

# Lag Times between Lymphoproliferative Disorder and Clinical Diagnosis of Chronic Lymphocytic Leukemia: A Prospective Analysis Using Plasma Soluble CD23

Rudolf Kaaks<sup>1</sup>, Disorn Sookthai<sup>1</sup>, Anna Łuczyńska<sup>2</sup>, Christopher C. Oakes<sup>3</sup>, Susen Becker<sup>1,4,5</sup>, Theron Johnson<sup>1</sup>, Annsofie Johansson<sup>6</sup>, Beatrice Melin<sup>6</sup>, Klas Sjöberg<sup>7</sup>, Dimitrios Trichopoulos<sup>8,9,10,†</sup>, Antonia Trichopoulou<sup>8,10</sup>, Pagona Lagiou<sup>9,10,11</sup>, Amalia Mattiello<sup>12</sup>, Rosario Tumino<sup>13</sup>, Giovanna Masala<sup>14</sup>, Claudia Agnoli<sup>15</sup>, Heiner Boeing<sup>16</sup>, Krasimira Aleksandrova<sup>16</sup>, Paul Brennan<sup>17</sup>, Silvia Franceschi<sup>18</sup>, Sandrine Roulland<sup>19</sup>, Delphine Casabonne<sup>20,21</sup>, Silvia de Sanjose<sup>20,21</sup>, María-José Sánchez<sup>21,22</sup>, José María Huerta<sup>21,23</sup>, Eva Ardanaz<sup>21,24</sup>, Nuria Sala<sup>25</sup>, Kim Overvad<sup>26</sup>, Anne Tjønneland<sup>27</sup>, Jytte Halkjær<sup>27</sup>, Elisabete Weiderpass<sup>28,29,30,31</sup>, H. B(as). Bueno-de-Mesquita<sup>32,33,34</sup>, Roel Vermeulen<sup>35</sup>, Petra H. Peeters<sup>34,36</sup>, Paolo Vineis<sup>34,37,38</sup>, Rachel S. Kelly<sup>38</sup>, Kay-Tee Khaw<sup>39</sup>, Ruth C. Travis<sup>40</sup>, Timothy J. Key<sup>40</sup>, Elio Riboli<sup>34</sup>, and Alexandra Nieters<sup>2</sup>

## Abstract

**Background:** Chronic lymphocytic leukemia (CLL) is a chronic disease that often progresses slowly from a precursor stage, monoclonal B-cell lymphocytosis (MBL), and that can remain undiagnosed for a long time.

**Methods:** Within the European Prospective Investigation into Cancer cohort, we measured prediagnostic plasma sCD23 for 179 individuals who eventually were diagnosed with CLL and an equal number of matched control subjects who remained free of cancer.

**Results:** In a very large proportion of CLL patients' plasma sCD23 was clearly elevated 7 or more years before diagnosis. Considering sCD23 as a disease predictor, the area under the ROC curve (AUROC) was 0.95 [95% confidence interval (CI), 0.90–1.00] for CLL diagnosed within 0.1 to 2.7 years after blood

measurement, 0.90 (95% CI, 0.86–0.95) for diagnosis within 2.8 to 7.3 years, and 0.76 (95% CI, 0.65–0.86) for CLL diagnosed between 7.4 and 12.5 years. Even at a 7.4-year and longer time interval, elevated plasma sCD23 could predict a later clinical diagnosis of CLL with 100% specificity at >45% sensitivity.

**Conclusions:** Our findings provide unique documentation for the very long latency times during which measurable B-cell lymphoproliferative disorder exists before the clinical manifestation of CLL.

**Impact:** Our findings have relevance for the interpretation of prospective epidemiologic studies on the causes of CLL in terms of reverse causation bias. The lag times indicate a time frame within which an early detection of CLL would be theoretically possible.

*Cancer Epidemiol Biomarkers Prev*; 24(3); 538–45. ©2014 AACR.

<sup>1</sup>Division of Cancer Epidemiology, German Cancer Research Center Heidelberg, Heidelberg, Germany. <sup>2</sup>Center for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Germany. <sup>3</sup>Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center, Heidelberg, Germany. <sup>4</sup>Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Leipzig, Leipzig, Germany. <sup>5</sup>Leipzig Research Center for Civilization Diseases (LIFE), University Leipzig, Leipzig, Germany. <sup>6</sup>Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden. <sup>7</sup>Department of Gastroenterology and Nutrition, Malmö, Skåne University Hospital, Lund University, Lund, Sweden. <sup>8</sup>Hellenic Health Foundation, Athens, Greece. <sup>9</sup>Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts. <sup>10</sup>Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece. <sup>11</sup>Department of Hygiene, Epidemiology, and Medical Statistics, University of Athens Medical School, Goudi, Athens, Greece. <sup>12</sup>Dipartimento di Medicina Clinica e Chirurgia, Federico II, University, Naples, Italy. <sup>13</sup>Cancer Registry and Histopathology Unit, "Civic—M.P. Arezzo" Hospital, ASP Ragusa, Italy. <sup>14</sup>Molecular and Nutritional Epidemiology Unit, Cancer Research and Prevention Institute—ISPO, Florence, Italy.

<sup>15</sup>Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori Via Venezian, Milano, Italy. <sup>16</sup>Department of Epidemiology, German Institute of Human Nutrition (DIfE) Potsdam-Rehbrücke, Nuthetal, Germany. <sup>17</sup>Section of Genetics, International Agency for Research on Cancer (IARC), Lyon, France. <sup>18</sup>Infections and Cancer Epidemiology Group, International Agency for Research on Cancer, Lyon, France. <sup>19</sup>Centre d'Immunologie de Marseille-Luminy, Aix-Marseille Université, Marseille, France. <sup>20</sup>Unit of Infections and Cancer (UNIC), IDIBELL, Institut Català d'Oncologia, L'Hospitalet de Llobregat, Barcelona, Spain. <sup>21</sup>Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. <sup>22</sup>Escuela Andaluza de Salud Pública Instituto de Investigación Biosanitaria de Granada (Granada. ibs), Granada, Spain. <sup>23</sup>Department of Epidemiology, Murcia Regional Health Council, Murcia, Spain. <sup>24</sup>Navarre Public Health Institute, Pamplona, Spain. <sup>25</sup>Translational Research Laboratory and Unit of Nutrition, Environment, and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO), Biomedical Research Institute (IDIBELL), Barcelona, Spain. <sup>26</sup>Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark. <sup>27</sup>Danish Cancer Society Research Center, Copenhagen, Denmark.

## Introduction

B-cell chronic lymphocytic leukemia (CLL), a specific form of non-Hodgkin lymphoma (NHL), is characterized as a clonal expansion of abnormal B lymphocytes in peripheral blood, bone marrow, and lymphoid tissues. Two landmark studies (1, 2) showed that all CLL develop through a precursor stage of monoclonal B-cell lymphocytosis (MBL)—an asymptomatic condition in which small numbers of clonal B cells are detectable in blood, and which affects more than 3% of individuals of age 50 and older (3). The majority of MBL share with CLL a common immunophenotype, including coexpression of the cluster of differentiation antigens CD5 and CD23 on the cell surface, close to undetectable levels of surface immunoglobulin, and a number of cytogenetic abnormalities (4, 5). MBL is differentiated from CLL by an absolute number of monoclonal B cells with a CLL phenotype less than 5,000/mL in the peripheral blood (6).

Considerable research is being directed toward the identification of phenotypic characteristics of MBL that can reliably predict the progression toward clinical CLL requiring medical treatment (4, 7). One established risk factor for progression of MBL into clinical CLL is the size of the abnormal B-cell clone, which generally correlates strongly with absolute B lymphocyte counts. Low-count MBL, where absolute B lymphocyte numbers are not increased much, but where a few monoclonal, CLL-like B cells (<100 cells/mL) are detectable, does not appear to represent a true preleukemic condition. By contrast, high-count MBL, in which typically 500 to 5,000 neoplastic B cells per microliter are counted (i.e., corresponding to a genuine condition of lymphocytosis), does progress at an estimated rate of around 1% to 2% per year to a stage of clinical CLL (1, 8, 9).

Typically, in clinical CLL cell surface concentrations of CD23 are strongly increased in comparison with normal B cells, as a result of abnormal upregulation of the CD23 gene (10). CD23, is a transmembrane glycoprotein corresponding to the low-affinity receptor for IgE and plays an important role in the capture and processing of antigen complexed with IgE and regulation of IgE responses (11, 12). CD23 can be cleaved from cell surfaces to yield soluble CD23 (sCD23) proteins that are released into extracellular fluids. In patients with CLL plasma sCD23 levels are highly (3- to 500-fold) increased compared with levels among healthy control subjects (10–13). Furthermore, among CLL patients higher plasma levels of sCD23 correlate with more advanced disease stage, shorter doubling times of lymphocytes, and poorer prognosis in terms of expected survival time (13, 14), and sCD23 levels are also a reflection of tumor mass and the size of the CD23<sup>+</sup> B-cell pool (15, 16).

In the present article, we present results from a large, population-based cohort of individuals initially free of cancer symptoms—the European Prospective Investigation into Cancer (EPIC)—on the longitudinal relationship of plasma sCD23 levels with the subsequent probability of being clinically diagnosed with CLL, within a prospective follow-up time ranging from 0.06 to 12.5 years.

## Materials and Methods

EPIC is a large prospective cohort study with more than 500,000 participants enrolled in 23 centres in Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom (17). Between 1992 and 1998, subjects were mostly invited from the general adult population residing in a given town or geographical area. There were, however, exceptions to this recruitment scheme. The French cohort was based on members of a national health insurance for teachers (with the aim of facilitating follow-up for incidence of cancer and other diseases). The cohorts in Utrecht (the Netherlands) and Florence (Italy) included women invited for a local population-based breast cancer screening program. In Oxford (UK) half of the cohort was recruited among subjects who did not eat meat, including vegans (who consume no animal products), lacto-ovo vegetarians, and fish eaters (i.e., consumers of fish but not meat). Standardized lifestyle and personal history questionnaires were collected from all study participants, and for more than 430,000 participants a blood sample was also collected at first recruitment. Blood samples were processed into fractions containing serum, plasma, red cells, and buffy coat and were frozen and stored in a central biorepository at the International Agency for Research on Cancer (Lyon, France, –196°C in liquid nitrogen) for all countries except Denmark (–150°C in nitrogen vapor) and Sweden (in –80°C freezer). All participants gave written informed consent to participate in the study, and the present study was approved by the local ethics committees in the participating countries and by the Internal Review Board of the International Agency for Research on Cancer (Lyon, France).

### Follow-up for cancer incidence and vital status

Data on vital status in EPIC are collected through record linkage with regional and/or national mortality registries, and systematic checks through municipal population registries. Cancer incidence is determined through record linkage with regional cancer registries (Denmark, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom, complete, for the present report, up to December 2004) or through the use of health insurance records or active follow-up of study participants, followed by contacts with

<sup>28</sup>Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway. <sup>29</sup>Department of Research, Cancer Registry of Norway, Oslo, Norway. <sup>30</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. <sup>31</sup>Samfundet Folkhälsan, Helsinki, Finland. <sup>32</sup>National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands. <sup>33</sup>Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, the Netherlands. <sup>34</sup>School of Public Health, Imperial College, London, United Kingdom. <sup>35</sup>Division of Environmental Epidemiology, Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, the Netherlands. <sup>36</sup>Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, the Netherlands. <sup>37</sup>Human Genetic Foundation (HuGeF), Turin, Italy. <sup>38</sup>MRC-HPA Centre for Environment and Health, Department of Epidemiology and Biostatistics, Imperial

College London, London, United Kingdom. <sup>39</sup>School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom. <sup>40</sup>Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom.

†Deceased.

**Corresponding Author:** Rudolf Kaaks, Department of Cancer Epidemiology, German Cancer Research Center (DKFZ) Heidelberg, Im Neuenheimer Feld 581, Heidelberg 69120, Germany. Phone: 49-6221-42-2219; Fax: 49-6221-42-2203; E-mail: r.kaaks@dkfz-heidelberg.de

doi: 10.1158/1055-9965.EPI-14-1107

©2014 American Association for Cancer Research.

clinical pathology registries (France, Germany, and Greece complete up to December 2006). All reported diagnoses of lymphoid cancers reported within the European EPIC project were centrally reviewed by hematologists. Cases with lymphoid cancers were originally classified according to the second revision of the International Classification of Diseases for Oncology (ICD-O-2), but were subsequently reclassified according to the WHO classification of hematopoietic and lymphoid tissue cancers, third edition (18), using a program available on the United States National Cancer Institute Surveillance Epidemiology and End Results (SEER) website (<http://seer.cancer.gov/lymphomarecode/lymphoma-who2008.html>). As staging data as well as B-cell counts were unavailable for EPIC CLL cases, our reclassification of hematologic disease endpoints into ICD-O-3 coding did not differentiate between progressive MBL (previously classified CLL RAI 0 with B-cell counts below 5,000/mm<sup>3</sup>) or CLL according to more recent (ICD-O-3) definitions. In none of the local EPIC subcohorts was there any organized form of screening for hematologic disorders. Participants of the French cohort ( $n = 24,371$  with available blood samples) were excluded from the study because the clinical verification and endpoint coding of lymphoid neoplasms, initially identified through health insurance records, was incomplete at the time the present study was started. After the additional exclusion of cases who had any other form of cancer except benign skin tumors before the diagnosis of lymphoma, a total of 181 cases of CLL were available for the present study.

#### Nested case-control design and participant selection

For each case of CLL, one control was selected by incidence density sampling from all cohort participants who were alive and without a cancer diagnosis at the time the corresponding index case with lymphoma was diagnosed (19), and matching the index case for study centre, sex, blood donor status, age at blood donation ( $\pm 12$  months), and date ( $\pm 3$  months). The incidence density sampling method allows control subjects to be sampled twice, or even to become a CLL case at a later time in prospective follow-up. This, however, did not actually occur in the case-control sample generated for the present study.

#### Laboratory analysis

Plasma concentrations of sCD23 were determined using a commercially available ELISA assay (BIOSOURCE EASIA; Invitrogen). All sCD23 measurements were performed in the same laboratory (Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany). CLL cases and their matched control subjects were systematically analyzed within the same immunoassay batch, and laboratory personnel were blinded to the case or control status of participants. Plasma samples from each case-control set were assayed within the same analytic batch. Two plasma quality control samples were analyzed within each batch. The within-batch coefficient of variation was 3.9% and the between-batch coefficient of variation was 8.7%.

According to the manufacturer, the BIOSOURCE EASIA assay provides a dynamic range of 0.10 to 20.0 U/mL, in which 1 U/mL corresponds to 1.25 ng/mL of sCD23 protein (25-Da isoform). Because of slight flattening of standard curve at the extreme ends, the highest level of sCD23 that could be reliably quantified was 17.07 U/mL and the lowest quantifiable level was 0.90 U/mL. Therefore, measurements at the lower limit of the detection range [2 cases (1.12%) and 17 control subjects (9.50%)] were set to the lowest quantifiable value (0.90 U/mL) and likewise, for those at

the higher limit [44 cases (24.58%) and 0 control subjects (0%)] values were set to the maximum quantifiable value of 17.07 U/mL. For two individuals, the serum samples available for this study had to be discarded after a technical liquid handling error, leading to the exclusion of two case-control pairs from the statistical analysis.

#### Statistical analysis

Data from 179 matched case-control pairs were available for statistical analysis. The distributions of plasma sCD23 of CLL cases and their matched control subjects were described using Box plots, and differences between the distributions between cases and controls or across covariate levels were evaluated by non-parametric ANOVA (Kruskal-Wallis/Wilcoxon) tests. Conditional logistic regression was used to calculate ORs and 95% confidence intervals (95% CI) of CLL diagnosis in relation to quintile levels of plasma sCD23, using quintile cutoff points based on the control subjects. As the vast majority of CLL cases had serum sCD23 measurements within the highest quintile, there were very sparse case numbers in the lower four quintiles. In addition, across the lower four quintiles there was no evidence for association between sCD23 and CLL diagnosis. We therefore combined the lower four quintiles into a single reference category. Associated *P* values were computed using likelihood ratio statistics, indicating the improvement of the model fit after including the variable in the model. Possible confounding effects by smoking status (categorical: life-long nonsmoker, current smoker, ex-smoker), alcohol consumption at the time of recruitment (g/d), and highest attained school level (categorical: none, primary school completed, technical/professional school, secondary school, longer education (incl. university degree) were examined by additional inclusion of these variables in the logistic regression model. However, none of these additional adjustments changed OR estimates by more than 8%, and therefore they were not retained for our final risk model. Tests for interaction between other risk factors and CD23 were done by adding the multiplicative terms in the model and tested by likelihood ratio statistics. However, no significant interactions were observed. Associations of CD23 with risk of CLL were additionally examined by duration of lag time between blood donation and diagnosis of CLL, and by age of the study subjects at the time of blood donation, using quartile cutoff points to form categories of lag time and age, and combining the middle (second and third) quartiles into a single category. Within each category of lag time and age, ORs of CLL by sCD23 level were calculated, and the predictive capacity of sCD23 for future diagnosis of CLL was evaluated by analysis of the relative or ROCs, plotting sensitivity against 1 minus specificity. The ability of sCD23 to discriminate between individuals who developed CLL or not is summarized by the AUC, with an AUC of 100% being the best possible test, and an AUC of 50% corresponding to a random test with no discriminatory power.

All analyses were performed using SAS version 9.2 (SAS Institute Inc.). All tests of statistical significance were two-sided, and *P* values <0.05 were considered significant.

## Results

In total, 179 incident cases of CLL cases and an equal number of matched control subjects were included in this study. The CLL cases were on average 58.3-years-old at baseline (range,

**Table 1.** Characteristics of CLL cases and their matched controls

Variable	N (%) <sup>b</sup>	Controls (N = 179) Median sCD23 (range)	P <sup>a</sup>	N (%) <sup>b</sup>	Cases (N = 179) Median sCD23 (range)	P <sup>a</sup>
sCD23 (U/mL)		1.60 (0.90–6.04)			7.26 (0.90–17.07)	<0.001
Age at recruitment, y		58.3 (33.5–75.0)			58.3 (33.6–75.0)	
Age at diagnosis, y					63.2 (41.5–82.6)	
Lag time (years)					5.10 (0.06–12.5)	
BMI (kg/m <sup>2</sup> )			0.50			0.56
Normal (≤25.0)	69 (39)	1.59 (0.90–3.29)		71 (40)	6.18 (0.90–17.07)	
Overweight (25.1–29.9)	77 (43)	1.60 (0.90–4.41)		73 (41)	8.76 (0.90–17.07)	
Obese (≥30.0)	33 (18)	1.65 (0.90–6.04)		35 (19)	7.32 (1.10–17.07)	
Gender, N (%)			0.74			0.56
Male	102 (57)	1.59 (0.90–4.46)		102 (57)	7.29 (0.90–17.07)	
Female	77 (43)	1.64 (0.90–6.04)		77 (43)	7.11 (0.90–17.07)	
Smoking, N (%)			0.46			0.45
No	69 (39)	1.70 (0.90–6.04)		89 (50)	8.41 (0.90–17.07)	
In the past	63 (35)	1.44 (0.90–4.46)		52 (29)	7.15 (0.98–17.07)	
Current (at blood donation)	40 (22)	1.62 (0.90–4.03)		38 (21)	5.78 (0.92–17.07)	
Alcohol consumption at blood donation (g/day)			<0.01			0.59
<3.2	55 (31)	1.75 (0.90–6.04)		62 (35)	6.16 (0.92–17.07)	
3.2–24	78 (44)	1.59 (0.90–4.41)		78 (44)	8.23 (0.90–17.07)	
≥25	45 (25)	1.26 (0.90–2.74)		39 (21)	7.26 (1.55–17.07)	
Highest school level			0.33			0.47
None	8 (4)	1.88 (1.14–6.04)		8 (4)	11.68 (1.58–17.07)	
Primary school	58 (32)	1.67 (0.90–4.03)		53 (30)	7.70 (0.96–17.07)	
Technical/professional school	51 (29)	1.64 (0.90–4.41)		47 (26)	8.09 (0.92–17.07)	
Secondary school	18 (10)	1.45 (0.90–3.60)		25 (14)	8.99 (0.90–17.07)	
Higher education	36 (20)	1.44 (0.90–3.92)		40 (22)	4.91 (0.90–17.07)	

<sup>a</sup>P values for differences in sCD23 distributions between case and control subjects and across variable categories were obtained by the Kruskal-Wallis test or Wilcoxon test, as appropriate.

<sup>b</sup>The percentages do not add up to 100% because of missing values.

33.6–75.0 years) and 63.2 years at diagnosis of CLL (range, 41.5–82.6 years). The median follow-up time between blood donation and date of diagnosis was 5.1 years (range, 0.06–12.5 years). Overall, CLL cases and controls showed no significant differences in level of formal education, smoking status, alcohol consumption, or body mass index (BMI; Table 1). Plasma sCD23 concentrations correlated weakly and positively with age, more strongly so among individuals who developed CLL (Spearman  $R = 0.32$ ; 95% CI, 0.18–0.44) than among the control subjects ( $R = 0.16$ ; 95% CI, 0.01–0.30), but showed no significant difference by sex, level of BMI, and smoking. Among the control subjects, but not among the CLL cases, level of alcohol consumption was inversely associated with plasma sCD23 concentrations.

On average, levels of sCD23 were significantly higher for subjects who had a diagnosis of CLL (median, 7.26 U/mL 5th–95th percentile range 1.19–17.07+ U/mL; capped) than among the controls (median, 1.60 U/mL 5th–95th percentile range 0.90–6.04 U/mL; Fig. 1A). Examining the data within three strata of age at the time of blood donation (33.5–53.3, 53.4–62.1, and 62.2–75.0 years), the case-control differences were larger for the older as compared with the younger subjects (Fig. 1C, Box plots). Likewise, the distributions of sCD23 of cases and controls differed more strongly for case-control sets in which the cases had been diagnosed within shorter lag time since blood donation [(0.1–2.7 years) as compared with longer lag times 2.8–7.3 years, or 7.4–12.5 years; Fig. 1B]. Of note, even in the case-control subset within the highest quartile of lag time (7.4–12.5 years) 44.2% of the cases had sCD23 values above the maximum value observed for the control subjects, and 55.8% of the cases had values above the 90th percentile of the controls.

Conditional logistic regression models showed a strong increase in the probability of CLL diagnosis only for the highest quintile of sCD23 and no association at lower levels of sCD23 (Table 2). No significant interactions between sCD23 and other risk factors were observed.

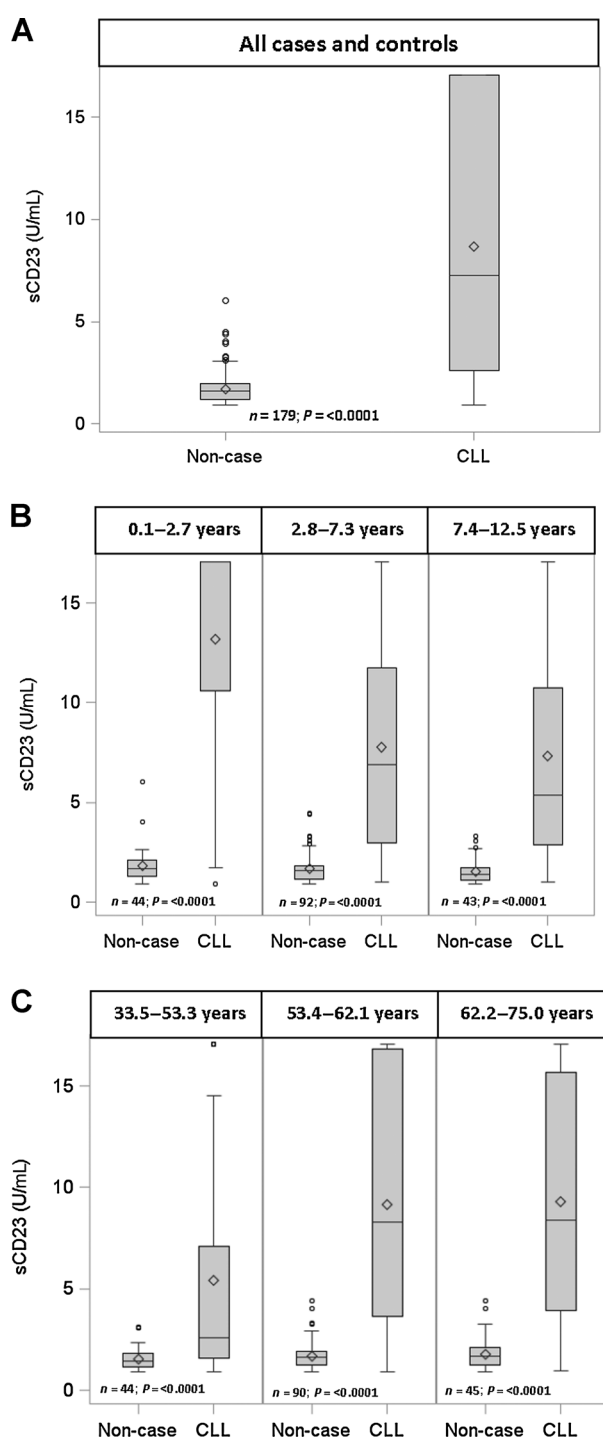
All case-controls sets combined, irrespective of age or lag time, the analysis of ROCs showed an AUC of 0.88 for a model including sCD23 as a predictor of CLL diagnosis (ROC figure not shown). Plotting ROC curves for case-control sets within three strata of lag time between blood donation and diagnosis (for the cases), the AUC equaled 0.95 for the lowest quartile of lag time (0.1–2.7 years), 0.90 for the middle two quartiles combined (2.4–7.3 years), and 0.76 for the highest quartile (7.4–12.5 years; Fig. 2). Plotting ROC curves for case-control sets by quartiles of different age at blood donation, the predictive capacity of sCD23 for diagnosis of CLL showed a clear increase with age, with AUC of 0.94 (95% CI, 0.89–0.99) for the oldest age group (>62 years at blood donation; Fig. 2). Adjustments for smoking, alcohol consumption, or BMI did not alter the AUROC estimates. Among the CLL cases, age at blood donation was only weakly inversely associated with lag time until diagnosis ( $R = -0.20$ ; 95% CI,  $-0.34$  to  $-0.05$ ), whereas lag time until diagnosis showed a moderate inverse correlation with sCD23 concentrations ( $R = -0.42$ ; 95% CI,  $-0.53$  to  $-0.29$ ).

## Discussion

In this prospective analysis among initially asymptomatic individuals, high blood levels of sCD23 were found to be a strong predictor for the later clinical diagnosis of CLL, and indicated long latency times for the existence of B-cell lymphoproliferative disorder preceding clinical disease. As judged by the AUC of an



Kaaks et al.

**Figure 1.**

Box plots of plasma sCD23 measurements by case-control status. A, all cases and controls; B, by three strata of lag time between blood draw and CLL diagnosis; C, by three strata of age at blood draw.

ROC curve, the prediction of CLL by sCD23 was very strong for CLL diagnoses within about 7 years or less after blood measurement ( $AUC = 0.92$ ), and was fairly strong even for diagnoses after a longer time interval (7.4–12.5 years;  $AUC = 0.76$ ). In addition,

**Table 2.** ORs (95% CI) for CLL diagnosis with overall matched case-control sets and with stratification by lag time and age at blood draw

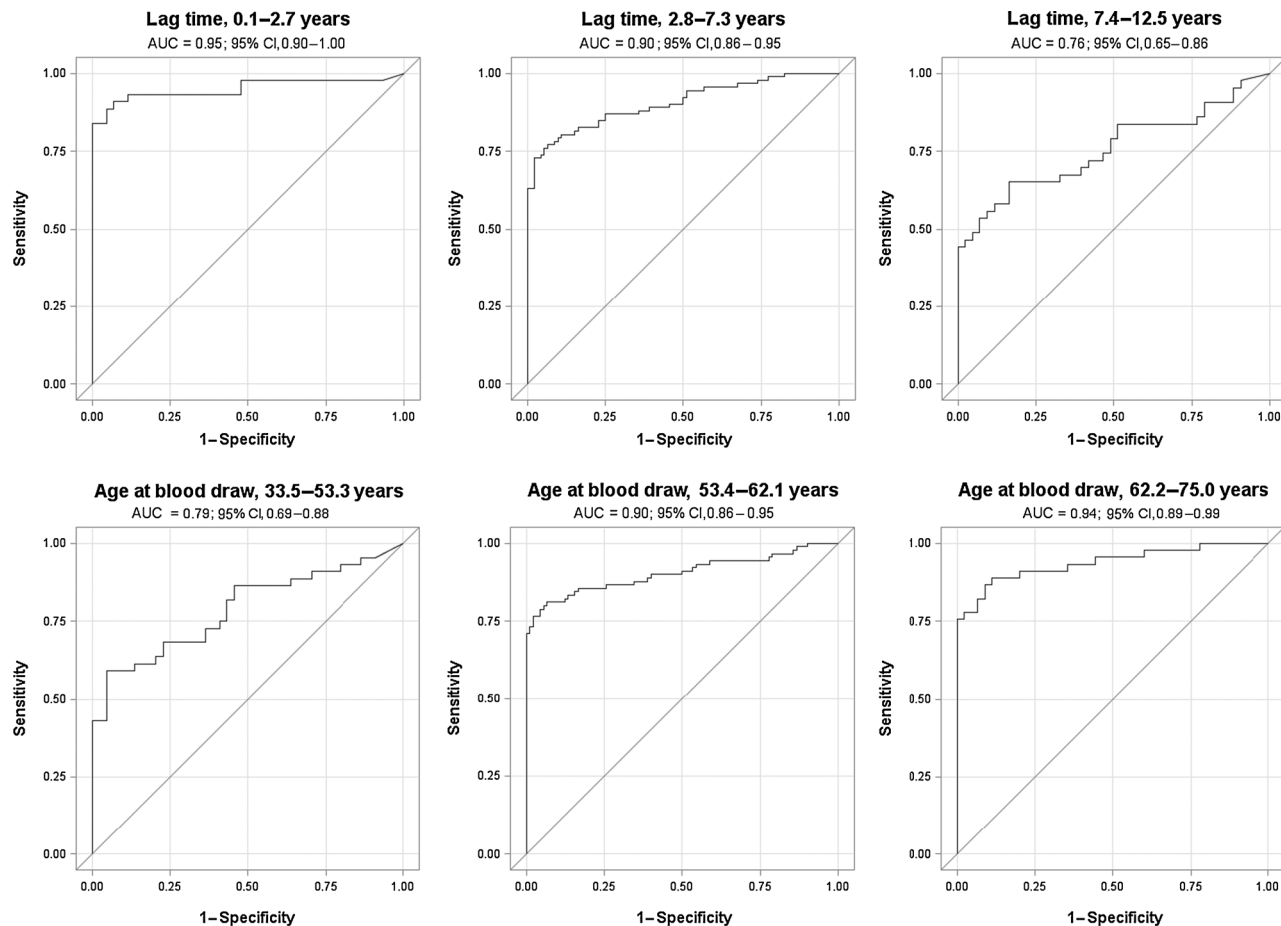
	The 1st–4th quintiles of sCD23 combined	The 5th quintile of sCD23	$P^a$
Overall			
ca/co	36/144	143/35	$<0.0001$
OR (95% CI)	Ref.	22.60 (9.29–55.35)	
Lag time			
0.1–2.7 years			$<0.0001$
ca/co	3/34	41/10	
OR (95% CI)	Ref.	32.00 (4.37–234.18)	
2.8–7.3 years			$<0.0001$
ca/co	16/74	76/18	
OR (95% CI)	Ref.	20.33 (6.38–64.80)	
7.4–12.5 years			$<0.0001$
ca/co	17/36	26/7	
OR (95% CI)	Ref.	20.00 (2.68–149.02)	
Age at blood draw			
33.5–53.3 years			$<0.0001$
ca/co	18/38	26/6	
OR (95% CI)	Ref.	7.67 (2.30–25.53)	
53.4–62.1 years			$<0.0001$
ca/co	14/76	76/14	
OR (95% CI)	Ref.	63.00 (8.74–454.23)	
62.2–75.0 years			$<0.0001$
ca/co	4/30	41/15	
OR (95% CI)	Ref.	27.00 (3.67–198.69)	

<sup>a</sup>A  $P$  value was based on the likelihood ratio test for overall effect of the covariate indicating the improvement in model fit.

the prediction of CLL diagnosis was also stronger for individuals who had provided blood at an older age (for subjects  $\geq 62$  years,  $AUC = 0.94$ ) as compared with younger age (for subjects  $< 53.4$  years,  $AUC = 0.79$ ).

At least three previous prospective studies have shown an increase in risk of B-cell NHL among patients with acquired immunodeficiency syndrome who had elevated serum sCD23 (20–22), but each of these studies had very small numbers of incident lymphoma cases and they included few, if any cases specifically of CLL. By contrast, in the United States one recent study within the Women's Health Initiative cohort—a prospective study of initially asymptomatic women recruited from the general population—showed an increased relative risk of developing CLL, small lymphocytic lymphoma, or prolymphocytic leukemia as a combined disease endpoint ( $N = 140$ ) among women with elevated sCD23 (OR, 16.1; 95% CI, 7.2–36.2), comparing the top quartile ( $\geq 75.0$  U/mL) versus bottom quartile ( $< 29.2$  U/mL) of sCD23. Quartile cutoff points determined in the control population; ref. 23). In this latter study, however, sCD23 was interpreted primarily together with other markers as an indicator of immune stimulatory host environment and B-cell activation.

Our finding that the majority of CLL cases must have had a preexisting progressive lymphoproliferative disorder up to 7 years and longer before diagnosis fits well to the findings of the landmark study by Landgren and colleagues (2) in the U.S. PLCO cohort, which had shown that the vast majority of individuals who developed CLL (44 of 45 patients) had high-count MBL up to 6 years before diagnosis. Contrary to this landmark study, however, our sCD23 data were not sufficient to indicate whether in all patients the characteristics of this preclinical disorder at the time of blood donation would have already corresponded to a diagnosis of full-blown CLL, or whether the disorder would have been classified as MBL. Among diagnosed patients with CLL, it is well documented that blood concentrations of sCD23 increase

**Figure 2.**

ROC curves for the prediction of future CLL diagnosis by plasma sCD23, by three strata of lag time between blood draw and CLL diagnosis, and by three strata of age at blood draw.

with progressing disease stage (24, 25). Furthermore, sCD23 levels have been shown to be a reflection of tumor mass and the size of the CD23 + B-cell pool (15, 16). It is, thus, very likely that sCD23 concentrations will also correlate strongly with the size of the B-cell clone in patients with re-CLL who have progressive MBL. However, we are not aware of published reports relating sCD23 concentrations to B-cell counts in MBL at various stages of progression of lymphocytosis, and the general lack of standardization of sCD23 immunoassays also prohibits cross-study comparisons of absolute sCD23 levels in relation to various stages of B-cell lymphoproliferative disease.

Our findings have major relevance with respect to the interpretation of prospective studies in EPIC and other cohorts that aim to identify possible causes of CLL. For example, prospective studies (26), including ours (27), have suggested an inverse relationship between serum levels of total and/or allergen specific IgE and risk of developing CLL and other forms of NHL. One proposed interpretation for this inverse relationship is that reduced immune surveillance, reflected by low IgE, could be a contributing cause of lymphoma (28, 29). In our data, we observed a fairly strong inverse correlation of total IgE with sCD23 in the prediagnostic blood of CLL cases ( $r = -0.50$  among the male patients and  $r = -0.17$  for the female patients), but not of the

control subjects. This suggests that low IgE levels are not necessarily a cause of CLL, but equally well could have been a consequence of preexisting disease. In a similar manner, as the expression of transmembrane CD23 on B cells and sCD23 production are regulated by many cytokines, including several ILs (IL4, IL13, IL5, IL9, and IL10) and IFN $\gamma$  (11, 30), it is possible that findings from prospective studies linking CLL risk to blood levels of proinflammatory cytokines (31) could also be a reflection of preexisting, but still latent disease.

Within the first 2.7 years of follow-up, and using a cutoff point of 6.2 U/mL, sCD23 predicted a clinical diagnosis of CLL with 100% specificity at >84% sensitivity (Fig. 2). This high specificity suggests that sCD23 would have potential as a marker for early detection of CLL. However, sCD23 levels can also be increased in conditions other than B-cell lymphoproliferative disorders, such as allergies, endometriosis, and autoimmune diseases such as rheumatoid arthritis or Sjögren's syndrome (32, 33), conditions in which the increase in sCD23 may arise as result of from polyclonal activation of B cells (11). Thus, sCD23 may lack specificity for a comprehensive detection of all CLL in an asymptomatic population, and cannot replace standard diagnostic tests based on blood cell counts. Furthermore, as there is currently no therapeutic intervention for CLL-

like MBL or early-stage CLL, screening for CLL cannot be recommended. However, the rapid discovery of somatic mutation spectra or epigenetic markers specific for CLL (34) may eventually allow further improvements in specificity of molecular subtypes of CLL that are most aggressive and for which earlier treatment could actually be beneficial.

A limitation of our study is the lack of characterization of study participants at baseline recruitment, in that these had not been screened for the presence of MBL or CLL in a clinically well-differentiated manner, including blood cell counts and cell sorting. Thus, we could not assess whether sCD23 is a better predictor than the B-cell count for the prediction of future clinical diagnosis of CLL. However, compared with existing clinical cohorts of patients with MBL, a strength of our study is the relatively large number of incident cases of newly diagnosed clinical CLL in a cohort of initially asymptomatic individuals with prediagnostic blood samples, with prospective follow-up times of up to 12.5 years.

Our study indicates that in the vast majority of patients with CLL diagnosed within the European general population, the diagnosis of CLL is preceded by high plasma levels of sCD23—a marker indicating the existence of prediagnostic lymphoproliferative disease—by a time period of up to 7 years or even longer. The length of the apparent lag times between detectable proliferative disorders until clinical diagnosis of CLL has major implications for the interpretation of prospective epidemiologic studies on the causes of CLL. Our findings also suggest relatively long time intervals over which early disease detection may be theoretically possible and relevant, as methods are being developed for the identification of aggressive CLL subtypes that would benefit from earlier treatment.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Authors' Contributions

**Conception and design:** R. Kaaks, R. Tumino, H. Boeing, E. Ardanaz, K. Overvad, A. Tjønneland, J. Halkjær, E. Weiderpass, H.B. Bueno-de-Mesquita, P.H. Peeters, P. Vineis, A. Nieters

**Development of methodology:** S. Becker, R. Tumino, P. Brennan, E. Ardanaz, E. Weiderpass, H.B. Bueno-de-Mesquita, P.H. Peeters, P. Vineis, A. Nieters

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** R. Kaaks, S. Becker, A. Johansson, B. Melin, D. Trichopoulos, A. Trichopoulou, P. Lagiou, A. Mattiello, R. Tumino, G. Masala, H. Boeing, P. Brennan, M.-J. Sánchez, E. Ardanaz, N. Sala, K. Overvad, A. Tjønneland, J. Halkjær, E. Weiderpass, H.B. Bueno-de-Mesquita, P.H. Peeters, K.-T. Khaw, R.C. Travis, T.J. Key

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** R. Kaaks, D. Sookthai, A. Łuczyńska, S. Becker,

B. Melin, K. Sjöberg, D. Trichopoulos, D. Casabonne, M.-J. Sánchez, E. Ardanaz, K. Overvad, E. Weiderpass, A. Nieters

**Writing, review, and/or revision of the manuscript:** R. Kaaks, D. Sookthai, C.C. Oakes, S. Becker, T. Johnson, B. Melin, K. Sjöberg, D. Trichopoulos, A. Trichopoulou, P. Lagiou, A. Mattiello, G. Masala, C. Agnoli, H. Boeing, K. Aleksandrova, S. Franceschi, S. Roulland, D. Casabonne, S. de Sanjose, M.-J. Sánchez, J.M. Huerta, E. Ardanaz, N. Sala, K. Overvad, A. Tjønneland, J. Halkjær, E. Weiderpass, H.B. Bueno-de-Mesquita, R. Vermeulen, P.H. Peeters, P. Vineis, R.S. Kelly, K.-T. Khaw, R.C. Travis, T.J. Key, E. Riboli, A. Nieters

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** R. Kaaks, D. Sookthai, T. Johnson, A. Trichopoulou, R. Tumino, M.-J. Sánchez, E. Ardanaz, E. Weiderpass, K.-T. Khaw

**Study supervision:** R. Kaaks, R. Tumino, E. Ardanaz, E. Weiderpass, H.B. Bueno-de-Mesquita, A. Nieters

## Grant Support

The study was partially supported by the German Federal Ministry of Education and Research (to R. Kaaks, A. Nieters, D. Sookthai, A. Łuczyńska, and T. Johnson; grant number A BMBF 01 EO 0803). This work was supported by WCRF UK and WCRF International, grant no 2009/39. The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer (to P. Brennan, S. Franceschi, and S. Roulland). The national cohorts are supported by Danish Cancer Society Denmark (to K. Overvad, A. Tjønneland, and J. Halkjær), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum, and Federal Ministry of Education and Research Germany (to R. Kaaks, A. Nieters, D. Sookthai, A. Łuczyńska, C.C. Oakes, S. Becker, T. Johnson, H. Boeing, and K. Aleksandrova), Ministry of Health and Social Solidarity, Stavros Niarchos Foundation, and Hellenic Health Foundation Greece (to D. Trichopoulos, A. Trichopoulou, and P. Lagiou), Italian Association for Research on Cancer (AIRC) and National Research Council Italy (to A. Mattiello, R. Tumino, G. Masala, and C. Agnoli), Dutch Ministry of Public Health, Welfare, and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), and Statistics Netherlands from the Netherlands (to H. B. Mesquita, R. Vermeulen, and P. H. Peeters), ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence program on Food, Nutrition, and Health Norway (to E. Weiderpass), Health Research Fund (FIS), Regional Governments of Andalucía, Asturias, Basque Country, Murcia (no. 6236), and Navarra, ISCIII RETIC (RD06/0020) Spain (to D. Casabonne, S. D. Sanjose, M. J. Sanchez, J. M. Huerta, and E. Ardanaz, and N. Sala), Swedish Cancer Society, Swedish Scientific Council and Regional Government of Skåne and Västerbotten Sweden (to A. Johansson, B. Melin, and K. Sjöberg), Cancer Research UK, Medical Research Council, Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency, and Wellcome Trust United Kingdom (to T.J. Keys, E. Riboli, R. C. Travis; K.T. Khaw, and R.S. Kelly, and P. Vineis).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 30, 2014; revised November 25, 2014; accepted December 5, 2014; published OnlineFirst December 26, 2014.

## References

- Rawstron AC, Bennett FL, O'Connor SJ, Kwok M, Fenton JA, Plummer M, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med* 2008;359:575–83.
- Landgren O, Albitar M, Ma W, Abbasi F, Hayes RB, Ghia P, et al. B-cell clones as early markers for chronic lymphocytic leukemia. *N Engl J Med* 2009;360:659–67.
- Shim YK, Middleton DC, Caporaso NE, Rachel JM, Landgren O, Abbasi F, et al. Prevalence of monoclonal B-cell lymphocytosis: a systematic review. *Cytometry B Clin Cytom* 2010;78:S10–8.
- Rawstron AC, Shanafelt T, Lanasa MC, Landgren O, Hanson C, Orfao A, et al. Different biology and clinical outcome according to the absolute numbers of clonal B-cells in monoclonal B-cell lymphocytosis (MBL). *Cytometry B Clin Cytom* 2010;78 Suppl 1:S19–23.
- Vardi A, Dagklis A, Scarfo L, Jelinek D, Newton D, Bennett F, et al. Immunogenetics shows that not all MBL are equal: the larger the clone, the more similar to CLL. *Blood* 2013;121:4521–8.
- Marti GE, Rawstron AC, Ghia P, Hillmen P, Houlston RS, Kay N, et al. Diagnostic criteria for monoclonal B-cell lymphocytosis. *Br J Haematol* 2005;130:325–32.
- Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. *Leukemia* 2010;24:512–20.

8. Rawstron AC. Monoclonal B cell lymphocytosis—what does it really mean? *Curr Hematol Malig Rep* 2013;8:52–9.
9. Mowery YM, Lanasa MC. Clinical aspects of monoclonal B-cell lymphocytosis. *Cancer Control* 2012;19:8–17.
10. Sarfati M, Bron D, Lagneaux L, Fonteyn C, Frost H, Delespesse G. Elevation of IgE-binding factors in serum of patients with B-cell–derived chronic lymphocytic leukemia. *Blood* 1988;71:94–8.
11. Platzer B, Ruiter F, van der Mee J, Fiebiger E. Soluble IgE receptors—elements of the IgE network. *Immunol Lett* 2011;141:36–44.
12. Acharya M, Borland G, Edkins AL, Maclellan LM, Matheson J, Ozanne BW, et al. CD23/FcεpsilonRII: molecular multi-tasking. *Clin Exp Immunol* 2010;162:12–23.
13. Saka B, Aktan M, Sami U, Oner D, Sanem O, Dincol G. Prognostic importance of soluble CD23 in B-cell chronic lymphocytic leukemia. *Clin Lab Haematol* 2006;28:30–5.
14. Knauf WU, Langenmayer I, Ehlers B, Mohr B, Adorf D, Nerl CH, et al. Serum levels of soluble CD23, but not soluble CD25, predict disease progression in early stage B-cell chronic lymphocytic leukemia. *Leukemia Lymphoma* 1997;27:523–32.
15. Callea V, Morabito F, Luise F, Piromalli A, Filangeri M, Stelitano C, et al. Clinical significance of sIL2 R, sCD23, sICAM-1, IL6 and sCD 14 serum levels in B-cell chronic lymphocytic leukemia. *Haematologica* 1996;81:310–5.
16. Sarfati M, Fournier S, Christoffersen M, Biron G. Expression of CD23 antigen and its regulation by IL-4 in chronic lymphocytic leukemia. *Leukemia Res* 1990;14:47–55.
17. Bingham S, Riboli E. Diet and cancer—the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer* 2004;4:206–15.
18. Jaffe ES. The 2008 WHO classification of lymphomas: implications for clinical practice and translational research. *Hematol Am Soc Hematol Educ Program* 2009;1:523–31.
19. Lubin JH, Gail MH. Biased selection of controls for case–control analyses of cohort studies. *Biometrics* 1984;40:63–75.
20. Yawetz S, Cumberland WG, van der Meyden M, Martinez-Maza O. Elevated serum levels of soluble CD23 (sCD23) precede the appearance of acquired immunodeficiency syndrome–associated non-Hodgkin's lymphoma. *Blood* 1995;85:1843–9.
21. Hussain SK, Hessol NA, Levine AM, Breen EC, Anastos K, Cohen M, et al. Serum biomarkers of immune activation and subsequent risk of non-Hodgkin B-cell lymphoma among HIV-infected women. *Cancer Epidemiol Biomarkers Prev* 2013;22:2084–93.
22. Breen EC, Hussain SK, Magpantay L, Jacobson LP, Detels R, Rabkin CS, et al. B-cell stimulatory cytokines and markers of immune activation are elevated several years prior to the diagnosis of systemic AIDS-associated non-Hodgkin B-cell lymphoma. *Cancer Epidemiol Biomarkers Prev* 2011;20:1303–14.
23. De Roos AJ, DK Mirick, Edlefsen KL, LaCroix AZ, Kopecky KJ, Madeleine MM, et al. Markers of B-cell activation in relation to risk of non-Hodgkin lymphoma. *Cancer Res* 2012;72:4733–43.
24. Reinisch W, Willheim M, Hilgarth M, Gasche C, Mader R, Szepefalusi S, et al. Soluble CD23 reliably reflects disease activity in B-cell chronic lymphocytic leukemia. *J Clin Oncol* 1994;12:2146–52.
25. Schwarzmeier JD, Shehata M, Hilgarth M, Marschitz I, Louda N, Hubmann R, et al. The role of soluble CD23 in distinguishing stable and progressive forms of B-chronic lymphocytic leukemia. *Leuk Lymphoma* 2002;43:549–54.
26. Melbye M, Smedby KE, Lehtinen T, Rostgaard K, Glimelius B, Munksgaard L, et al. Atopy and risk of non-Hodgkin lymphoma. *J Natl Cancer Inst* 2007;99:158–66.
27. Nieters A, Luczynska A, Becker S, Becker N, Vermeulen R, Overvad K, et al. Association between prediagnostic total and specific Immunoglobulin E levels and risk of lymphoma—results of the European Prospective Investigation into Cancer and Nutrition. *Carcinogenesis* 2014;35:2716–22.
28. Jensen-Jarolim E, Achatz G, Turner MC, Karagiannis S, Legrand F, Capron M, et al. AllergoOncology: the role of IgE-mediated allergy in cancer. *Allergy* 2008;63:1255–66.
29. Grulich AE, Vajdic CM, Cozen W. Altered immunity as a risk factor for non-Hodgkin lymphoma. *Cancer Epidemiol Biomarkers Prev* 2007;16:405–8.
30. Kara IO, Sahin B, Gunesacar R. Expression of soluble CD27 and interleukins-8 and -10 in B-cell chronic lymphocytic leukemia: correlation with disease stage and prognosis. *Adv Ther* 2007;24:29–40.
31. Purdue MP, Lan Q, Bagni R, Hocking WG, Baris D, Reding DJ, et al. Prediagnostic serum levels of cytokines and other immune markers and risk of non-Hodgkin lymphoma. *Cancer Res* 2011;71:4898–907.
32. Bansal AS, MacGregor AJ, Pumphrey RS, Silman AJ, Ollier WE, Wilson PB. Increased levels of sCD23 in rheumatoid arthritis are related to disease status. *Clin Exp Rheumatol* 1994;12:281–5.
33. Amrol DJ, Hagaman DD, Sheller JR, Murray JJ. Soluble CD23 and interleukin-1 receptor antagonist in human asthmatics following antigen challenge. *J Asthma* 2005;42:73–6.
34. Quesada V, Ramsay AJ, Rodriguez D, Puente XS, Campo E, Lopez-Otin C. The genomic landscape of chronic lymphocytic leukemia: clinical implications. *BMC Med* 2013;11:124.



# Cancer Epidemiology, Biomarkers & Prevention

## Lag Times between Lymphoproliferative Disorder and Clinical Diagnosis of Chronic Lymphocytic Leukemia: A Prospective Analysis Using Plasma Soluble CD23

Rudolf Kaaks, Disorn Sookthai, Anna Luczynska, et al.

*Cancer Epidemiol Biomarkers Prev* 2015;24:538-545. Published OnlineFirst December 26, 2014.

**Updated version** Access the most recent version of this article at:  
doi:[10.1158/1055-9965.EPI-14-1107](https://doi.org/10.1158/1055-9965.EPI-14-1107)

**Cited articles** This article cites 34 articles, 9 of which you can access for free at:  
<http://cebp.aacrjournals.org/content/24/3/538.full.html#ref-list-1>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, contact the AACR Publications Department at [permissions@aacr.org](mailto:permissions@aacr.org).