 female participants aged 35 to 65 years that were recruited from the local general population. Baseline examinations were carried out from 1994 through 1998 and included blood sampling (stored in liquid nitrogen), anthropometric measurements and self-administered questionnaires on diet, lifestyle and reproductive health. The study was approved by the ethics committee of the Medical School of the University of Heidelberg and all participants gave written informed consent. [28, 29] Incident cancer cases were ascertained by follow-up questionnaires and by record linkage, and all cases were verified by study physicians based on medical records. Further details on follow-up procedures of EPIC-Heidelberg have been described elsewhere [30].

A case-cohort study embedded in EPIC-Heidelberg was set up for the present project. After exclusion of prevalent cancer cases, the study population comprised 150 cases of lung (ICD-10: C34), 194 cases of colorectal (ICD-10: C18-20), 410 cases of breast (ICD-10: C50), and 394 cases of prostate cancer (ICD-10: C61) that occurred between baseline examination and December 31, 2006 as well as a random sub-cohort of 813 subjects, who had initially been drawn for the Europe-wide EPIC-InterAct case-cohort study [31]. For the primary analysis, study subjects were excluded when there were missing covariate data (cases: n=2 colorectal, n=1 prostate) or quality control in qPCR analysis failed (cases: n=1 lung, n=7 colorectal, n=11 breast, n=15 prostate; non-cases:

n= 6). Thus, statistical analyses were performed on 149, 185, 399, 378 cases of lung, colorectal, breast, and prostate cancer, respectively and 807 sub-cohort members. As a result of random selection, the sub-cohort eventually contained twenty two incident cancer cases (lung: n=4, breast: n=2, prostate: n=16).

Laboratory Methods

Details on storage and processing of blood samples are described in Supplement S2.

For epigenetic analysis, genomic DNA from pre-diagnostic buffy coat specimens was chemically modified by sodium bisulphite. In this reaction, unmethylated cytosine is converted to uracil while methylated cytosine remains unchanged. For each sample, about 1.6 µg DNA was bisulfite converted using the EpiTect Bisulfite Kit (Qiagen) following the manufacturer's protocol.

DNA fragments corresponding to unmethylated, bisulfite converted DNA at the Foxp3, CD3 and GAPDH loci were cloned into vector pUC57 (Genescript Inc.). The resulting 3787bp plasmid was linearized and used for the qPCR reactions detailed below in the following serially diluted final concentrations of 12.97, 2.59, 0.52, 0.1, 0.02 and 0.01 pg/ml yielding 15625, 3135, 625, 125, 25 and 15 plasmid copies. Epigenetic qPCR reactions contained 7.5 pmol forward and reverse primers, 1.25 pmol hydrolysis probe, 1x Roche LightCycler 480 Probes Master and approximately 70 ng bisulfite converted DNA or the above final concentrations of plasmid for standard curve design. Each reaction was performed in a final volume of 5 µl. Tregs, tTLs and all leukocytes (GAPDH) for all samples were analyzed in triplicate using a LightCycler 480 System (Roche). Cycling conditions were: 1 time 95° C preheating for 10 min and 50 cycles of 95 °C for 15 s followed by 1 min at 61° C. Template copy numbers were estimated from

standard curve by linear regression on crossing point (CP) using second derivative maximum method as defined by Roche Light Cycler 480 Software.

The proportion of a specific cell type was determined as follows: Using bisulfite converted DNA as substrate, qPCRs were designed and performed for the selected cell type-specific demethylated loci (Foxp3 and CD3) and for a locus known to be demethylated in all cell types (GAPDH) [26].Then, the ratio of Foxp3 and CD3 values was determined, and is referred to as ImmunoCRIT.

Statistical Analyses

Selected baseline characteristics of cases and sub-cohort members were presented as means ± standard deviations or proportions. ImmunoCRIT values were displayed as medians and interquartile ranges. For analyses on cancer risk, ImmunoCRIT measurements were categorized into tertiles using cut points based upon the distribution in the sub-cohort and subjects in the lowest tertile were considered as the reference group. Prentice-weighted Cox proportional hazards regression models, with age as the underlying time scale, were used to estimate hazard ratios (HRs) and 95% CIs for the association between ImmunoCRIT and cancer risk [32]. The person-time each participant contributed was calculated as difference between age at recruitment and age at diagnosis or age at censoring, i.e., death, loss-to follow-up or censoring at end of the follow-up period, respectively. For each endpoint, 2 regression models are presented. The first model includes age and sex. The second model was then fitted additionally including all those potential confounding variables that changed the risk estimates by more than 10%, or were clearly associated with the exposure; these are indicated in the

footnotes to Table 2. Tests for linear trend were carried out based on continuous values of the ImmunoCRIT on the log2 scale, thus calculating the HR associated with a doubling of the ImmunoCRIT.

For the analyses on breast cancer risk, heterogeneity by estrogen receptor (ER) status was assessed by the Cochran`s Q-statistic test. Multiplicative statistical interactions with risk factors were tested for by including cross-product terms along with the main effect terms into the multivariable adjusted models. Sensitivity analyses were conducted excluding cases that were diagnosed within the first 2 years of follow-up. To examine mid- and long-term partial correlations between ImmunoCRIT values assessed at baseline, after 14 years and after 15 years, Spearman correlation coefficients, adjusted for age at baseline and sex, were calculated within a substudy of 100 subjects. Additional information on these correlation analyses is provided in supplement S1.

All statistical tests were 2-sided and p-values below 0.05 were considered statistically significant. All analyses were performed using SAS 9.3 (SAS Institute, Cary, NC).

Results

Descriptive statistics

As compared to the sub-cohort, subjects who developed cancer were characterized by a higher prevalence of unfavorable lifestyle behaviors, as shown in Table 1. The mean lag time from blood donation to time of diagnosis was 6.3, 6.4, 6.7, and 7.0, respectively, for the cases of breast, colorectal, lung and prostate cancer. The proportions of women among lung cancer cases (30%) and colorectal cancer cases (35%) were smaller than in the sub-cohort (54%). In participants of the sub-cohort, geometric means of ImmunoCRIT were significantly higher in women and ever smokers, with some indication for an increase by both cumulative lifetime smoking history and current smoking status at the time of blood sampling (Supplementary Table S1.1). There were no differences in ImmunoCRIT levels across strata of age, waist circumference, alcohol intake, and current NSAID use. Among women, ImmunoCRIT values did not differ significantly by menopausal status, exogenous hormone use and pregnancy-related factors (e.g. parity, supplementary Table S1.2).

Stability of the ImmunoCRIT over time

Over one year, intra-individual values showed good reproducibility with a Spearman coefficient of 0.67 after adjustment for sex and age. Long-term correlations were also reasonable, both after 14 years (time point 1; r=0.57) and 15 years (time point 2; r=0.52), even though different substrates were used for the latter analyses (buffy coat at baseline vs. PBMC at time points 1 and 2).

ImmunoCRIT values and cancer risk

The distribution of ImmunoCRIT measurements among cases and participants in the sub-cohort is visualized with box plots in Figure 1. Median ImmunoCRIT values were 5.7 % in breast, 5.3 % in colorectal, 5.9 % in lung, and 4.8 % in prostate cancer cases, whereas the sub-cohort showed median values of 5.1 % in total, as well as 5.5 % and 4.7 % among female and male subjects, respectively.

Associations between the ImmunoCRIT and risks of lung, colorectal, breast (overall and by estrogen receptor status), and prostate cancer are presented in Table 3. After multivariable adjustment, Cox regression analyses showed significant positive associations between ImmunoCRIT values and lung cancer risk (highest vs. lowest tertile; HR: 1.98, 95%CI: 1.06–3.69; p_{trend} = 0.03) as well as colorectal cancer risk (HR: 1.59, 95%CI: 0.99–2.54; p_{trend} = 0.007). For colorectal cancer, associations in the crude and multivariable adjusted model were of similar magnitude, whereas associations between the ImmunoCRIT and lung cancer risk were attenuated through adjustment, particularly when smoking was accounted for. There were no associations of the ImmunoCRIT with breast and prostate cancer risk in multivariable models. Sensitivity analyses excluding cases, which occurred within the first 2 years of follow-up, showed no major change in any of the risk estimates, as presented in supplementary Table S2.

There were no significant interactions between the ImmunoCRIT and any of the adjustment factors. However, significant heterogeneity in the associations between ImmunoCRIT and breast cancer risk by ER status (p_{het} = 0.02) was observed. Subgroup analyses by ER status revealed a positive association between the ImmunoCRIT and the risk of ER-negative breast cancer (HR: 3.3, 95% CI: 1.52–7.35; p_{trend} = < 0.001; see

Table 3), but no significant association with the risk of ER-positive breast cancer. A significant direct association of the ImmunoCRIT with breast cancer risk was found among women diagnosed at age less than 50 years (highest vs. lowest tertile; HR: 2.26, 95% CI: 1.12 – 4.46; p_{trend} = 0.04, see supplementary Table S3), but not women older than 50 years.

Discussion

To our knowledge, this prospective study in initially healthy subjects was the first to address the relationship between inter-individual variations in peripheral immune tolerance and cancer risk. We observed that an increased Treg-to-tTL ratio (ImmunoCRIT) was clearly associated with a higher risk of lung and colorectal cancer. Moreover, there was a significant direct association between elevated ImmunoCRIT values and the risk of ER-negative breast cancer. No statistically significant relationships were found with respect to ER-positive breast and prostate cancer.

Overall, our findings are in line with the notion that Treg-mediated immune tolerance plays an important role throughout cancer development. In fact, the observed significant associations were still present after excluding subjects diagnosed with cancer within two years after blood draw, which indicates that increased immune tolerance may facilitate carcinogenesis rather than being a consequence of tumor manifestation. The present study adds to the limited prospective data on the importance of immunological host defense mechanisms against cancer among healthy humans. So far, only one study by Imai et al. reported that natural cytotoxic activity of peripheral blood lymphocytes was inversely associated with overall cancer risk [33].

While the associations observed for lung, colorectal, and ER-negative breast cancer in the present study point to a role of immune tolerance as a global cancer risk factor, explanations for the lack of association between the ImmunoCRIT and the risks of prostate and ER-positive breast cancers are required. One could argue that the strength of association with the ImmunoCRIT varies depending on tumor aggressiveness and/or anatomical tumor localization determining the degree of peripheral interaction. In

agreement with this notion, we have previously found an increase of cellular tolerance in patients with tumors that are often associated with distant, hematogenous metastases, such as lung and colon carcinomas [27]. By contrast, ovarian cancer, which unfolds at an immune privileged site and is not immediately challenged by the peripheral immune system, does not commonly exhibit hematogenous metastases. Here, an increase of peripheral tolerance was not observed at primary diagnosis and occurred only upon prolonged treatment and when patients developed distant metastases outside the abdominal cavity [27]. Those observations led to the assumption that an interaction between tumor microenvironment and peripheral immune system may influence the ability of hematogenous dissemination. Actually, such interpretation is compatible with our finding that the risks of ER-positive breast and prostate tumors, which are less aggressive [34], and have a lesser tendency to form distant metastasis [35], respectively, were not associated with increased ImmunoCRIT. Yet, even though this hypothesis of tumor-specific susceptibility to peripheral immune control in carcinogenesis is appealing, it remains somewhat speculative.

Another finding of the present study that deserves further consideration is the utility of the ImmunoCRIT as a long-term biomarker of immune tolerance, as the balance between Tregs and tTLs may be affected by acute rather than chronic immune activation only. Notably, the good 1-year and 14-year within-subject reproducibility of the ImmunoCRIT suggests that a single measurement may provide a good proxy for long-term values, which is in line with the theory of a tight homeostatic control of Tregs [36].

Assuming that the ratio of peripheral Treg levels and tTLs is rather stable over time, it still appears worthwhile to identify potential modulators of the ImmunoCRIT. In our

study, the only factors which showed significant associations with ImmunoCRIT values were sex and smoking. Heterogeneity of the ImmunoCRIT by sex could be due to the fact that one of the two Foxp3-TSDR alleles, which should be methylated as a result of X-inactivation in women, may be not entirely inactivated [37]. Since methylation of the CD3D gene is not sex-dependent, this may well imply that clinical reference ranges of ImmunoCRIT values should indeed be sex-specific. Nonetheless, our analyses showed no significant heterogeneity of the association between ImmunoCRIT and cancer risk by sex. Our observation of a positive association between ImmunoCRIT values and smoking is in line with previous findings of elevated Treg levels in female smokers [38] and experimental data pointing to comprised immunity induced by smoking [39, 40]. In this context, it is of note that statistical adjustment for smoking led to substantial attenuation of the association between ImmunoCRIT values and lung cancer risk in our study. While we acknowledge that smoking assessment is prone to measurement error, and that residual confounding may have influenced our results on lung cancer risk, it is also plausible that the adverse effect of smoking with respect to lung cancer may in part be mediated by immune suppression [41]. However, associations between ImmunoCRIT and covariates in our study were merely cross-sectional. Therefore, further research on the possible interaction between ImmunoCRIT and lifestyle factors is needed.

The prospective design of this study and the novel epigenetic assay enabling the quantification of immune cells in buffy coat samples after long-term storage provided a unique opportunity to clarify the relationship between Treg-mediated tolerance and cancer risk. The long-term within-subject reproducibility of ImmunoCRIT values was demonstrated based on repeated blood draws, and correlation analyses. While our main analyses on breast, prostate, lung, and colorectal cancer were well-powered, it must be

noticed that our subgroup finding of an association between ImmunoCRIT values and ER-negative breast cancer requires replication in a larger sample.

In summary, the present case-cohort study nested within the EPIC-Heidelberg cohort indicates that an increased ImmunoCRIT may promote early events of carcinogenesis independent from well-established risk factors, at least with regard to colorectal, lung, and ER-negative breast cancer. Overall, our findings imply a role of a positively skewed Treg/T-lymphocyte ratio in suppressing immune surveillance of human carcinomas at selected sites by inducing immune tolerance. Consequently, the reduction of peripheral tolerance might be a promising target for the prevention of cancer and the ImmunoCRIT may serve as a pre-diagnostic biomarker for the identification of individuals at higher cancer risk.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646-74.

2. Cavallo F, De Giovanni C, Nanni P*, et al.* 2011: the immune hallmarks of cancer. Cancer Immunology, Immunotherapy 2011;60(3):319-326.

3. Galon J, Angell HK, Bedognetti D*, et al.* The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. Immunity 2013;39(1):11-26.

4. Zitvogel L, Galluzzi L, Smyth Mark J*, et al.* Mechanism of Action of Conventional and Targeted Anticancer Therapies: Reinstating Immunosurveillance. Immunity 2013;39(1):74-88.

5. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12(4):252-264.

6. Sakaguchi S, Yamaguchi T, Nomura T*, et al.* Regulatory T Cells and Immune Tolerance. Cell 2008;133(5):775-787.

7. Sakaguchi S, Wing K, Yamaguchi T. Dynamics of peripheral tolerance and immune regulation mediated by Treg. Eur J Immunol 2009;39(9):2331-6.

8. Yamaguchi T, Sakaguchi S. Regulatory T cells in immune surveillance and treatment of cancer. Seminars in Cancer Biology 2006;16(2):115-123.

9. Bates GJ, Fox SB, Han C*, et al.* Quantification of regulatory T cells enables the identification of high-risk breast cancer patients and those at risk of late relapse. J Clin Oncol 2006;24(34):5373-80.

10. Curiel TJ, Coukos G, Zou L*, et al.* Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. Nat Med 2004;10(9):942-9.

11. Flammiger A, Weisbach L, Huland H*, et al.* High tissue density of FOXP3+ T cells is associated with clinical outcome in prostate cancer. Eur J Cancer 2013;49(6):1273-9.

12. Suzuki H, Chikazawa N, Tasaka T*, et al.* Intratumoral CD8(+) T/FOXP3 (+) cell ratio is a predictive marker for survival in patients with colorectal cancer. Cancer Immunol Immunother 2010;59(5):653-61.

13. Atkins MB, Lotze MT, Dutcher JP*, et al.* High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: analysis of 270 patients treated between 1985 and 1993. J Clin Oncol 1999;17(7):2105-16.

14. Rosenberg SA. Interleukin 2 for patients with renal cancer. Nat Clin Pract Oncol 2007;4(9):497.

15. Ahmadzadeh M, Rosenberg SA. IL-2 administration increases CD4+ CD25(hi) Foxp3+ regulatory T cells in cancer patients. Blood 2006;107(6):2409-14.

16. Jacobs JF, Nierkens S, Figdor CG*, et al.* Regulatory T cells in melanoma: the final hurdle towards effective immunotherapy? Lancet Oncol 2012;13(1):e32-42.

17. Savage PA, Malchow S, Leventhal DS. Basic principles of tumor-associated regulatory T cell biology. Trends Immunol 2013;34(1):33-40.

18. Kono K, Kawaida H, Takahashi A*, et al.* CD4(+)CD25high regulatory T cells increase with tumor stage in patients with gastric and esophageal cancers. Cancer Immunol Immunother 2006;55(9):1064-71.

19. Hiraoka N, Onozato K, Kosuge T*, et al.* Prevalence of FOXP3+ regulatory T cells increases during the progression of pancreatic ductal adenocarcinoma and its premalignant lesions. Clin Cancer Res 2006;12(18):5423-34.

20. Miller AM, Lundberg K, Ozenci V*, et al.* CD4+CD25high T cells are enriched in the tumor and peripheral blood of prostate cancer patients. J Immunol 2006;177(10):7398-405.

21. Liyanage UK, Goedegebuure PS, Moore TT*, et al.* Increased prevalence of regulatory T cells (Treg) is induced by pancreas adenocarcinoma. J Immunother 2006;29(4):416-24.

22. Ormandy LA, Hillemann T, Wedemeyer H*, et al.* Increased populations of regulatory T cells in peripheral blood of patients with hepatocellular carcinoma. Cancer Res 2005;65(6):2457-64.

23. Hasegawa T, Suzuki H, Yamaura T*, et al.* Prognostic value of peripheral and local forkhead box P3 regulatory T cells in patients with non-small-cell lung cancer. Mol Clin Oncol 2014;2(5):685-694.

24. Ling KL, Pratap SE, Bates GJ*, et al.* Increased frequency of regulatory T cells in peripheral blood and tumour infiltrating lymphocytes in colorectal cancer patients. Cancer Immun 2007;7:7.

25. Wieczorek G, Asemissen A, Model F*, et al.* Quantitative DNA methylation analysis of FOXP3 as a new method for counting regulatory T cells in peripheral blood and solid tissue. Cancer Res 2009;69(2):599-608.

26. Sehouli J, Loddenkemper C, Cornu T*, et al.* Epigenetic quantification of tumor-infiltrating T-lymphocytes. Epigenetics 2011;6(2):236-46.

27. Turbachova I, Schwachula T, Vasconcelos I*, et al.* The cellular ratio of immune tolerance (immunoCRIT) is a definite marker for aggressiveness of solid tumors and may explain tumor dissemination patterns. Epigenetics 2013;8(11):1226-35.

28. Riboli E, Hunt KJ, Slimani N*, et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5(6b):1113-24.

29. Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43(4):205-15.

30. Bergmann MM, Bussas U, Boeing H. Follow-up procedures in EPIC-Germany--data quality aspects. European Prospective Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43(4):225-34.

31. Langenberg C, Sharp S, Forouhi NG*, et al.* Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. Diabetologia 2011;54(9):2272-82.

32. PRENTICE RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73(1):1-11.

33. Imai K, Matsuyama S, Miyake S*, et al.* Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. Lancet 2000;356(9244):1795-9.

34. Alqaisi A, Chen L, Romond E*, et al.* Impact of estrogen receptor (ER) and human epidermal growth factor receptor-2 (HER2) co-expression on breast cancer disease characteristics: implications for tumor biology and research. Breast Cancer Res Treat 2014; 10.1007/s10549-014-3145-x.

35. Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. Int J Cancer 2000;85(1):60-7.

36. Liston A, Gray DH. Homeostatic control of regulatory T cell diversity. Nat Rev Immunol 2014;14(3):154-65.

37. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 2005;434(7031):400-4.

38. Hampras SS, Nesline M, Wallace PK*, et al.* Predictors of immunosuppressive regulatory T lymphocytes in healthy women. J Cancer Epidemiol 2012;2012:191090.

39. Sopori M. Effects of cigarette smoke on the immune system. Nat Rev Immunol 2002;2(5):372-7.

40. Hernandez CP, Morrow K, Velasco C*, et al.* Effects of cigarette smoke extract on primary activated T cells. Cell Immunol 2013;282(1):38-43.

41. Stampfli MR, Anderson GP. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. Nat Rev Immunol 2009;9(5):377-84.

Figures

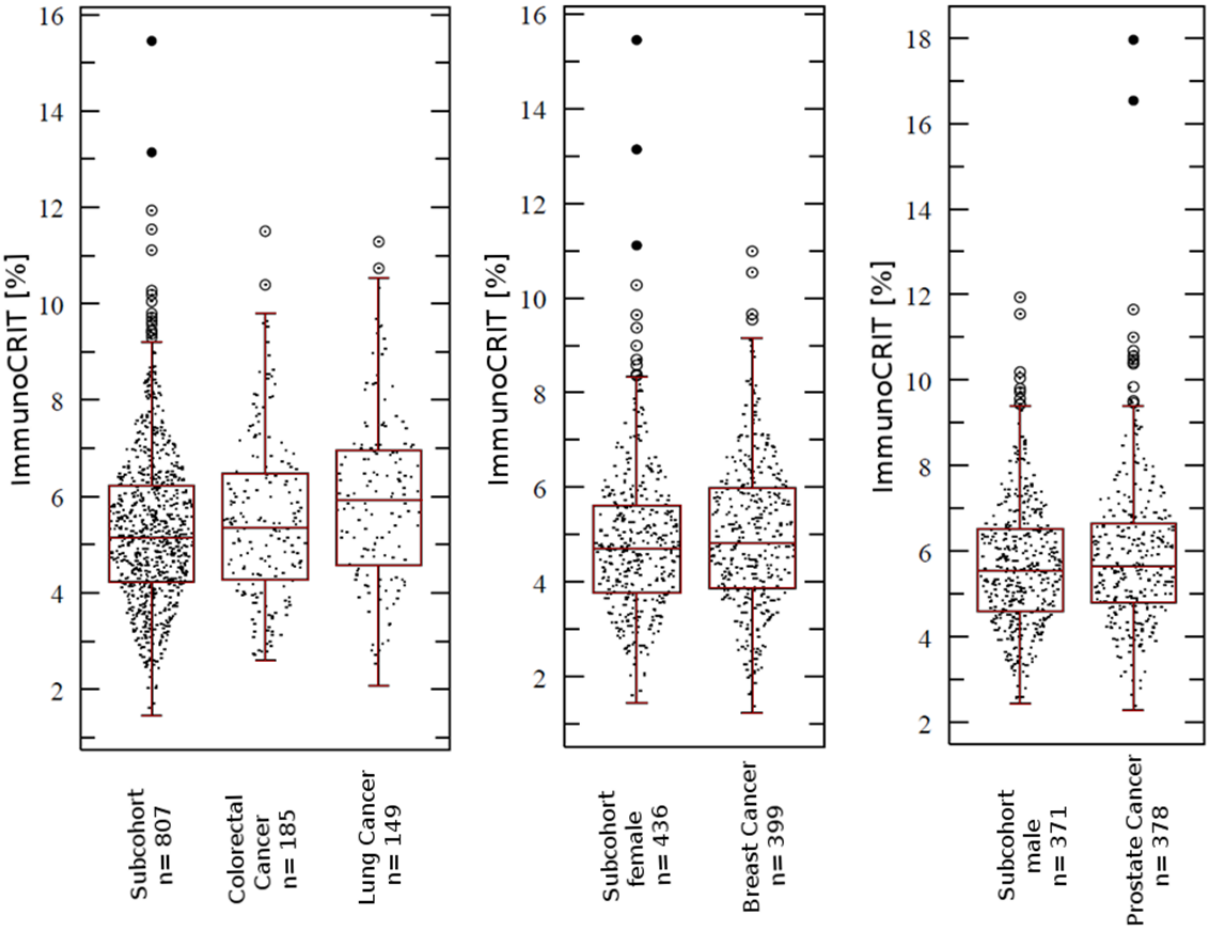

Figure 1: Boxplots showing ImmunoCRIT data by cancer type

Tables

Table 1 Characteristics and laboratory results of the study population

	Incident cancer cases				Sub-cohort		
	Lung	**Colorectum**	**Breast**	**Prostate**	**Total**	**Women**	**Men**
N	149	185	399	378	807	436	371
Socio-demographic factors							
Women, %	30	35	100	-	54	100	-
Age at baseline, years	55.1± 7.4	56.1 ± 6.3	51.6± 7.8	57.7± 5.4	50.7± 8.0	49.1± 8.4	52.5± 7.0
Education level, %							
Primary	46	34	27	35	25	25	25
Secondary	39	36	47	31	43	52	34
University	15	30	26	34	32	23	41
Case characteristics							
Age at diagnosis, years	61.8± 7.5	62.4± 6.8	57.8± 7.8	64.7± 5.4	-	-	-
Stage at diagnosis, %							
Local	19	39	61	71	-	-	-
Regional	26	41	34	24	-	-	-
Distant	45	19	2	4	-	-	-
Unknown	10	1	3	1	-	-	-
Lifestyle factors, %							
Abdominal adiposity *	33	41	26	26	23	23	24
Physically Inactive †	52	50	48	47	44	45	43
Ever smokers	91	67	44	58	57	50	66
Heavy drinkers ‡	39	46	34	44	35	31	39
Laboratory measurements, median (IQR)							
% ImmunoCRIT	5.9 (2.1,11.3)	5.3 (2.6,11.5)	5.7 (2.3,18.0)	4.8 (1.2,11.0)	5.1 (1.4,15.5)	5.5 (2.4,11.9)	4.7 (1.4,15.5)

Values are means ± standard deviations or proportions, unless otherwise stated. IQR: Interquartile range

Education data missing for one breast cancer case. Smoking data missing for one colorectal as well as one breast cancer case and three sub-cohort members.

* defined by waist circumference ≥ 102cm for men and ≥ 88 cm for women according to WHO cut-offs.

† summary variable for inactive and moderately inactive

‡ defined by alcohol intake at baseline >24g/d for men and >12 g/d for women

Table 2: Hazard ratios and 95% confidence intervals [HR (95% CI)] of solid cancers across tertiles of ImmunoCRIT

	Tertiles*			HR (95% CI) $_{log2}$	p_{trend}
	1 (referent)	2	3		
Lung cancer					
Tertile median (range)	3.9 (1.4,4.6)	5.1 (4.6,5.8)	6.7 (5.8,15.5)		
N Cases / Sub-cohort†	35 / 265	36 / 265	78 / 264		
Model 1	1.00	1.43 (0.82,2.51)	3.45 (1.97,6.04)	3.44 (2.11,5.62)	<.0001
Model 2	1.00	0.99 (0.52,1.89)	1.98 (1.06,3.69)	1.95 (1.08, 3.52)	0.0263
Colorectal cancer					
Tertile median (range)	3.9 (1.4,4.6)	5.1 (4.6,5.8)	6.6 (5.8,15.5)		
N Cases / Sub-cohort†	59 / 265	55 / 264	71 / 264		
Model 1	1.00	1.31 (0.85,2.02)	1.70 (1.1,2.64)	1.81 (1.22, 2.70)	0.0035
Model 2	1.00	1.32 (0.83,2.11)	1.59 (0.99,2.54)	1.81 (1.18, 2.77)	0.0069
Breast cancer					
Tertile median (range)	4.3 (2.4,4.9)	5.5 (5.0,6.2)	7.1 (6.2,11.9)		
N Cases / Sub-cohort†	117 / 145	129 / 145	153 / 145		
Model 1	1.00	1.10 (0.76,1.59)	1.23 (0.86,1.75)	1.47 (0.99, 2.19)	0.056
Model 2	1.00	1.03 (0.69,1.53)	1.11 (0.77,1.61)	1.34 (0.90, 2.01)	0.162
ER-positive breast cancer‡					
N Cases / Sub-cohort†	94 / 144	98 / 146	115 / 144		
Model 1	1.00	1.00 (0.67,1.50)	1.10 (0.75,1.62)	1.30 (0.84, 2.02)	0.242
Model 2	1.00	0.91 (0.59,1.41)	0.99 (0.67,1.46)	1.18 (0.75,1.85)	0.481
ER-negative breast cancer‡					
N Cases / Sub-cohort†	14 / 144	23 / 145	34 / 144		
Model 1	1.00	1.99 (0.89,4.44)	3.09 (1.46,6.55)	3.44 (1.76, 6.72)	0.0003
Model 2	1.00	2.28 (1.05,4.93)	3.34 (1.52,7.35)	3.73 (1.80,7.73)	0.0004
Prostate cancer					
Tertile median (range)	3.4 (1.4,4.2)	4.7 (4.2,5.3)	6.1 (5.3,15.5)		
N Cases / Sub-cohort†	124 / 122	102 / 123	152 / 122		
Model 1	1.00	0.92 (0.60,1.40)	1.11 (0.75,1.66)	1.01 (0.71, 1.45)	0.95
Model 2	1.00	0.90 (0.58,1.39)	1.39 (0.91,2.13)	1.17 (0.80, 1.70)	0.44

* Tertile cut-points were based on the distribution in the sub-cohort. † Exclusions due to missing covariates for cases/sub-cohort: lung cancer (0/13), colorectal cancer (2/14), overall breast cancer (0/1), prostate cancer (1/4). ‡ Receptor status could not be determined for n=21 breast cancers.

Model 1 was stratified by age at recruitment (in 1-year integers) and adjusted for sex if appropriate.

Model 2 was further adjusted for: *Lung cancer:* smoking status (never smokers, former smokers ≥10 years, former smokers <10 years, smokers <15 cigarettes/day, smokers ≥ 15 cigarettes/day), smoking duration (years), NSAID use (yes/no), history of myocardial infarction or stroke (yes/no), red meat consumption (g/day) and height (cm); *Colorectal cancer*: waist circumference (cm), alcohol intake (g/d), processed meat consumption (g/d), hyperlipidemia (yes/no), height (cm), smoking status (see above) and smoking duration (years); *Breast cancer*: exogenous hormone use (OC/HRT; yes/no), height (cm), menopausal status (pre-,peri-,postmenopausal including surgical hysterectomy), and NSAID use (yes/no); *ER-positive*: see overall breast cancer model; *ER-negative*: menopausal status (see above), exogenous hormone use (yes/no) and time between menarche and first birth (years); *Prostate cancer*: smoking status (see above), smoking duration (years), calcium intake (mg/day) and wholegrain intake (g/day).

Supporting information

Supplement S1: Description of reproducibility study

Supplement S2: Sample storage and processing

Table S1.1 & S.1.2: Geometric mean of ImmunoCRIT by common cancer risk factors

Table S2: Sensitivity analyses excluding cases that were diagnosed within 2 years from blood draw

Table S3: Breast cancer risk stratified by age at diagnosis

Table S4. Estrogen and progesterone receptor status of breast cancer cases and their corresponding ImmunoCRIT values

Supplement S1: Description of reproducibility study

Between 2011 and 2012, around 14 years after baseline, a sub study of EPIC-Heidelberg on diet, physical activity, and body composition as assessed by magnet resonance imaging (MRI) was initiated. Participants were re-invited according to a rectangular sampling scheme with the aim to recruit 300 men and 300 women equally distributed over three 10-year categories of baseline age (35-44 years, 45-54 years and 55-64 years). Subjects with contraindications to MRI (metal implants, defibrillators, stents, subcutaneous chips, tattoos, dementia, hemophilia, claustrophobia, BMI greater than 42, dialysis, diagnosis of a serious diseases in the past 12 months, and pregnant women) were not eligible for the sub study. Overall, 613 subjects attended the MRI examination and 592 provided a blood sample. The participation rate was 47%. Between 2012 and 2013, i.e. 1 year after the first re-invitation and around 15 years after baseline, all sub study participants were re-invited to a second but shorter re-examination during which another blood sample was drawn. Finally, a blood sample from both re-examinations was available for 592 subjects (50% women).

Out of these, the first 50 men and women who attended the second re-examination were selected for the reproducibility study. Blood samples from both re-examinations during the sub study were processed (plasma, serum, peripheral blood mononuclear cells [PBMC], erythrocytes) and stored in freezers at -80^0 Celsius.

Supplement S2: Sample storage and processing

Preparation of samples and of genomic and bisulfite converted DNA

At study baseline, ten milliliters of blood were centrifuged in anticoagulant-containing Monovettes at room temperature and 1500xg for 20 minutes. Buffy coats were separated from the interphase and aliquoted into 500 µl portions which were stored under liquid nitrogen in the EPIC-Heidelberg biobank. DNA was isolated by LGC Limited, (Hoddesdon, UK). DNA solutions for each subject were returned in 2D barcoded tubes on dry ice and stored at -80°C until DNA analyses took place. DNA concentration and quality was measured using Quant-iT PicoGreen dsDNA Assay (Life Technologies, Darmstadt, Germany).

Supplementary Table 1.1: ImmunoCRIT values across strata of baseline characteristics in the sub-cohort (n = 807)

Characteristic		N	Mean† (95% CI)	P_{trend} ‡
Sex [a]	Male	371	4.6 (4.5-4.7)	**<0.0001**
	Female	436	5.5 (5.4-5.7)	
Age at blood donation [b]	< 46	223	4.9 (4.7-5.1)	0.17
	46-55	314	5.1 (4.9-5.2)	
	>55	270	5.1 (4.9-5.3)	
Highest level of education	Primary	203	4.9 (4.7-5.1)	0.88
	Secondary	350	5.2 (5.0-5.3)	
	University	254	4.9 (4.8-5.1)	
Waist circumference (WHO)	<94 cm (M); <80 cm (W)	404	5.1 (4.9-5.2)	0.91
	≥94 cm (M); ≥80 cm (W)	216	4.9 (4.7-5.1)	
	≥102 cm (M); ≥88 cm (W)	187	5.1 (4.9-5.3)	
Smoking status	Never smokers	344	4.9 (4.8-5.1)	**<0.0001**
	Long-term quitters (≥10 years)	179	4.8 (4.6-5.0)	
	Short-term quitters (< 10 years)	88	5.2 (4.9-5.5)	
	Light smokers (< 15 cig/d)	80	5.2 (4.9-5.6)	
	Heavy smokers (≥ 15 cig/d)	103	5.6 (5.3-6.0)	
Alcohol intake [c]	None- or light drinkers	365	5.0 (4.9-5.2)	0.72
	Moderate drinkers	163	5.0 (4.8-5.3)	
	Heavy drinkers	279	5.1 (4.9-5.2)	
Physical activity	Inactive	89	5.0 (4.7-5.4)	0.99
	Moderately Inactive	269	5.1 (4.9-5.3)	
	Moderately Active	234	4.9 (4.8-5.1)	
	Active	215	5.1 (4.9-5.3)	
NSAID use	No	733	5.0 (4.9-5.1)	0.11
	Yes	74	5.3 (5.0-5.7)	
History of CVD [d]	No	791	5.0 (4.9-5.1)	0.26
	Yes	16	5.5 (4.7-6.3)	
Menopausal status [a]	Premenopausal	210	5.6 (5.4-5.9)	0.2
	Perimenopausal	66	5.5 (5.2-5.9)	
	Postmenopausal	160	5.3 (5.0-5.6)	
Use of Exogenous hormones [a]	No	316	5.6 (5.4-5.7)	
	Yes	119	5.3 (5.0-5.5)	0.09

† ImmunoCRIT values are geometric means (5-95% percentile range) adjusted for sex and age;

‡ Differences in geometric means between categories were tested for by using Generalized Linear Models;

Data on smoking status and exogenous hormone use are missing for 13 and 1 subject, respectively;

[a] Value is not adjusted for sex. [b] Value is not adjusted for age.

[c] Sex-specific categories of alcohol consumption were created. Men: Non- or light drinkers (<12g/d; reference), moderate drinkers (>12-24 g/d), and heavy drinkers (>24 g/d); Women: non- or light drinkers (<6 g/d; reference), moderate drinkers (>6-12 g/d), and heavy drinkers (>12 g/d).

[d] Cardiovascular diseases (CVD) comprise myocardial infarction and stroke.

Supplementary Table S1.2: ImmunoCRIT values across strata of selected, further reproductive factors of the female sub-cohort (n = 436), adjusted for age and smoking status

Characteristic		N	ImmunoCRIT% Mean† (95% CI)	P_{trend} ‡
Age at menarche (years)	< 12	162	5.6 (5.3-5.8)	0.12
	13	119	5.6 (5.3-5.9)	
	14	95	5.7 (5.4-6.1)	
	≥ 15	59	5.2 (4.8-5.5)	
Age at menopause (years) [a]	< 50	70	5.5 (5.1-6)	0.37
	≥ 50	57	5.8 (5.3-6.3)	
Number of full term pregnancies	Nulliparous	81	5.5 (5.2-5.9)	0.53
	1	97	5.6 (5.3-5.9)	
	2	184	5.5 (5.3-5.8)	
	≥ 3	73	5.7 (5.4-6.1)	
Age at first birth (years) [b]	< 23	86	5.4 (5.1-5.8)	0.67
	23-25	89	5.7 (5.4-6.1)	
	26-28	74	5.4 (5.2-5.9)	
	≥ 29	103	5.6 (5.3-5.9)	
Median time between menarche and first full-term childbirth (years) [b]	<12	146	5.5 (5.2-5.7)	0.25
	≥ 12	208	5.6 (5.4-5.9)	
Breastfeeding [b]	Yes	286	5.6 (5.4-5.7)	0.72
	No	64	5.6 (5.3-6.0)	
Ever used oral contraceptives [c]	No	78	5.3 (5-5.7)	0.15
	Yes	357	5.6 (5.4-5.8)	
Ever used postmenopausal hormones [d]	No	79	5.5 (5.2-5.9)	0.62
	Yes	129	5.6 (5.3-5.9)	
Menstrual cycle phase [c]	follicular (days 0-11)	80	5.4 (5.1-5.8)	0.69
	periovulatory (days 12-16)	41	5.5 (5.1-6.0)	
	luteal (days 17-39)	69	5.4 (5.1-5.8)	
	unknown	20	5.8 (5.2-6.5)	

† ImmunoCRIT values are geometric means (5-95% percentile range).

‡ Differences in geometric means between categories were tested for by using generalized linear models, using linear trend tests if appropriate.

Number of missing values dependent on covariates: age at menarche (n=1), age at menopause (n=33), parity (n=1), age at first birth (n=1), breastfeeding (n=4), ever used oral contraceptives (n=1), ever used postmenopausal hormones (n=18).

a among postmenopausal women only (n=160).

b among parous women only (n=354).

c among premenopausal women only (n=210).

d among peri- and postmenopausal women only (n=226).

Supplementary Table S2: Adjusted Hazard ratios and 95% confidence intervals [HR (95% CI)] of solid cancers across tertiles of ImmunoCRIT, after exclusion of cases diagnosed within 2 years from blood draw

	Tertiles*			HR (95% CI) $_{log2}$	p_{trend}
	1 (referent)	2	3		
Lung					
N cases ≤ 2years	3	5	10		
N Cases / Sub-cohort	32/265	31/265	68/264		
Model 2	1.00	0.90 (0.46,1.74)	1.87 (0.99,3.55)	1.97 (1.08, 3.58)	0.0265
Colorectum					
N cases ≤ 2years	10	6	6		
N Cases / Sub-cohort	49/265	49/264	65/264		
Model 2	1.00	1.52 (0.93,2.49)	1.87 (1.13,3.09)	2.05 (1.30, 3.22)	0.0019
Breast					
N cases ≤ 2years	18	19	23		
N Cases / Sub-cohort	99/145	110/145	130/145		
Model 2	1.00	1.03 (0.67,1.59)	1.14 (0.77,1.69)	1.35 (0.88, 2.07)	0.1673
ER-positive breast cancer †					
N cases ≤ 2years	14	11	12		
N Cases / Sub-cohort	80/144	87/146	103/144		
Model 2	1.00	0.95(0.6,1.51)	1.06(0.7,1.61)	1.12 (0.71, 1.77)	0.6378
ER-negative breast cancer †					
N cases ≤ 2years	0	6	8		
N Cases / Sub-cohort	14/144	17/145	26/144		
Model 2	1.00	1.66(0.73,3.79)	2.33(1.04,5.21)	3.28 (1.47, 7.33)	0.0037
Prostate					
N cases ≤ 2years	10	10	15		
N Cases / Sub-cohort	114/122	92/123	137/122		
Model 2	1.00	0.90 (0.58,1.38)	1.36 (0.88,2.1)	1.16 (0.79, 1.70)	0.4428

* Tertile cut-points were based on the distribution in the sub-cohort (excluding early cases). † Receptor status could not be determined for n=21 breast cancers.

Number of incident cases diagnosed within 2 years from blood collection:

Lung (n=18), Colon (n=22), Breast (n=60; 37 ER+ | 14 ER- | 9 unknown), Prostate (n=35)

Adjustment in Model 2 retained corresponding adjustment factors reported in Table 2.

Supplementary Table S3: Hazard ratios and 95% confidence intervals [HR (95% CI)] of breast cancer risk across tertiles of ImmunoCRIT, stratified by age at diagnosis

	Tertiles; median (range)				
	1 (referent)	**2**	**3**		
Age at diagnosis†	4.3 (2.4,4.9)	5.5 (5.0,6.2)	7.1 (6.2,11.9)	HR (95% CI) $_{log2}$	p_{trend}
< 50 years (n = 71)					
N Cases / Sub-cohort*	17/144	19/145	35/144		
Model 1	1.00	1.05(0.5,2.19)	2.37(1.2,4.68)	2.36 (1.11, 5.06)	0.0265
Model 2	1.00	0.95(0.42,2.12)	2.26(1.12,4.56)	2.27 (1.04, 4.95)	0.0388
≥ 50 years (n = 328)					
N Cases / Sub-cohort	100/145	109/145	119/145		
Model 1	1.00	1.11 (0.73,1.69)	1.06 (0.71,1.58)	1.31 (0.83, 2.07)	0.2467
Model 2	1.00	1.03 (0.66,1.62)	0.92 (0.60,1.40)	1.16 (0.72, 1.85)	0.5445

* 2 breast cancer cases with age at diagnosis **≥** 50 years were excluded from the sub-cohort.

† Distribution of ER-status by age at diagnosis: < 50 (24% ER^-, 76% ER^+), **≥** 50(18% ER^-, 82% ER^+).

Model 1 was stratified by age at recruitment (in 1-year categories).

Model 2 was further adjusted for: exogenous hormone use (yes/no), height (cm), menopausal status (pre-,peri-,postmenopausal including surgical hysterectomy), and NSAID use (yes/no).

For analysis of subgroup "< 50" a simplified model was used including only those covariates yielding the best model fit (AIC); thus adjusting for height (cm), exogenous hormone use (yes/no) and NSAID use (yes/no).

Supplementary Table S4:
Estrogen and progesterone receptor status of breast cancer cases (n=399) and their corresponding ImmunoCRIT values

Hormone receptor status	N	ImmunoCRIT±
ER^+/PR^+ , ER^+/PR^-	307	5.8 (5.6-6.0)
ER^-/PR^-, ER^-/PR^+	71	6.2 (5.8-6.6)
Unknown	21	5.4 (4.8-6.2)
ER^+/PR^+	269	5.8 (5.6-6.0)
ER^+/PR^-	38	5.7 (5.2-6.2)
ER^-/PR^-	63	6.2 (5.7-6.6)
ER^-/PR^+	8	6.5 (5.4-8.0)

±Values are geometric means (95% CI) adjusted for age and smoking status.

Funding

The EPIC-Heidelberg study is supported by the German Federal Ministry of Education and Research (BMBF), Grant number 01ER0809.