IJC
International Journal of Cancer

# Serum levels of *hsa-miR-16-5p*, *hsa-miR-29a-3p*, *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-223-3p* and subsequent risk of chronic lymphocytic leukemia in the EPIC study

Delphine Casabonne[1,2], Yolanda Benavente[1,2], Julia Seifert[3], Laura Costas[2], María Armesto[3], María Arestin[3], Caroline Besson[4,5,6], Fatemeh S. Hosnijeh[7,8], Eric J. Duell[9], Elisabete Weiderpass[10], Giovanna Masala[11], Rudolf Kaaks[12], Federico Canzian[13], María-Dolores Chirlaque[1,14], Vittorio Perduca[4,5,6,15], Francesca R. Mancini[4,5], Valeria Pala[16], Antonia Trichopoulou[17], Anna Karakatsani[17,18], Carlo La Vecchia[17,19], Maria-Jose Sánchez[1,20,21], Rosario Tumino[22], Marc J. Gunter[23], Pilar Amiano[1,24], Salvatore Panico[25], Carlotta Sacerdote[26], Julie A. Schmidt[27], Heiner Boeing[28], Matthias B. Schulze[29,30], Aurelio Barricarte[1,31,32], Elio Riboli[33], Anja Olsen[34], Anne Tjønneland[34,35], Roel Vermeulen[7], Alexandra Nieters[36], Charles H. Lawrie[3,37,38] and Silvia de Sanjosé[1,2,39]

[1]Centro de Investigación Biomédica en Red: Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[2]Unit of Molecular and Genetic Epidemiology in Infections and Cancer, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain
[3]Molecular Oncology Group, Biodonostia Research Institute, San Sebastián, Spain
[4]CESP, Faculté de Médecine, Université Paris-Sud, UVSQ, INSERM, Université Paris-Saclay, Villejuif, France
[5]Gustave Roussy, Villejuif, France
[6]Department of Hematology and Oncology, Hospital of Versailles, Le Chesnay, France
[7]Institute for Risk Assessment Sciences, Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands
[8]Department of Immunology, Laboratory Medical Immunology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands
[9]Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO-IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain
[10]International Agency for Research on Cancer, World Health Organization, Lyon, France
[11]Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, Italy
[12]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[13]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany
[14]Department of Epidemiology, Regional Health Council, IMIB-Arrixaca, Murcia University, Murcia, Spain
[15]Laboratoire de Mathématiques Appliquées MAP5 (UMR CNRS 8145), Université Paris Descartes, Paris, France
[16]Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy
[17]Hellenic Health Foundation, Athens, Greece
[18]Pulmonary Medicine Department, School of Medicine, National and Kapodistrian University of Athens, "ATTIKON" University Hospital, Haidari, Greece
[19]Dept. of Clinical Sciences and Community Health, Università Degli Studi di Milano, Milan, Italy
[20]Andalusian School of Public Health (EASP), Granada, Spain
[21]Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Universidad de Granada, Granada, Spain
[22]Cancer Registry and Histopathology Department, Azienda Sanitaria Provinciale (ASP), Ragusa, Italy
[23]Section of Nutrition and Metabolism, IARC, International Agency for Research on Cancer, Lyon, France
[24]Public Health Division of Gipuzkoa, BioDonostia Research Institute, San Sebastian, Spain
[25]Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy
[26]Unit of Cancer Epidemiology, Città Della Salute e Della Scienza University-Hospital and Center for Cancer Prevention (CPO), Turin, Italy
[27]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom
[28]Department of Epidemiology, German Institute of Human Nutrition (DIfE) Postdam-Rehbrücke, Nuthetal, Germany
[29]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany
[30]Institute of Nutritional Sciences, University of Potsdam, Nuthetal, Germany
[31]Navarra Public Health Institute, Pamplona, Spain

**Additional Supporting Information** may be found in the online version of this article.

**Key words:** chronic lymphocytic leukemia, circulating miRNA, serum, prospective study

**Abbreviations:** AUC: area under the curve; CI: confidence interval; CLL: chronic lymphocytic leukemia; CT: cycle threshold; EPIC: the European Prospective Investigation into Cancer and Nutrition; ICD-O: International Classification of Diseases for Oncology; MBL: monoclonal B-cell lymphocytosis; miRNA: microRNA; OR: odds ratio; PCA: principal component analysis; PCR: polymerase chain reaction; ROC: receiver operating characteristics

**DOI:** 10.1002/ijc.32894
**History:** Received 29 Aug 2019; Accepted 17 Dec 2019; Online 3 Feb 2020

**Correspondence to:** Delphine Casabonne, E-mail: dcasabonne@iconcologia.net

[32]Navarra Institute for Health Research (IdiSNA), Pamplona, Spain
[33]School of Public Health, Imperial College London, London, United Kingdom
[34]Danish Cancer Society Research Center, Copenhagen, Denmark
[35]Department of Public Health, The Copenhagen University, Copenhagen, Denmark
[36]Institute for Immunodeficiency (IFI) Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany
[37]Radcliffe Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom
[38]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[39]Reproductive Health, PATH, Seattle, WA



**Chronic lymphocytic leukemia (CLL) is an incurable disease accounting for almost one-third of leukemias in the Western world. Aberrant expression of microRNAs (miRNAs) is a well-established characteristic of CLL, and the robust nature of miRNAs makes them eminently suitable liquid biopsy biomarkers. Using a nested case–control study within the European Prospective Investigation into Cancer and Nutrition (EPIC), the predictive values of five promising human miRNAs (*hsa-miR-16-5p*, *hsa-miR-29a-3p*, *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-223-3p*), identified in a pilot study, were examined in serum of 224 CLL cases (diagnosed 3 months to 18 years after enrollment) and 224 matched controls using Taqman based assays. Conditional logistic regressions were applied to adjust for potential confounders. The median time from blood collection to CLL diagnosis was 10 years (p25–p75: 7–13 years). Overall, the upregulation of *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-29a-3p* was associated with subsequent risk of CLL [$OR_{1\triangle Ct\text{-unit increase}}$ (95%CI) = 1.42 (1.18–1.72), 1.64 (1.31–2.04) and 1.75 (1.31–2.34) for *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-29a-3p*, respectively] and the strongest associations were observed within 10 years of diagnosis. However, the predictive performance of these miRNAs was modest (area under the curve <0.62). *hsa-miR-16-5p* and *hsa-miR-223-3p* levels were unrelated to CLL risk. The findings of this first prospective study suggest that *hsa-miR-29a*, *hsa-miR-150-5p* and *hsa-miR-155-5p* were upregulated in early stages of CLL but were modest predictive biomarkers of CLL risk.**

**What's new?**

**Aberrant circulating microRNA (miRNA) levels are a well-established characteristic of chronic lymphocytic leukemia (CLL), but pre-diagnosis data remain scarce. In this nested case–control study within the European Prospective Investigation into Cancer and Nutrition Study, circulating *hsa-miRNA-29a-3p*, *-150-5p*, *and -155-5p* were upregulated up to 10 years before CLL diagnosis compared to controls, suggesting a role in early disease progression. However, these biomarkers were suboptimal to discriminate CLL from controls. No difference was observed in *hsa-miRNA-16-5p* and *-223-3p* expression between pre-CLL and controls. The ability to detect pre-clinical biomarkers may help better understand CLL development and open new avenues for personalized treatment.**

## Introduction

Chronic lymphocytic leukemia (CLL) is the most common form of leukemia in the Western world with an incidence of around 5 per 100,000 per year.[1] CLL is essentially an incurable disease characterized by an accumulation of abnormal B-cell in the blood, bone marrow and lymphoid tissues.[2] Increasing age, being Caucasian, a familial history of CLL and male sex are all well-established risk factors but the etiology of CLL remains unknown. Current diagnosis of CLL depends upon either the presentation of clinical symptoms or is mainly made serendipitously during routine blood testing.

CLL could be preceded by a precursor stage, monoclonal B-cell lymphocytosis (MBL), characterized by the presence of less than 5,000 clonal B-cells per μl in the peripheral blood and no other signs of a lymphoproliferative disorder. With the advancements of technology, two subgroups of CLL precursors have been defined (*i*) high count MBL ($MBL^{hi}$), with 500–5,000 circulating clonal B-cell per μl with a 1.1% rate of progression to CLL requiring treatment[3] and (*ii*) low count MBL ($MBL^{lo}$), with less than 500 circulating clonal B-cell per μl in otherwise healthy individuals. $MBL^{lo}$ is detected within research screening programs using highly sensitive flow-cytometry techniques in up to 12% of individuals aged 40 or over and its prevalence increases with age.[4] The reasons for progression to a larger B-cell clonal population from $MBL^{lo}$ to higher stages are not known. There is currently no early diagnosis or screening test for CLL, and after diagnosis, most patients are managed following a watch-and-wait approach. The ability to detect preclinical biomarkers in the blood of otherwise healthy individuals will contribute to a better understanding of the development and progression of MBL and CLL and might open new avenues to optimize personalized treatment strategies.

MicroRNAs (miRNAs), short (~22 nucleotides) noncoding single chain RNA, act as key regulators in almost every

biological processes examined to date and are aberrantly expressed in many, if not all, cancers including hematological malignancies.[5] Binding to near complementary sequences of mRNAs, miRNAs regulate gene expression.[6] In 2002, CLL was the first tumor to be associated with aberrant miRNA expression.[7] The loss of the expression of *hsa-miR-15a* and *hsa-miR-16-1*, located at chromosome 13q14, a locus often deleted in CLL patients (>50%), lead to the overexpression of the inhibitor of apoptosis BCL2.[7,8] Encoded by eukaryotic nuclear DNA, miRNAs can also be detected in body fluids[9,10] and their presence in blood in direct contact with tumor cells might provide molecular markers of the tumor cell microenvironment. Altered expression of miRNAs is observed in CLL patients and is associated with CLL progression and prognosis factors in many studies.[7] Due to their inherent chemical stability, and ability to complex with either protective proteins or sequestration in extracellular vesicles, miRNAs are eminently suitable liquid biopsy cancer biomarkers that have generated a lot of interest in recent years.[11] To the best of our knowledge, serum levels of miRNAs have never been examined prior to CLL diagnosis in a large prospective cohort study.[12]

In order to test the ability of miRNAs as predictive biomarkers of CLL, we investigated the levels of five promising miRNAs (*hsa-miR-16-5p*, *hsa-miR-29a-3p*, *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-223-3p*), selected following a pilot study[12–15] (Supporting Information Data S1), in a nested case–control study of 224 pre-CLL cases and 224 matched controls obtained from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.

## Materials and Methods

### Study population

EPIC is an ongoing prospective cohort study involving 23 centers from 10 European countries (Denmark, France, Germany, Greece, the Netherlands, Italy, Norway, United Kingdom, Spain and Sweden). The rationale, methods and study design have been described previously.[16,17] In brief, 521,324 subjects, aged between 30 and 70 years, were recruited between 1992 and 2000. Written informed consent was provided by all participants. The ethical review boards from the International Agency for Research on Cancer (IARC) and from all local centers approved the study. Prior to analysis, the following exclusions were made: participants with a prevalent cancer ($n$ = 25,184) and with missing follow-up information ($n$ = 4,148). Thus, our initial study population included 491,992 EPIC participants. Controls and cases provided a serum sample at recruitment.

### Follow-up and nested case–control design

Incident lymphoma cancer cases were identified by population cancer registries for Denmark, Italy, the Netherlands, Norway, Spain, Sweden and the United Kingdom. A combination of methods was used in France, Germany and Greece, as detailed previously.[17] Mortality data were also obtained from regional or national mortality registries. The follow-up period was defined from the age at recruitment to the age at first hematological cancer diagnosis, death or last complete follow-up, depending on which occurred first. Censoring dates for the last complete follow-up ranged from June 2008 to December 2013, depending on the EPIC center. Initially, the diagnosis of lymphoma cases was based on the second revision of the International Classification of Diseases for Oncology (ICD-O-2). Later, all cases were reclassified into the ICD-O-3 using a conversion program available on the web site of the Surveillance Epidemiology and End Results (SEER) program (http://seer.cancer.gov/tools/conversion/ICD02-3manual.pdf) and involving a pathology expert and experts from the EPIC centers. Finally, the InterLymph Pathology Working Group classification, which is based in the current 2008 WHO classification, was used to categorize lymphoma histologic subtypes. CLL and small lymphocytic lymphoma are considered the same underlying disease.

A case–control study nested within the EPIC study was set-up and cases were people with CLL from whom a serum sample was available from prior to diagnosis. Controls were alive and without a cancer diagnosis (other than non-melanoma skin cancer) at the time of the diagnosis of the CLL case. One control was randomly chosen for each case (1:1) using incidence density sampling and matched for sex, study center, age at blood collection (±12 months), date of blood collection (±3 months), time of blood collection (±3 hr) and fasting at blood collection (no/yes). The present study includes 224 CLL cases and 224 controls. CLL cases were diagnosed 3 months–18 years after blood collection and the median follow-up time was 10.0 years. In July 2011, the Danish biobank was flooded and the samples might have been affected with more freeze and thaw cycles.

### Selection of the miRNAs: EpiLymph-Spain pilot study

The EpiLymph-Spain study has been described in more detail previously.[18] Following a literature search (January 2002–February 2014), we run a pilot study examining 12 promising miRNA (*hsa-miR-15a-5p*, *hsa-miR-20a-5p*, *hsa-miR-21-5p*, *hsa-miR-23a-3p*, *hsa-miR-29a-3p*, *hsa-miR-34a-5p*, *hsa-miR-146a-5p*, *hsa-miR-150-5p*, *hsa-miR-155-5p*, *hsa-miR-181b-5p*, *hsa-miR-223-3p* and *hsa-miR-451a*) in 48 untreated CLL patients and 42 frequency-matched (age, sex and region) hospital-based controls from the Spanish data of the EpiLymph study. Methods and results can be found in Supporting Information Data S1. Based on the findings, we limited the final selection for the prospective analysis to these four most promising miRNA (*hsa-miR-29a-3p*, *hsa-miR-150-5p*, *hsa-miR-155-5*p and *hsa-miR-223-3p*) as well as *hsa-miR-16-5p* and *hsa-miR-24-5p* for potential normalization.

### Laboratory methods

Total RNA was purified from serum samples (200 µl) collected at recruitment using the Exiqon miRCURY RNA

Isolation Kit-Biofluids (Vedbaek, Denmark) according to the manufacturer's protocol. RNA was then reverse-transcribed using Megaplex RT primers pool A v.2.1 (Thermo Fisher, St. Louis, MO) according to manufacturer's protocol. Levels of individual miRNAs were measured using specific Taqman probes according to manufacturer's protocols (Thermo Fisher) in a CFX384 Touch Real-Time PCR Detection System (Bio-Rad). *Hsa-miR-24-5p* was used for normalize the Ct values. Samples were measured in triplicate in a blinded fashion without knowledge of the case–control status of the sample. The ΔCt for each individual was defined as: $(\sum_1^n hsa\_miR\_ref)/n - (\sum_1^n hsa\_miR\_of\_interest)/n$, where $n$ was the number of replicates ($n = 3$) and *hsa-miR-ref* was *hsa-miR-24-5p*.

### Statistical analysis

Distribution of cases and controls by selected putative risk factors (educational level: secondary or higher *vs*. other, body mass index (BMI): <25, 25–30, ≥30, physical activity: inactive *vs*. active based on the Cambridge physical activity index, smoking status: never, former and current, and alcohol consumption at recruitment, height and total energy intake as continuous) was examined. Heterogeneity for categorical exposures was assessed using chi-squared test and *t*-test or nonparametric test for normal or non-normal continuous exposures, respectively. Among controls, the association between ΔCt distribution and the selected variables was examined using linear regression adjusted for the matching variables. Spearman correlation quantified the association between miRNAs and principal component analysis (PCA) was applied to visualize patterns across miRNAs.

The association between pre-CLL participants and their matched controls was analyzed using the overall database and then stratifying by years from blood collection to disease diagnosis (<5, 5–9, 10 or more; in years). ΔCt country-specific (Danish and non-Danish) tertiles were obtained based on the distribution of ΔCt values in controls. Tests for trend in tertiles were performed using categories 1, 2 and 3, respectively. Unadjusted odds ratios (OR) and 95% confidence intervals (95% CI) for one unit ΔCt values and for tertiles of ΔCt were estimated with conditional logistic regression. Percentage of missing for all variables was <4% in both pre-CLL and controls. Adjustment for the selected risk factors did not change the OR more than 7%; therefore, further adjustment was not considered (Supporting Information Data S2).

Receiver operating characteristic (ROC) curves analysis and the area under the curve (AUC) were used to evaluate the accuracy of each miRNA to discriminate controls from pre-CLL. A threefold cross-validation was performed using the R library *caret* in individuals with their diagnosis within 5-year of recruitment.

As sensitivity analysis, we tested heterogeneity in ΔCt values of the miRNA among countries by means of Cochran's Q test. To take into account the potential impact of the flooding in Denmark, we examined the differences in the medians of ΔCt values between the 218 Danish and 230 non-Danish samples using Mann–Whitney test and also examined the heterogeneity of the results stratifying by Danish and non-Danish countries. Finally, we ran again all analyses excluding extreme values defined by mean ΔCt value ± 3 standard deviations. Two-sided *p*-values were reported with statistical significance set at $p < 0.05$. All analyses were performed by using STATA statistical software, version 15 (Stata Corporation, College Station, TX) and R v3.4.0.

### Data availability

For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

## Results

### Study sample description

Most of the 224 paired samples came from Denmark ($n = 109$, 44%; Table 1). The mean age at recruitment was 57 years and the male/female ratio was 50%. The median follow-up time between blood collection and diagnosis of CLL was 10 years (p25–p75: 7–13 years). No significant differences in the potential confounders were observed between pre-CLL and controls. Regarding raw miRNA serum levels, there were strong differences in ΔCt means of *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-29a-3p* (*p*-values < 0.001), but not *hsa-miR-16-5p* and *hsa-miR-223-3p* between pre-CLL and their matched controls. No association was observed for the selected sociodemographic variables (smoking, education, BMI, physical activity, height and alcohol) and miRNA levels at recruitment (data not shown). In relation to the matching variables, statistically significant differences between Danish and non-Danish crude expression levels were observed for all miRNA with higher median ΔCt values for *hsa-miR-16-5p*, *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-29a-3p* and inversely for *hsa-miR-223-3p* (Supporting Information Data S3). Samples were collected over a 6-year period (1993–1998); however, no differences in miRNA levels were observed by year of sample collection (data not shown). Overall, younger individuals had higher *hsa-miR-223-3p* and lower *hsa-miR-155-5p* levels. Females tended to have higher *hsa-miR-223-3p* and lower *hsa-miR-155-5p*, *hsa-miR-16-5p and hsa-miR-29a-3p* crude levels than males (Supporting Information Data S3).

### MiRNA correlation

High correlations were observed between *hsa-miR-150-5p* and *hsa-miR-155-5p* (ρ = 0.82) and between *hsa-miR-29a-3p* and *hsa-miR-150-5p* and *hsa-miR-155-5p* (ρ = 0.75 and 0.66, respectively; data not shown). These results were reflected by the two components of the PCA explaining 79% of the miRNAs expression levels variability (Fig. 1). PC1 was dominated by *hsa-miR-150-5p*,*hsa-miR-155-5p*, *hsa-miR-29a-3p* and to a lesser extent by *hsa-miR16-5p* while PC2, mainly defined by *hsa-miR223-3p*, explained 25% of the variability. CLL diagnosed in <5 years from blood collection had high values of PC1.

**Table 1.** Demographic and lifestyle characteristics of pre-CLL cases and their matched controls at recruitment

| | Controls | Pre-CLL | $p^1$ |
|---|---|---|---|
| Overall (*n*) | 224 | 224 | |
| Female, *n* (%) | 113 (50) | 113 (50) | – |
| Country, *n* (%) | | | – |
| Denmark | 109 (49) | 109 (49) | |
| United Kingdom | 31 (14) | 31 (14) | |
| Italy | 24 (11) | 24 (11) | |
| Spain | 22 (109) | 22 (10) | |
| Germany | 15 (7) | 15 (7) | |
| The Netherlands | 19 (8) | 19 (8) | |
| Greece | 4 (2) | 4 (2) | |
| Age at recruitment (years), mean (SD) | 56.6 (7.1) | 56.6 (7.1) | |
| ≤53 | 70 (31) | 74 (33) | |
| 54–59 | 69 (31) | 65 (29) | |
| ≥60 | 85 (38) | 85 (38) | 0.89 |
| Secondary school or higher education, *n* (%) | 72 (34) | 61 (29) | 0.24 |
| Height (cm)[2], mean (SD) | 168.0 (9.34) | 168.7 (9.24) | 0.48 |
| Body mass index (kg/m$^2$), mean (SD) | 25.9 (3.71) | 26.7 (4.18) | 0.08 |
| Total energy | 2,150.1 (607.88) | 2,161.3 (613.39) | 0.66 |
| Physically inactive *n* (%) | 40 (18) | 38 (17) | 0.80 |
| Current smoker, *n* (%) | 54 (24) | 63 (28) | 0.33 |
| Alcohol intake (g/day), mean (SD) | 15.8 (17.81) | 17.6 (21.33) | 0.75 |
| *hsa-miR-16-5p*[3], ΔCt mean (SD) | 4.8 (1.19) | 4.8 (1.09) | 0.39 |
| *hsa-miR-150-5p*[3], ΔCt mean (SD) | −1.04 (1.23) | −0.65 (1.46) | 0.002 |
| *hsa-miR-155-5p*[3], ΔCt mean (SD) | −6.57 (1.11) | −6.08 (1.6) | 0.003 |
| *hsa-miR-223-3p*[3], ΔCt mean (SD) | 4.77 (0.81) | 4.73 (0.89) | 0.20 |
| *hsa-miR-29a-3p*[3], ΔCt mean (SD) | −2.44 (0.8) | −2.14 (1.1) | 0.001 |

[1]*p* of heterogeneity for categorical variables and *t*-test or median test for continuous variables.
[2]Anthropometric data were adjusted to reduce heterogeneity due to protocol differences in clothing worn during measurement.
[3]Normalized by *hsa-miR-24-5p*.
% of missing for all variables was less than 2% in both pre-CLL and controls except for educational levels (4 and 3% in controls and pre-CLL, respectively).
Abbreviations: CLL, chronic lymphocytic leukemia; SD, standard deviation.

### Serum *hsa-miR-150-5p*, *hsa-miR-155-5p* and *hsa-miR-29a-3p* levels elevated years before CLL diagnosis

At recruitment, pre-CLL were more likely to have higher ΔCt values of *hsa-miR-29a-3p*, *hsa-miR-150-5p* and *hsa-miR-155-5p* than controls ($OR_{1\Delta Ct\ unit\ increase}$: 1.75, 95% CI: 1.31–2.34; OR:1.42, 95% CI: 1.18–1.72 and OR: 1.64, 95% CI: 1.31–2.04, respectively; Fig. 2 and Supporting Information Data S2) whereas no associations were observed with *hsa-miR-223-3p* and *hsa-miR-16-5p*. Following mutual adjustment for the other four miRNAs, *hsa-miR-150-3p* did not remain associated with CLL ($OR_{1\Delta Ct\ unit\ increase}$: 0.85, 95% CI: 0.60–1.20; Supporting Information Data S2). Evaluating the model including the two principal components, only PC1 was associated with CLL risk ($OR_{1\ unit\ increase}$: 1.44; 95% CI: 1.21–1.72; and $OR_{1\ unit\ increase}$: 0.93; 95% CI: 0.71–1.23; for PC1 and PC2, respectively; Supporting Information Data S4).

By time from recruitment to diagnosis, the association between serum *hsa-miR-29a-3p*, *hsa-miR-150-5p* and *hsa-miR-155-5p* levels and pre-CLL strengthened with decreasing time to diagnosis ($OR_{<5\ years\ of\ CLL}$ = 3.38, 95% CI = 1.44–7.90; OR = 2.18, 95% CI = 1.24–3.38; OR = 2.32, 95% CI = 1.30–4.12, respectively; Fig. 3 and Supporting Information Data S4). An increased risk of CLL was also observed in individuals diagnosed more than 10 years after recruitment, but estimates did not reach statistical significance. No interactions by age at recruitment and sex were detected (data not shown).

### Modest predictive performances of miRNAs

Overall, the five miRNAs (alone or altogether) had modest predictive performances and limited ability to discriminate between pre-CLL and controls (AUC ranged from 0.52 to 0.62, Table 2). *Hsa-miR-155-5p*, *hsa-miR-150-5p* and *hsa-miR-29a-3p* as well as both PCs gained good predictive performances with decreasing time to CLL diagnosis reaching AUCs within 5 years of CLL diagnosis of around 0.80/0.90. However,

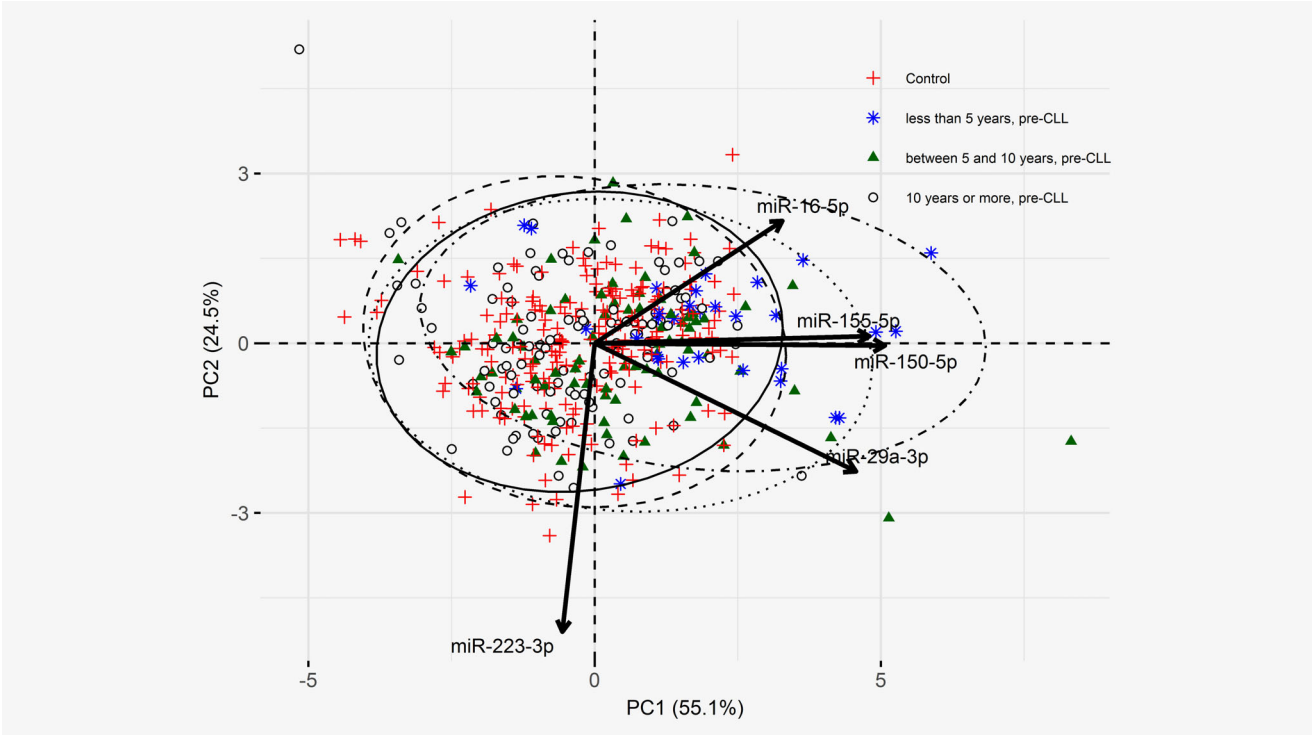


**Figure 1.** Principal component analysis of miRNA normalized ΔCt values. The two principal components obtained were: PC1, highly loaded on *hsa-miR-150-5p* (32%), *hsa-miR-155-5p* (29%), *hsa-miR-29a-3p* (26%) and to a lesser extent *hsa-miR-16-5p* (13%), explained 55% of the variability; PC2, highly loaded on *hsa-miR-223-3p* (−73%), explained 25% of the variability. 95% confidence ellipses are represented by solid line (controls), dash-dotted line (<5 years, pre-CLL), dotted line (between 5 and 10 years, pre-CLL), dashed line (10 years or more, pre-CLL). [Color figure can be viewed at wileyonlinelibrary.com]

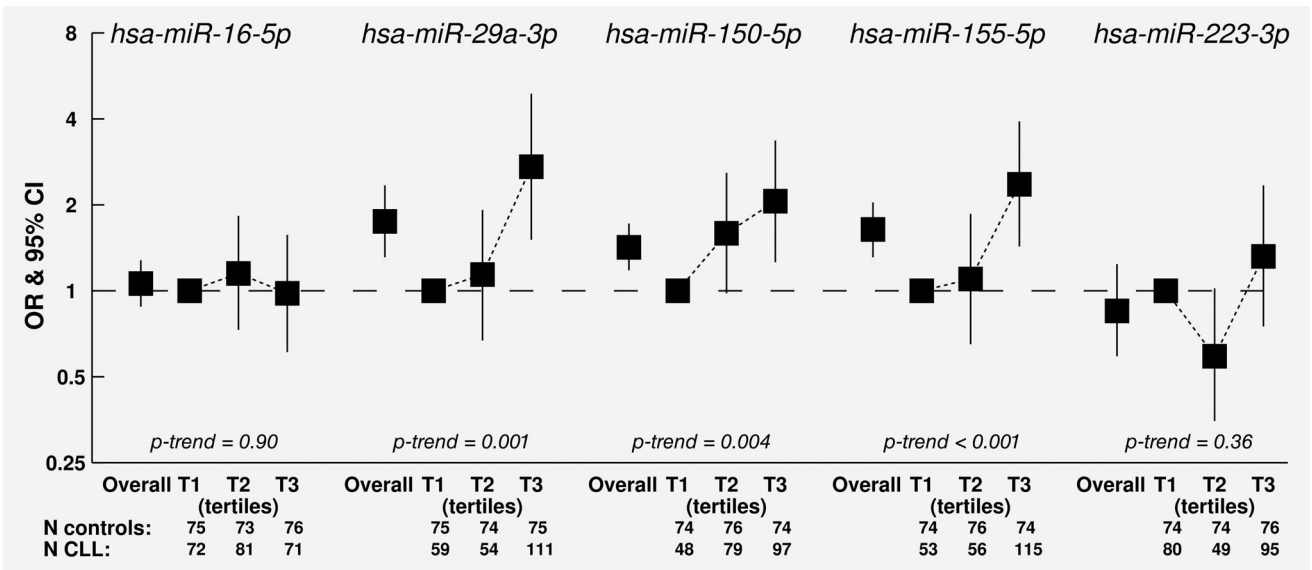


**Figure 2.** Odds ratio of CLL in relation to prediagnostic levels of the five selected miRNAs, by tertile distribution at recruitment. Abbreviations: CI, confidence interval; CLL, chronic lymphocytic leukemia; miR, microRNA; *n*, number; OR, odds ratio; *p*-trend, *p*-value for trend. 1: Tertiles (T) based on country-specific distribution of hsa-miR expression in controls (Danish and non-Danish origin). ORs were estimated using conditional logistic regression models (each case was matched by sex, study center, age at blood collection (1 year), date of blood collection (±3 months), time of blood collection (±3 hr) and fasting at blood collection (no, yes)). Data of the graph can be found in Supporting Information Data S2. Overall test for trend was based on a 1 unit ΔCt increase and tests for trend across tertiles were done using categories 1, 2 and 3 for each tertile, respectively.

following a threefold cross-validation analysis the prediction capacity for this time period dropped sharply for all miRNAs ($AUC_{<5\ years} < 0.61$; Table 2).

### Sensitivity analyses

No heterogeneity by country for the analysis on CLL and 1-ΔCt value increase for any of the five miRNA was detected (All $p$ values for heterogeneity >0.10; Supporting Information Data S5) and the results were not materially modified after exclusion of extreme values (representing <1% of each miRNA; data not shown).

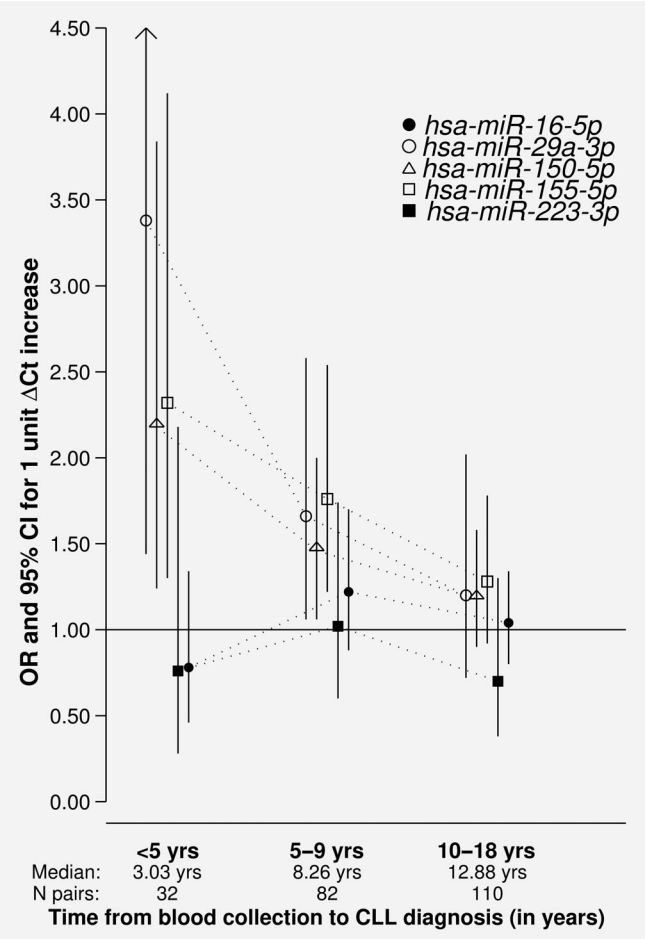


**Figure 3.** Odds ratio of CLL in relation to prediagnostic levels of the five selected miRNAs by time (years) from blood collection to CLL. Abbreviations: CI, confidence interval; CLL, chronic lymphocytic leukemia; miR, microRNA; *n*, number; OR, odds ratio; yrs, years. ORs were estimated using conditional logistic regression models (1:1 matched by sex, study center, age at blood collection (1 year), date of blood collection (±3 months), time of blood collection (±3 hr) and fasting at blood collection (no, yes)). Data of the graph can be found in Supporting Information Data S4.

**Table 2.** Area under the curve (AUC) for each miRNA by time (years) from blood collection to CLL

| | *N* pair | hsa-miR-16-5p | hsa-miR-29a-3p | hsa-miR-150-5p | hsa-miR-155-5p | hsa-miR-223-3p | All miRs | Both PC[1] |
|---|---|---|---|---|---|---|---|---|
| Overall | | 0.53 (0.47–0.58) | 0.59 (0.54–0.65) | 0.59 (0.54–0.65) | 0.61 (0.56–0.67) | 0.52 (0.47–0.57) | 0.62 (0.57–0.68) | 0.58 (0.53–0.63) |
| Time from sample collection to diagnosis (years) | | | | | | | | |
| ≥10 | 110 | 0.55 (0.47–0.63) | 0.56 (0.48–0.63) | 0.56 (0.48–0.63) | 0.57 (0.50–0.65) | 0.55 (0.47–0.762) | 0.59 (0.52–0.67) | 0.57 (0.49–0.65) |
| Between 5 and 10 | 82 | 0.55 (0.46–0.64) | 0.61 (0.52–0.70) | 0.61 (0.53–0.70) | 0.65 (0.56–0.73) | 0.53 (0.44–0.62) | 0.65 (0.57–0.74) | 0.65 (0.56–0.73) |
| ≤5 | 32 | 0.65 (0.51–0.78) | 0.83 (0.72–0.93) | 0.79 (0.68–0.91) | 0.79 (0.68–0.90) | 0.59 (0.45–0.73) | 0.88 (0.80–0.96) | 0.82 (0.72–0.93) |
| ≤5[2] | 32 | 0.42 (0.25–0.60) | 0.61 (0.47–0.75) | 0.59 (0.56–0.62) | 0.56 (0.46–0.67) | 0.43 (0.33–0.52) | 0.52 (0.38–0.66) | 0.49 (0.29–0.59) |

ROC curves adjusted by match variables: sex, study center, age at blood collection (1 year), date of blood collection (±3 months), time of blood collection (±3 hr) and fasting at blood collection (no, yes).
[1]PC1, highly loaded on *hsa-miR-150-5p* (32%), *hsa-miR-155-5p* (29%), *hsa-miR-29a-3p* (26%) and to a lesser extent *hsa-miR-16-5p* (13%), explained 55% and PC2, highly loaded on *hsa-miR-223-3p* (~73%), explained 25% of the variability.
[2]Threefold cross-validation.
Abbreviations: CLL, chronic lymphocytic leukemia; ROC, receiver operating characteristic; PC, principal component.

## Discussion

In this case–control study nested within the prospective EPIC study, circulating *hsa-miRNA-29a-3p*, *hsa-miR-150-5p* and *hsa-miR-155-5p* were found to be deregulated up to 10 years before CLL diagnosis. Our results lend further support for a role in disease progression of these three miRNAs in early CLL stages. However, these biomarkers were suboptimal to discriminate CLL from controls. No difference was observed in *hsa-miRNA-16-5p* and *hsa-miR-223-3p* expression between pre-CLL and controls.

Both *hsa-miR-150-5p* and *hsa-miR-155-5p* have been intensively studied in CLL and immune system deregulation.[19,20] Their expression has been found to increase with B-cell receptor (BCR) and T cell activation, hematopoiesis and inflammation. In particular, *hsa-miR-155-5p* has been associated with the regulation of activation-induced deaminase (AID)-mediated oncogenic translocations and class-switched B cells.[21] Furthermore, *hsa-miR-150-5p* could inhibit expression of genes associated with BCR signaling like forkhead box P1 (FOXP1) and GRB2-associated binding protein 1 (GAB1).[22] Upregulation of *hsa-miR-155-5p* in CLL cells and serum is well-documented[23,24] and was also found to be a valid prognosis marker for CLL, independently of the established markers of poor clinical outcome: ZAP-70 and IgHV mutational status.[25] Likewise, higher levels of *hsa-miR-150-5p* were also detected in serum[24,26] and plasma[27] of CLL patients compared to healthy individuals. Upregulation of *hsa-miR-150-5p* was associated with some (i.e., CD38+), but not all (ZAP-70) prognostic markers in serum.[26] In contrast, Moussay *et al.* reported an association between higher levels of hsa-miR-150-5p and disease severity as well as ZAP70 positive patients.[27] The high correlation between *hsa-miR-150-5p* and *hsa-miR-155-5p* levels suggests that these molecules might act jointly in CLL to enhance tumor growth. Nevertheless, after mutual adjustment only the association between CLL and *hsa-miR-155-5p* remained, suggesting a stronger role of *hsa-miR-155-5p* than *hsa-miR-150-5p*. However, since the two miRNA were highly correlated, the mutual adjustment may represent overadjustement. Compared to circulating soluble CD23, a strong predictor of subsequent CLL risk,[28] neither *hsa-miR-155-5p* nor *hsa-miR-150-5p* levels were good to discriminate controls from CLL. The increasing ORs with decreasing time to diagnosis might reflect the slowly developing disease as reported by Landgren *et al.*[29] with the presence of $MBL^{hi}$ 6.4 years before CLL diagnosis and might indicate an increase in the clonal B-cell population.[30]

In previous studies, *hsa-miR-155-5p* was overexpressed in B-cells of individuals with $MBL^{hi}$.[31] However, to the best of our knowledge, there is no data looking at serum before CLL diagnosis. Concurrently, two groups have reported at the American Society of Hematology 2018 meeting higher risk of infections as well as higher mortality rates in individuals with $MBL^{lo}$ compared to the general population as well as to age-sex matched individuals without $MBL^{lo}$.[32,33] CLL is associated with the deterioration of CD8+ T-cell function, the so-called "T-cell exhaustion",[34] and in turn with higher risk of infections[35,36] leading to a vicious circle of infections, host susceptibility and worsened host immune system.[37] Interestingly, *hsa-miR-155-5p* was found to regulate the CD8+ T-cell exhaustion in chronic and acute bacterial and viral infections.[38–40] Likewise, *hsa-miR-150-5p* was found to regulate memory CD8+ T cell through the transcription factor MYB(c-Myb)[41] and forkhead box O1 (Foxo1).[42] It would have been interesting to examine the association of miRNAs levels and the personal history of infectious diseases of the participants but, the EPIC study does not have this information.

Compared to *hsa-miR-150-5p* and *hsa-miR-155-5p*, little is known on the association between *hsa-miR-29a-3p* and CLL. We found here that *hsa-miR-29a-3p* showed a strong and stable association with CLL even following mutual adjustment for other miRNAs. The deregulation of the *miR-29* family has been mainly described as tumor suppressor for many cancers and has been implicated in the regulation of the acquired immune system.[43] Using 94 samples of CLL cells, *Calin et al.*[15] reported that a miRNA signature of 13 miRNAs, including *miR-29* family (*miR*-29a, *miR*-29b, *miR*-29c, all downregulated), could discriminate between CLL samples with poor (expression of ZAP-70 and unmutated IgHV) and good (no expression of ZAP-70 and mutated IgHV) prognosis. To the best of our knowledge and in line with our pilot study, only one study examining serum miRNA levels in 22 CLL patients and eight healthy volunteers reported higher levels of *hsa-miR-29a-3p* compared to controls.[24] T-cell leukemia/lymphoma 1 (TCL1) plays a central role in lymphogenesis[44] and Pekarsky *et al.* reported that TCL1 overexpressed in CLL might be regulated by the miR-29 family.[45] Our findings from the EPIC prospective study support *hsa-miR-29a-3p* as a possible player in CLL development, but again the biomarker does not provide good predictive power. This result warrants further investigation to fully understand the role of *miR-29* family in CLL. In particular, miR-29c that has been associated with CLL severity and different clinical subgroups should also be included in future investigation.

Interestingly, contrary to our hypothesis and assuming that miRNA expression in serum reflects the oncogenic activity, circulating *hsa-miR-223-3p* was not deregulated before diagnosis suggesting that this molecule might not be involved in early disease stages. Indeed, the downregulation of *hsa-miR-223-3p* has been associated with poor prognostic factors[46] and with a subsequent development of autoimmune hemolytic anemia in CLL.[47] Hence, *hsa-miR-223-3p* might act only as a prognosis marker. Finally, *hsa-miR-16-5p* as well as *hsa-miR-15-5p*, encoded within the commonly deleted locus in CLL 13q14 (seen in around 55%), have tumor-suppressor function associated with BCL2.[48] In accordance with our pilot study no differences were observed between pre-CLL and controls for these miRNAs.

The five selected miRNA had poor CLL prediction capacity both when the miRNA were examined separately or jointly.

Classification accuracy for *hsa-miR-29a-3p*, *hsa-miR-150-5p* and *hsa-miR-155-5p* increased when we restricted to pre-CLL recruited <5 years before diagnosis but these findings did not remain after using cross-validation. Since the use of this technique in small sample size is controversial,[49] external validation is warranted.

The strength of our study includes the use of a very well-characterized prospective study with samples taken up to 18 years before diagnosis. Also, the relative large number of CLL cases and the inclusion of well-matched controls are valuable assets compared to other published work on circulating miRNAs. Extracellular vesicles such as exosomes are key players in cell–cell communication microenvironment and miRNA transport. While the detection of circulating miRNA does not inform us on the cell origin of miRNAs, our findings suggest that serum miRNA levels, easily obtained compared to the expensive and time-consuming process of exosome purification, might be a valid way to examine miRNA alteration and to reflect miRNA expression in tumor cells. The main limitation of our study is the lack of information on clinical data for the CLL cases such as ZAP-70 status, IGVH mutational status, CD38 expression, cytogenetics (13q(del), 11(del), 17(del), trisomy 12), driver mutations (NOTCH1, SF3B1 or BIRC3) or disease stages (Rai or Binet) to identify subgroups of CLL affected by deregulation of these miRNAs prior to diagnosis. Also, information on main CLL risk factors such as a family history of hematological cancer was not collected in the EPIC study. Moreover, CLL and $MBL^{hi}$ are often detected by chance in routine blood test and we cannot exclude that among the participants included in the control group single individuals might have had CLL or MBL. Also, MBL could precede CLL and a clonal population might have been present at recruitment. Finally, we focused this work on five promising miRNA but further work on the interaction between a larger number of miRNA is crucial.

In conclusion, this large prospective study identifying the deregulation of circulating *hsa-miR-29a-3p*, *hsa-miR-150-5p* and *hsa-miR-155-5p* before CLL diagnosis provides strong support for an association between these molecules and CLL. The added value of these biomarkers for disease discrimination before diagnosis, however, seems limited. Future studies should assess these circulating biomarkers in both $MBL^{hi}$ and $MBL^{lo}$ to monitor and to understand the progression from early stages to full cancer development. A better understanding of the interplay between these miRNAs and the host immune system might open new avenues for more effective therapies.

## Acknowledgements

We thank all participants of the EPIC study, Bertrand Hémon at IARC for his valuable work and technical support with the EPIC database. The authors would like to thank the participants of the EpiLymph-Spain study as well as Dr Eva Domingo (Catalan Institute of Oncology, l'Hospitalet de Llobregat, Spain) and Dr Ramón Bosch (Hospital de Tortosa Verge de la Cinta, Tortosa, Spain) for their support in patient recruitment. The Umeå-EPIC node at Umeå University, Sweden is acknowledged for their valuable contribution and ongoing support to the EPIC Study. We thank the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands, for their contribution and ongoing support to the EPIC Study. The EPIC-Norfolk study (DOI 10.22025/2019.10.105.00004) is grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research. Our study was supported by the : European Commission (DG-SANCO), International Agency for Research on Cancer, Danish Cancer Society (Denmark), Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Generale de l'Education Nationale, Institut National de la Sante et de la Recherche Medicale (INSERM) (France), German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany), the Hellenic Health Foundation (EPIC-Greece), Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy), Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands, for their contribution and ongoing support to the EPIC Study, World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands; ERC2009-AdG 232997), German Federal Ministry of Education and Research (BMBF 01EO1303), Spanish Ministry of Economy and Competitiveness—Carlos III Institute of Health cofunded by FEDER funds/European Regional Develpment Fund (ERDF)—a way to build Europe (PI13/00061 [to Granada], PI13/01162 [to EPICMurcia, Regional Governments of Andalucıa, Asturias, Basque Country, Murcia and Navarra], PI17/01280 and PI14/01219 [to Barcelona]), Centro de Investigación Biomédica en Red: Epidemiología y Salud Pública (CIBERESP, Spain), Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), CERCA Programme/Generalitat de Catalunya for institutional support (2017SGR1085), Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Vasterbotten (Sweden), Cancer Research UK (14136 [to EPIC-Norfolk], C8221/A19170 [to EPIC-Oxford]) and Medical Research Council Canada (MR/N003284/1, MC-PC_13048 and MC-UU_12015/1 [to EPIC-Norfolk study; doi: 10.22025/2019.10.105.00004), MR/M012190/1 [to EPIC-Oxford, UK]).

## Conflict of interest

The authors declare no competing financial interests.

## References

1. Sant M, Minicozzi P, Mounier M, et al. Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EUROCARE-5, a population-based study. *Lancet Oncol* 2014;15:931–42.
2. Chiorazzi N, Dameshek W, Rai KR, et al. Cell proliferation and death: forgotten features of chronic lymphocytic leukemia B cells. *Best Pract Res Clin Haematol* 2007;20:399–413.
3. Rawstron AC, Bennett FL, O'Connor JMS, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med* 2008;359:575–83.
4. Nieto WG, Almeida J, Romero A, et al. Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B-cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach. *Blood* 2009;114:33–7.
5. Lawrie CH. MicroRNAs and haematology: small molecules, big function. *Br J Haematol* 2007;137:503–12.
6. Storz G. An expanding universe of noncoding RNAs. *Science (80- )* 2002;296:1260–3.

7. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA* 2002;99:15524–9.
8. Pekarsky Y, Balatti V, Croce CM. BCL2 and miR-15/16: from gene discovery to treatment. *Cell Death Differ* 2018;25:21–6.
9. Lawrie CH, Gal S, Dunlop HM, et al. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol* 2008;141:672–5.
10. Lawrie CH. MicroRNAs in hematological malignancies. *Blood Rev* 2013;27:143–54.
11. Larrea E, Sole C, Manterola L, et al. New concepts in Cancer biomarkers: circulating miRNAs in liquid biopsies. *Int J Mol Sci* 2016;17:627.
12. Solé C, Arnaiz E, Lawrie CH. MicroRNAs as biomarkers of B-cell lymphoma. *Biomark Insights* 2018;13:1177271918806840.
13. van Roosbroeck K, Calin GA, Author SO. MicroRNAs in CLL: miRacle or miRage for prognosis and targeted therapies? HHS public access Author manuscript. *Semin Oncol* 2016;43:209–14.
14. Yeh Y-Y, Ozer HG, Lehman AM, et al. Characterization of CLL exosomes reveals a distinct microRNA signature and enhanced secretion by activation of BCR signaling. *Blood* 2015;125:3297–305.
15. Calin GA, Ferracin M, Cimmino A, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005;353:1793–801.
16. Riboli E, Kaaks R. The EPIC project: rationale and study design. European prospective investigation into Cancer and nutrition. *Int J Epidemiol* 1997;26(Suppl 1):S6–14.
17. 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;5:1113–24.
18. de Sanjose S, Shah KV, Domingo-Domenech E, et al. Lack of serological evidence for an association between simian virus 40 and lymphoma. *Int J Cancer* 2003;104:522–4.
19. He Y, Jiang X, Chen J. The role of miR-150 in normal and malignant hematopoiesis. *Oncogene* 2014;33:3887–93.
20. Due H, Svendsen P, Bødker JS, et al. miR-155 as a biomarker in B-cell malignancies. *Biomed Res Int* 2016;2016:1–14.
21. Vigorito E, Kohlhaas S, Lu D, et al. miR-155: an ancient regulator of the immune system. *Immunol Rev* 2013;253:146–57.
22. Mraz M, Chen L, Rassenti LZ, et al. miR-150 influences B-cell receptor signaling in chronic lymphocytic leukemia by regulating expression of GAB1 and FOXP1. *Blood* 2014;124:84–95.
23. Fulci V, Chiaretti S, Goldoni M, et al. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood* 2007;109:4944–51.
24. Filip AA, Grenda A, Popek S, et al. Expression of circulating miRNAs associated with lymphocyte differentiation and activation in CLL—another piece in the puzzle. *Ann Hematol* 2016;96:33–50.
25. Cui B, Chen L, Zhang S, et al. MicroRNA-155 influences B-cell receptor signaling and associates with aggressive disease in chronic lymphocytic leukemia. *Blood* 2014;124:546–54.
26. Stamatopoulos B, Van Damme M, Crompot E, et al. Opposite prognostic significance of cellular and serum circulating microRNA-150 in chronic lymphocytic leukemia patients. *Mol Med* 2015;21:123–33.
27. Moussay E, Wang K, Cho J-H, et al. MicroRNA as biomarkers and regulators in B-cell chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 2011;108:6573–8.
28. Kaaks R, Sookthai D, Luczynska A, et al. Lag times between lymphoproliferative disorder and clinical diagnosis of chronic lymphocytic leukemia: a prospective analysis using plasma soluble CD23. *Cancer Epidemiol Biomarkers Prev* 2015;24:538–45.
29. Landgren O, Albitar M, Abbasi F, et al. B-cell clones as early markers for chronic lymphocytic leukemia. *N Engl J Med* 2009;360:659–67.
30. Vardi A, Dagklis A, Scarfo L, et al. Immunogenetics shows that not all MBL are equal: the larger the clone, the more similar to CLL. *Blood* 2013;121:4521–8.
31. Ferrajoli A, Shanafelt TD, Ivan C, et al. Prognostic value of miR-155 in individuals with monoclonal B-cell lymphocytosis and patients with B chronic lymphocytic leukemia. *Blood* 2013;122:1891–9.
32. Orfao A, Criado I, Rodríguez-Caballero A, et al. Low-count monoclonal B-cell lymphocytosis persists after 7 years of follow-up and confers a higher risk of death. *Blood* 2017;130:4299.
33. Shanafelt T, Kay N, Hanson C, et al. Prevalence of low count (LC) monoclonal B cell lymphocytosis (MBL) and serious infections in a population-based cohort of U.S. adults participating in a large bio-repository. *Blood* 2017;130:831.
34. Riches JC, Davies JK, McClanahan F, et al. T cells from CLL patients exhibit features of T-cell exhaustion but retain capacity for cytokine production. *Blood* 2013;121:1612–21.
35. Anderson LA, Landgren O, Engels EA. Common community acquired infections and subsequent risk of chronic lymphocytic leukaemia. *Br J Haematol* 2009;147:444–9.
36. Landgren O, Rapkin JS, Caporaso NE, et al. Respiratory tract infections and subsequent risk of chronic lymphocytic leukemia. *Blood* 2007;109:2198–201.
37. Beldomenico PM, Begon M. Disease spread, susceptibility and infection intensity: vicious circles? *Trends Ecol Evol* 2010;25:21–7.
38. Gracias DT, Stelekati E, Hope JL, et al. The microRNA miR-155 controls CD8(+) T cell responses by regulating interferon signaling. *Nat Immunol* 2013;14:593–602.
39. Tsai C-Y, Allie SR, Zhang W, et al. MicroRNA miR-155 affects antiviral effector and effector memory CD8 T cell differentiation. *J Virol* 2013;87:2348–51.
40. Lind EF, Elford AR, Ohashi PS. Micro-RNA 155 is required for optimal CD8+ T cell responses to acute viral and intracellular bacterial challenges. *J Immunol* 2013;190:1210–6.
41. Chen Z, Stelekati E, Kurachi M, et al. miR-150 regulates memory CD8 T cell differentiation via c-Myb. *Cell Rep* 2017;20:2584–97.
42. Ban YH, Oh S-C, Seo S-H, et al. miR-150-mediated Foxo1 regulation programs CD8 + T cell differentiation. *Cell Rep* 2017;20:2598–611.
43. Schmitt MJ, Margue C, Behrmann I, et al. MiRNA-29: a microRNA family with tumor-suppressing and immune-modulating properties. *Curr Mol Med* 2013;13:572–85.
44. Virgilio L, Narducci MG, Isobe M, et al. Identification of the TCL1 gene involved in T-cell malignancies. *Proc Natl Acad Sci USA* 1994;91:12530–4.
45. Pekarsky Y, Santanam U, Cimmino A, et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. *Cancer Res* 2006;66:11590–3.
46. Stamatopoulos B, Meuleman N, Haibe-Kains B, et al. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. *Blood* 2009;113:5237–45.
47. Ferrer G, Navarro A, Hodgson K, et al. MicroRNA expression in chronic lymphocytic leukemia developing autoimmune hemolytic anemia. *Leuk Lymphoma* 2013;54:2016–22.
48. Mertens D, Stilgenbauer S. CLL and deletion 13q14: merely the miRs? *Blood* 2012;119:2974–5.
49. Isaksson A, Wallman M, Göransson H, et al. Cross-validation and bootstrapping are unreliable in small sample classification. *Pattern Recognit Lett* 2008;29:1960–5.