

## ARTICLE OPEN



# Lifetime exposure to brominated trihalomethanes in drinking water and swimming pool attendance are associated with chronic lymphocytic leukemia: a Multicase-Control Study in Spain (MCC-Spain)

Carolina Donat-Vargas<sup>1,2,3,4</sup>, Manolis Kogevinas<sup>1,2,3,5</sup>, Yolanda Benavente<sup>3,6</sup>, Laura Costas<sup>3,6</sup>, Elias Campo<sup>7</sup>, Gemma Castaño-Vinyals<sup>1,2,3,5</sup>, Guillermo Fernandez-Tardon<sup>3,8</sup>, Javier Llorca<sup>3,9</sup>, Inés Gómez-Acebo<sup>3,9,10</sup>, Nuria Aragonés<sup>3,11</sup>, Marina Pollán<sup>3,12</sup>, Delphine Casabonne<sup>3,6</sup> and Cristina M. Villanueva<sup>1,2,3,5</sup>✉

© The Author(s) 2023

**BACKGROUND:** Chronic lymphocytic leukemia (CLL) etiology is poorly understood, and carcinogenic chemicals in drinking and recreational water are candidates.

**OBJECTIVE:** To evaluate the association between drinking-water exposure to trihalomethanes (THMs) and nitrate as well as lifetime swimming pool attendance and CLL.

**METHODS:** During 2010–2013, hospital-based CLL cases and population-based controls were recruited in Spain, providing information on residential histories, type of water consumed and swimming pool attendance. Average THMs and nitrate levels in drinking water were linked to lifetime water consumption. Odds ratios (OR) and 95% confidence intervals (CI) were estimated using mixed models.

**RESULTS:** Final samples for residential tap water analyses and swimming pool attendance analyses were 144 cases/1230 controls and 157 cases/1240 controls, respectively. Mean (SD) values for average lifetime residential brominated THMs and chloroform in tap water ( $\mu\text{g/L}$ ), and ingested nitrate (mg/day) were 48.1 (35.6), 18.5 (6.7) and 13.7 (9.6) respectively in controls; and 72.9 (40.7), 17.9 (5.4), and 14.1 (8.8) in CLL cases. For each 10  $\mu\text{g/L}$  increase of brominated THMs and chloroform lifetime-average levels, the ORs (95% CI) were 1.22 (1.14, 1.31) and 0.54 (0.34, 0.87), respectively. For each 5 mg/day increase of ingested nitrate, the OR of CLL was 0.91 (0.80, 1.04). The OR of lifetime pool users (vs. non-users) was 2.38 (1.61, 3.52). Upon performing annual frequency of attending pools analysis through categorization, the second and third categories showed an ORs of 2.36 (1.49, 3.72) and 2.40 (1.51, 3.83), respectively, and P-trend of 0.001.

**IMPACT STATEMENT:** This study identifies an association of long-term exposure to THMs in drinking water, at concentrations below the regulatory thresholds and WHO guidelines, and swimming pool attendance, with chronic lymphocytic leukemia (CLL). These unprecedented findings are highly relevant since CLL is an incurable cancer with still unknown etiology and because the widespread exposure to chlorination by-products that remain in drinking and recreational water worldwide. Despite the demonstrated carcinogenicity in animals of several chlorination by-products, little is known about their potential risks on human health. This study makes a significant contribution to the search for environmental factors involved in the etiology of CLL and to the evidence of the health impact of these high prevalent water contaminants.

**Keywords:** Water contaminants; Trihalomethanes; Nitrate; swimming pools; Cancer; Chronic lymphocytic leukemia

*Journal of Exposure Science & Environmental Epidemiology*; <https://doi.org/10.1038/s41370-023-00600-7>

<sup>1</sup>ISGlobal, Barcelona, Spain. <sup>2</sup>Universitat Pompeu Fabra (UPF), Barcelona, Spain. <sup>3</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. <sup>4</sup>Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. <sup>5</sup>IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. <sup>6</sup>Unit of Molecular and Genetic Epidemiology in Infections and Cancer (UNIC-Molecular), Cancer Epidemiology Research Programme, IDIBELL, Catalan Institute of Oncology, 08908 L'Hospitalet de Llobregat, Spain. <sup>7</sup>Haematopathology Section, Hospital Clinic of Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), University of Barcelona, Centre d'Investigació Biomèdica en Red de Càncer (CIBERONC), Barcelona, Spain. <sup>8</sup>Health Research Institute of Asturias, ISPA, Oviedo, Spain. <sup>9</sup>Faculty of Medicine, University of Cantabria, Santander, Spain. <sup>10</sup>IDIVAL. Instituto de Investigación Sanitaria Valdecilla, 39011 Santander, Spain. <sup>11</sup>Epidemiology Section, Public Health Division, Department of Health of Madrid, Madrid, Spain. <sup>12</sup>Cancer and Environmental Epidemiology Unit, National Centre for Epidemiology, Carlos III Institute of Health, Madrid, Spain. ✉email: [cristina.villanueva@isglobal.org](mailto:cristina.villanueva@isglobal.org)

Received: 4 April 2023 Revised: 22 August 2023 Accepted: 23 August 2023

Published online: 19 September 2023

## INTRODUCTION

Chronic lymphocytic leukemia (CLL) is a hematologic neoplasm of memory B cells characterized by the clonal proliferation and accumulation of mature and typically CD5-positive B cells within the blood, bone marrow, lymph nodes, and/or spleen [1]. It is the most common type of leukemia in developed countries, with a current age-adjusted incidence of ~5 per 100 000 population [2]. CLL is more common in males than in females, in the elderly, with a mean age of diagnosis around 73 years, and, unlike other types of leukemia, CLL appears to be unique to persons of predominantly European descent [3].

While enormous progress has been made in the management of CLL in the past decade to improve clinical outcomes, it is still an incurable cancer whose etiology is poorly understood. Family history of non-Hodgkin lymphoma or lymphoproliferative disorder, hepatitis C virus seropositivity, respiratory tract infections, history of atopic health disorders, reduced sun exposure, living on a farm or exposure to pesticides, and occupation as hairdresser are factors linked to CLL development [4, 5]. The elevated risk observed for occupational exposures of farming and hairdressers suggest that carcinogenic agents are likely involved in the pathogenesis of this blood disease. Thus, carcinogens contained both in drinking water and in recreational pools around the world such as some disinfection by-products (DBPs) and nitrate, are potential candidates.

DBPs constitute a complex mixture of chemicals formed as unintended by-products of the disinfectants applied to both drinking and recreational water to inactivate microbial pathogens [6]. Chlorine is a cost-effective disinfectant widely used worldwide and trihalomethanes are the most prevalent DBPs formed after chlorination. However, although disinfection is essential to avoid waterborne disease, it raises a public health issue: the potential carcinogenicity of DBPs [7]. Moreover, DBP concentrations in swimming pool water often exceed those found in drinking water. The primary routes of DBP uptake during swimming are dermal absorption and inhalation, that result in higher and more persistent levels of DBPs in the bloodstream compared to oral exposure, owing to their direct access to the bloodstream [8].

Nitrate occurrence in the water cycle is steadily increasing on a global scale, primarily due to the widespread use of nitrogen fertilizers and the intensification of agricultural practices [9]. Human exposure to nitrate mainly occurs through ingestion of food and drinking water [10]. When ingested, nitrate is reduced to nitrate, which can interact with endogenous molecules to form N-nitroso compounds such as nitrosamines and nitrosamides. Ingested nitrate leads to endogenous nitrosation to form nitrogenous compounds such as nitrosamines and nitrosamides, which are classified as probably carcinogenic to humans [11, 12]. Within this context and to provide more insight about the health impact of prevalent water contaminants, we explored the link between long-term exposures to THMs and nitrate in drinking water, as well as, the recreational use of swimming pools and CLL development.

## METHODS

### Study design and population

CLL cases were recruited within the multicentric multicase-control (MCC) study in Spain (<http://www.mccspain.org>) in collaboration with the International Cancer Genome Consortium on Chronic Lymphocytic Leukemia Project (ICGC-CLL, [www.cllgenome.es](http://www.cllgenome.es) and [www.icgc.org](http://www.icgc.org)). Briefly, cases and controls were recruited between January 2010 and January 2013. CLL cases were identified in 11 hospitals from different Spanish regions, together with a set of frequency-matched controls by sex, age, and region. Controls were selected from the general population, identified from the lists of randomly selected family practitioners in primary health centers located within the catchment area of the hospitals recruiting CLL [13]. Response rates were 87% for cases and 53% for controls. Inclusion criteria required participants to be 20–85 years old, be able to understand

and answer the questionnaire, and have lived for at least 6 months in the study area.

### Data collection

At recruitment, participants answered a structured computer-assisted questionnaire administered by trained personnel in a face-to-face interview to gather information on anthropometrics (self-reported), socio-demographics, lifestyle factors, and personal and family medical history. Cases and controls provided full address, year start and stop living in all the residences where they lived for at least 12 months since age 18 until the time of the interview, and the type of water consumed in each residence (municipal, bottled, well, other). Amount (glasses/day) of water ingested on average lifetime at home, work and other places was ascertained. Information related to swimming pool attendance included ever attendance (defined as more than 10 times in the lifetime), year (or age) when swimming pool attendance started and ended, average attendance frequency in summer and the rest of the year separately, average time in the water, and type of swimming pool (indoor, outdoor, both). Dietary habits the year before the interview were collected through a self-administered semiquantitative food frequency questionnaire, including a total of 140 food items, previously validated in Spain [14]. Questionnaires used are available online (<http://mccspain.org>). A final section evaluating the reliability of the interview was completed by the interviewer.

### Outcome definition

CLL cases were diagnosed according to the criteria of the International Workshop on Chronic Lymphocytic Leukemia [15]. CLL and small lymphocytic lymphoma are considered the same underlying disease. All diagnoses were morphologically and immunologically confirmed using flow cytometry immunophenotype and complete blood count. In the vast majority of CLL cases, we could differentiate between Rai stage 0 or I–IV tumors. Given the generally indolent course of CLL, with most of the patients having a slow progression disease and with no treatment need, CLL cases were invited to participate independently of their date of diagnosis. Patients who received a CLL diagnosis within three years prior to the start of the study (January 2010) were eligible for inclusion. Incident cases were defined as patients diagnosed at during the study recruitment period (January 2010 – January 2013), while prevalent cases were defined as participants diagnosed  $\leq 3$  years prior to the start of the study (January 2010).

### THM and nitrate levels in municipal drinking water

We designed a structured questionnaire aimed at water utilities, local and/or health authorities to collect drinking water source (surface/ground water proportion) and treatment in the study areas back to 1940, as well as available THMs and nitrate measurements. We targeted data collection among study municipalities that contributed up to 80% of person-years. In addition, centralized routine monitoring data on THMs (chloroform, bromodichloromethane, dibromochloromethane and bromoform) and nitrate was provided by the SINAC (Spanish acronym of the National Information System on Water for Consumption) for the years 2004–2010. This data corresponded to samples collected at drinking water treatment plants and the distribution network, and included information at the water zone level fed by water supply operators from public or private companies or municipalities, and public or private laboratories. The water zone, that mostly corresponds to municipality, was defined as a geographical area supplied by water with a homogeneous source and treatment, and whose quality in the water distributed in the networks can be considered homogeneous. We linked each postal code from the residence to the corresponding water zone. The distribution of the sampling points and the sampling frequency varied greatly depending on the population served, extension of the water zone and the year, and could be more than once a day (e.g., Madrid), up to once every three months, or once a year in less populated areas. Measurements below the analytical limit of quantification (QL) were substituted with half the QL (QL/2) [16]. If the QL was missing, we imputed half of the most frequently reported QL.

### THM and nitrate levels in non-municipal drinking water

We measured nitrate in the 9 most-consumed bottled water brands in Spain using UV spectrophotometry, with 0.5/0.1 mg/L detection/quantification limit. Nitrate concentrations were in the range of 2.3–15.6 mg/l [17]. THMs were previously measured in 15 popular bottled water brands in Spain, through purge-and-trap and gas chromatography. Mean

concentrations for chloroform and brominated THMs were  $\leq 0.1 \mu\text{g/L}$  [18]; and limits of detection were 0.015 (chloroform), 0.004 (bromodichloromethane), 0.005 (dibromochloromethane) and  $0.011 \mu\text{g/L}$  (bromoform). We used THM data from 56 measurements in different Spanish areas that were supplied by chlorinated ground water (e.g. wells). Average concentrations were 0.3, 0.3, 0.8, and  $1.8 \mu\text{g/L}$  for chloroform, bromodichloromethane, dibromochloromethane, and bromoform, respectively. Nitrate data in private wells was not available.

### Estimation of long-term levels in municipal drinking water

We calculated the annual average level of nitrate and THMs at the water zone level. Years without measurements were assigned the average of all available measurements in the water zone if the water source and treatment did not change over the years. In the case of changes in the water source and/or treatment, procedures to back extrapolate were applied.

For THMs, since their concentrations in surface water are generally higher than in ground sources [19], we used surface water percentage as a weight to back extrapolate individual THM concentrations when water source changed through linear interpolation, assuming that concentrations increased proportionately to the percentage of surface water. Likewise, water zones with changes in treatment over the years and THM measurements were used to estimate the change percentage of THM concentrations after introducing such treatments. These percentages were applied as a weight to back-extrapolate THM concentrations in areas with changes in these specific treatments when measurements were unavailable. Before chlorination started, THMs concentrations were assumed to be zero. Total THMs (TTHM) levels were calculated by adding up chloroform, bromodichloromethane, dibromochloromethane, and bromoform concentrations.

For years without nitrate measurements in water zones where water source changed over the years, the ground water percentage was used as a weight to back-extrapolate concentrations using linear interpolation, assuming that nitrate levels increased with ground water proportion [9]. In municipalities without any nitrate measurement (covering  $\sim 0.5\%$  of the total person-years), we imputed the levels of neighboring municipalities supplied with similar ground water proportion  $\pm 10\%$ .

### Individual indices of THM and nitrate exposure in the study population

*Average THMs and nitrate levels in residential tap water.* We used municipality and year to link levels in drinking water with residential histories of study participants from age 18 years to 2 years before the interview. We estimated the average concentration of nitrate (mg/L) and THMs ( $\mu\text{g/L}$ ) for this period, henceforth referred to as "lifetime" or "long-term exposure". In general, participants were assigned to different water zones based on their residence patterns. On average, individuals lived in three different residences during the exposure window period, with the residence at the time of the interview being the longest, spanning approximately 30 years.

*Average ingested nitrate.* To calculate waterborne ingested nitrate (mg/day), we assigned levels in drinking water by year according to the water type consumed at home, including municipal (tap), bottled, and private well/other water. Nitrate levels in municipal water were assigned for tap water consumption. Nitrate levels in the sampled bottled waters (range 2.3–15.6 mg/l) [17] were averaged using the sales frequency of each brand as a weight, leading to 6.1 mg/L of nitrate, that was assigned to study participants consuming bottled water. Since nitrate levels in well water were not available, waterborne ingested nitrate was considered missing for years when well water consumption was reported ( $\sim 2\%$ ). The annual nitrate estimates were averaged from age 18 to 2 years before the interview and multiplied by the average daily water intake at the residence. Total amount of ingested water was ascertained as the number of water glasses per day consumed on average by the participant at home (L/day, assuming 200 mL/glass). Water intakes = 0 and above the 99th percentile (4 L/day), considered implausible, were treated as missing values in the analyses.

### Covariables

We collected information on self-reported age, education, weight and height 1 year before the interview to compute body mass index (BMI,  $\text{kg/m}^2$ ), family history of hematologic malignancy, smoking and physical activity.

Smokers were defined as those smoking at least one cigarette/day for  $\geq 6$  months. Former smokers were defined as those who quit smoking  $\geq 1$  year before the interview. Physical activity was ascertained through open questions on any type of physical activity practiced in life, years, and frequency (hours/week), to calculate metabolic equivalents (METs) from age 16 to 2 years before the interview.

### Statistical Analysis

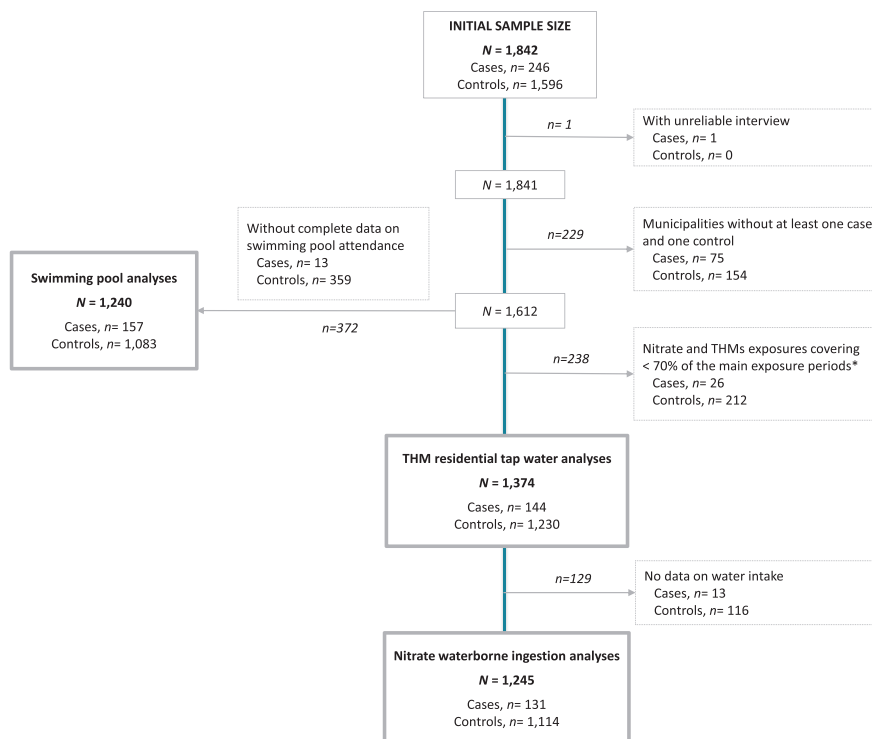
The initial sample consisted of 1,842 participants, comprising 246 CLL cases (92 incident, 154 prevalent) and 1,596 controls. Number of controls was greater than the number of cases because controls were selected to also be matched to other cancer sites. We excluded 1 subject with unreliable interview as qualified by the trained interviewers, and, in order to have a similar geographical distribution of cases and controls, only municipalities with at least one case and one control were included ( $n = 229$  excluded). For the residential tap water analyses, participants with THMs or nitrate estimates covering less than 70% of the years between age 18 to 2 years before the interview were also excluded ( $n = 238$ ), resulting in a sample of 1374 (144 cases, 1230 controls). For the nitrate waterborne ingestion analyses the sample was made up of 1245 (131 cases, 1114 controls) due to missing data on water consumption in 129 participants. Finally, out of 1612 there were 372 participants without complete data on swimming pool attendance and the sample for these analyses it was made up of 1240 (157 cases, 1803 controls) (Fig. 1).

As exposure to nitrate in drinking water is exclusively through the oral route, waterborne nitrate ingestion was evaluated as nitrate exposure. However, since inhalation and dermal absorption are suggested as relevant routes for THMs (volatile and permeable to the skin), residential tap water THM levels were used as main exposure. As a supplementary analysis, waterborne THM ingestion was additionally explored. Average lifetime residential tap water of total THMs, brominated THMs, chloroform ( $\mu\text{g/L}$ ), and waterborne nitrate ingestion (mg/day), were treated as continuous and as tertiles defined according to the distribution among controls.

Our study also encompassed an evaluation of swimming pool attendance. Firstly, we conducted a comparative analysis between individuals classified as 'users' (defined as those who have attended the pool  $\geq 10$  times in their lifetime) and 'non-users' (defined as those who have attended the pool  $< 10$  times in their lifetime). Subsequently, we performed a secondary analysis by categorizing participants based on the average annual frequency of pool attendance. Given that over half of the participants fell into the non-user category, the reference category was designated as the group of 'non-users', while users were further divided based on their attendance frequency, both below and above the median of the yearly average times of attendance. To achieve a finer-grained measurement, we further evaluated each successive 25th percentile increment of yearly pool attendance frequency, distinguishing between summer and rest of the year.

We estimated odds ratios (OR) and 95% confidence intervals (CI) of CLL using mixed models with recruitment area (Barcelona, Asturias, Cantabria) as random effect. To test for linear trends (P-trend) across increasing categories of exposure, the median concentration within each category was treated as a continuous variable in the model. Smoothed spline with three degrees of freedom from general additive models (GAM) were used to visually display the exposure–response relationships on continuous variables.

The selection of adjustment variables was made based on prior knowledge about CLL risk or preventive factors and variables potentially correlated with exposures of interest. Thus, all models were adjusted for recruitment area (introduced as random effect in the mixed model), age ( $< 55$ , 55–64, 65–69, 70–74,  $\geq 75$  y) sex, and education (primary school or lower, secondary school, university). Further adjustment included family history of hematologic malignancy (yes/no), BMI ( $\text{kg/m}^2$ ), smoking (never, former, current), physical activity (inactive, low, moderate, very active), ever worked in farming or agriculture (yes/no) and alcohol consumption (never, former, current moderate consumption [ $\leq 20$  g/day men;  $\leq 10$  g/day women]), current high consumption [ $> 20$  g/day men;  $> 10$  g/day women]). An additional model was reported with mutual adjustments between nitrate waterborne ingestion, residential tap water chloroform and brominated THMs levels. Concerning the analysis of swimming pool attendance, the second model was not adjusted for physical activity, while the subsequent third model did include adjustments for physical activity. We used stochastic regression (which adds a random error term that appropriately reproduces the correlation between X and Y) to impute



**Fig. 1** Flow chart showing exclusions of study participants from the Multicase-Control Study in Spain (MCC-Spain). Cases were defined as patients with CLL recruited within three years from diagnosis to interview (diagnosed  $\leq 3$  y prior to recruitment). \* From 18 years of age to 2 years before the interview.

missing values in BMI (4%), physical activity (<1%) and alcohol consumption (18%). Treating the missing of the alcohol variable as a separate category had no impact on any estimate.

We explored the stratification by sex and by rai stage (0 and I–IV) and some sensitivity analyzes were further conducted: (i) excluding participants with family history of hematologic malignancy, to completely avoid any genetic influence and (ii) only including incident cases (i.e., CLL diagnosed after recruitment), to minimize the possibility of reverse causation. For this supplementary analysis, exposures were exclusively investigated in their continuous form, as this approach was necessitated by a considerable reduction in the sample size. All p values presented are two-tailed;  $<0.05$  was considered statistically significant. Analyses were performed using STATA version 16.0 (Stata Corp, College Station, TX).

## RESULTS

The recruitment area contributing with the largest population was Barcelona. The average age was 64 (11) years old for controls and 66 (10) for CLL cases. On average, cases had more often family history of hematologic malignancy, were less often current smokers, less physically active, worked more often in farming or agriculture, and consumed less alcohol (Table 1). The excluded participants exhibited no discernible differences in comparison to the selected sample (Supplementary Table 1).

Mean (SD) values for average lifetime residential tap water concentrations of brominated THMs ( $\mu\text{g/L}$ ) and chloroform ( $\mu\text{g/L}$ ), and waterborne ingested nitrate (mg/day) were 48.1 (35.6), 18.5 (6.7) and 13.7 (9.6), respectively, in controls; and 72.9 (40.7), 17.9 (5.4) and 14.1 (8.8) in CLL cases. While 58% of the cases exhibited a history of pool attendance over their lifespan, 39% of the control group manifested such attendance. Among users, average annual frequencies were slightly higher in controls than in cases (Table 1). Pool attendance during summer was 31% indoors, 67% outdoors and 2% mixed, while during the rest of the year, 98.5% was indoors, and 1.5% was outdoors (data not in tables).

## Trihalomethanes in residential tap water and nitrate waterborne ingestion

The ORs of CLL were 1.24 (95%CI: 1.16, 1.34) and 1.22 (1.14, 1.31) for each  $10 \mu\text{g/L}$  increase of, respectively, total and brominated THMs in residential tap water. Upon performing exposure analysis through tertile categorization, the third tertile showed an OR exceeding 2 for both total and brominated THMs. However, statistical significance was not attained. In contrast, an inverse association was identified between chloroform levels present in residential tap water and CLL, OR of 0.54 (0.34, 0.87) for  $10 \mu\text{g/L}$  chloroform increase (Table 2, Fig. 2). These associations remained consistent and unchanged in both the stratified analyses and the sensitivity analyses (Table 3). However, the ingestion of waterborne THMs (total, brominated and chloroform) was not found to be associated with CLL (Supplementary Table 2).

Waterborne ingested nitrate was consistently not associated with CLL. The OR associated for each 5 mg/day increase was 0.91 (0.80, 1.04), and when comparing tertile 3 vs. 1, the OR was 1.04 (0.57, 1.89; P-trend 0.64) (Table 2 and Fig. 2). Null associations for nitrate exposure remained in stratified and sensitivity analyses (Table 3).

## Swimming pool attendance

A positive association was observed between ever swimming pool attendance and CLL. The OR of swimming pool users (vs. non-users) was 2.38 (1.61, 3.52). Upon performing annual frequency of attending pools analysis through categorization, the second and third categories showed an ORs of 2.36 (1.49, 3.72) and 2.40 (1.51, 3.83), respectively (P-trend 0.001). For each 25-fold increment in the annual frequency of attending pools, the odds of CLL increased by 5%, OR = 1.05 (0.99, 1.12) (Table 4). This association was sustained both in attendance during the summer and winter months and remained similar in stratified and sensitivity analyses (Table 3).

**Table 1.** Characteristics of the study population.

<b>N = 1612</b>	<b>Controls N = 1442</b>	<b>Cases N = 170</b>
Recruitment area, <i>n</i> (%)		
Barcelona	984 (68.2)	137 (80.6)
Asturias	221 (15.3)	26 (15.3)
Cantabria	237 (16.4)	7 (4.12)
Age (years), mean (SD)		
<55, <i>n</i> (%)	288 (20.0)	27 (15.9)
55–64, <i>n</i> (%)	349 (24.2)	36 (21.2)
65–69, <i>n</i> (%)	286 (19.8)	34 (20.0)
70–74, <i>n</i> (%)	236 (16.4)	36 (21.2)
≥75, <i>n</i> (%)	283 (19.6)	37 (21.8)
Females, <i>n</i> (%)	622 (43.1)	56 (33.0)
Educational level, <i>n</i> (%)		
Primary school or lower	764 (53.0)	101 (59.4)
Secondary school	420 (23.1)	40 (23.5)
University	258 (17.9)	29 (17.1)
Family history of hematologic malignancy, <i>n</i> (%)		
Body Mass Index (kg/m <sup>2</sup> ), mean (SD)	26.9 (4.2)	27.1 (3.8)
Smoking, <i>n</i> (%)		
Never	614 (42.6)	67 (39.4)
Former	555 (38.5)	74 (43.5)
Current smoker	273 (18.9)	29 (17.1)
Physical activity, <i>n</i> (%)		
Inactive	395 (27.4)	54 (31.8)
Low	555 (38.5)	71 (41.8)
Moderate	210 (14.6)	20 (11.8)
Very active	282 (19.6)	25 (14.7)
Ever worked in farming or agriculture, <i>n</i> (%)		
Alcohol consumption, <i>n</i> (%)	323 (22.4)	53 (31.2)
Never	197 (13.7)	17 (10.0)
Former	95 (6.6)	12 (7.1)
Current moderate consumption (≤20 g/day men; ≤10 g/day women)	782 (54.2)	113 (66.5)
Current high consumption (>20 g/day men; >10 g/day women)	368 (25.5)	28 (16.5)
Rai stage, <i>n</i> (%)		
0	—	105 (61.8)
I–IV	—	61 (35.9)
Unknown	—	4 (2.35)
<b>Average lifetime residential tap water, mean (SD) N = 1374</b>	<b>Controls N = 1230</b>	<b>Cases N = 144</b>
Nitrate (mg/L)	8.82 (3.7)	9.7 (3.6)
Total trihalomethanes (µg/L)	66.6 (37.8)	90.8 (42.1)
Brominated trihalomethanes (µg/L)	48.1 (35.3)	72.9 (40.7)
Chloroform (µg/L)	18.5 (6.7)	17.9 (5.4)
<b>Average lifetime waterborne ingestion, mean (SD) N = 1245</b>	<b>Controls N = 1114</b>	<b>Cases N = 131</b>
Nitrate (mg/d)	13.7 (9.6)	14.1 (8.8)
Total trihalomethanes (µg/d)	46 (55.2)	54.1 (72.4)

**Table 1.** continued

Brominated trihalomethanes (µg/d)	32 (42.9)	42.5 (62.9)
Chloroform (µg/d)	14 (15.9)	11.6 (12)
<b>Swimming pool attendance N = 1240</b>	<b>Controls N = 1083</b>	<b>Case N = 157</b>
Users (≥10 times in life), <i>n</i> (%)	425 (39.2)	89 (57.7)
Average annual frequency (among users) (times/year), mean (SD)	89.9 (83.5)	77.5 (69.9)
Average frequency during summer (among users) (times/year), mean (SD)	41.0 (32.1)	35.4 (28.8)
Average frequency the rest of the year (among users) (times/year), mean (SD)	110.1 (65.8)	112.1 (50.1)

Brominated THMs includes bromodichloromethane, dibromochloromethane, and bromoform.

Total THMs (TTHMs) includes chloroform, bromodichloromethane, dibromochloromethane, and bromoform.

Cases of chronic lymphocytic leukemia (CLL) and controls from the MCC-Spain (Barcelona, Asturias, Cantabria).

## DISCUSSION

Despite the extensive use of chlorination as a cost-effective water disinfection method that has been applied for decades, there is surprising little evidence about its potential risks on human health. To our knowledge, this is the first study to evaluate the relationship between exposure to water contaminants, specifically THMs and nitrate in drinking water and swimming pool attendance, with CLL risk.

Elevated long-term total and brominated THM levels in residential tap water exhibited a positive association with an increased likelihood of developing CLL. Conversely, chloroform demonstrated an inverse association with CLL and no significant association was observed between nitrate and CLL. The absence of an association between THMs waterborne ingestion and CLL, in contrast to the observed association with residential THM levels, which account for different exposure routes, may suggest that inhalation and/or dermal exposures may play a more significant role in elevating the risk of CLL than THM ingestion. Swimming pool attendance was also associated with increased risk of CLL.

Chloroform, bromodichloromethane, dichloroacetic acid, and trichloroacetic acid have been classified as possible human carcinogens (Group 2B) by the International Agency for Research on Cancer [7, 20]. However, evidence of DBP carcinogenicity in animals differs by chemical, animal species, organ, and administration route. Since some DBPs (e.g., THMs) are volatile and skin permeable, exposure occurs through dermal absorption and inhalation when showering, bathing, hand dishwashing, cooking or attending swimming pools [21], in addition to ingestion. Long-term exposure to THMs, as a marker of DBP exposure, has been consistently associated with increased bladder cancer risk in epidemiological studies [22]. DPB exposure has also been linked to colorectal cancer, miscarriage, and birth defects, although evidence is less consistent [23, 24]. Exposure to swimming pool waters has also been related to eye and skin irritation, asthma and respiratory disorders [25, 26].

Drinking water chlorination form chloroform but also brominated THMs in the presence of bromide. In this study, mean levels of brominated THMs in drinking water were twice the levels of chloroform. Our data suggests that exposure to brominated THMs, but not chloroform, increases the risk of CLL, which could be expected given their different mechanisms of action and toxicological potential [26, 27]. Brominated THMs are more cytotoxic than chloroform in addition to being also genotoxic

**Table 2.** Association of trihalomethanes (THMs) in residential tap water and nitrate waterborne ingestion with chronic lymphocytic leukemia (CLL).

Exposure	Controls	Cases	OR (95% CI) Multivariable adjusted <sup>a</sup>	OR (95% CI) Multivariable adjusted <sup>b</sup>	OR (95% CI) Multivariable adjusted <sup>c</sup>
Total THMs in residential tap water ( $\mu\text{g/L}$ ), $N = 1374$					
Tertile 1 (<52.2)	411	31	Ref. [1]	Ref. [1]	Ref. [1]
Tertile 2 (52.2–83.2)	409	23	0.48 (0.20, 1.19)	0.49 (0.19, 1.25)	0.57 (0.22, 1.48)
Tertile 3 (>83.2)	410	90	2.11 (0.91, 4.89)	2.13 (0.89, 5.14)	2.22 (0.90, 5.48)
P-trend	1230	144	<0.001	<0.001	<0.001
Per 10 $\mu\text{g/L}$	1230	144	1.24 (1.16, 1.32)	1.23 (1.15, 1.32)	1.24 (1.16, 1.34)
Brominated THMs in residential tap water ( $\mu\text{g/L}$ ), $N = 1374$					
Tertile 1 (<30.0)	411	31	Ref. [1]	Ref. [1]	Ref. [1]
Tertile 2 (30.0–60.8)	409	23	0.50 (0.21, 1.22)	0.52 (0.21, 1.29)	0.85 (0.31, 2.36)
Tertile 3 (>60.8)	410	90	2.16 (0.94, 4.94)	2.23 (0.95, 5.26)	2.32 (0.91, 5.91)
P-trend	1230	144	<0.001	<0.001	<0.001
Per 10 $\mu\text{g/L}$	1230	144	1.25 (1.17, 1.33)	1.24 (1.17, 1.3)	1.22 (1.14, 1.31)
Chloroform in residential tap water ( $\mu\text{g/L}$ ), $N = 1374$					
Tertile 1 (<17.3)	410	54	Ref. [1]	Ref. [1]	Ref. [1]
Tertile 2 (17.3–22.3)	410	62	0.69 (0.45, 1.06)	0.70 (0.45, 1.08)	0.81 (0.50, 1.31)
Tertile 3 (>22.3)	410	28	0.23 (0.13, 0.40)	0.23 (0.13, 0.41)	0.42 (0.22, 0.82)
P-trend	1230	144	<0.001	<0.001	0.02
Per 10 $\mu\text{g/L}$	1230	144	0.44 (0.32, 0.61)	0.44 (0.31, 0.61)	0.54 (0.34, 0.87)
Nitrate waterborne ingestion (mg/day), $N = 1245$					
Tertile 1 (<8.6)	372	34	Ref. [1]	Ref. [1]	Ref. [1]
Tertile 2 (8.6–16.8)	371	57	1.54 (0.91, 2.61)	1.57 (0.92, 2.68)	1.52 (0.87, 2.64)
Tertile 3 (>16.8)	371	40	0.94 (0.53, 1.66)	1.00 (0.56, 1.79)	1.04 (0.57, 1.89)
P-trend	1114	131	0.36	0.51	0.64
Per 5 mg/day	1114	131	0.93 (0.83, 1.04)	0.95 (0.84, 1.06)	0.91 (0.80, 1.04)

Odds ratios (OR) and 95% confidence intervals (CI) were calculated using mixed models with residential area (Barcelona, Asturias, Cantabria) as random effect. Brominated THMs includes bromodichloromethane, dibromochloromethane, and bromoform.

Total THMs includes chloroform, bromodichloromethane, dibromochloromethane, and bromoform.

ORs Odds ratios, 95% CIs 95% confidence intervals.

<sup>a</sup>Model adjusted for age, sex, and educational level.

<sup>b</sup>Model further adjusted for family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture and alcohol consumption.

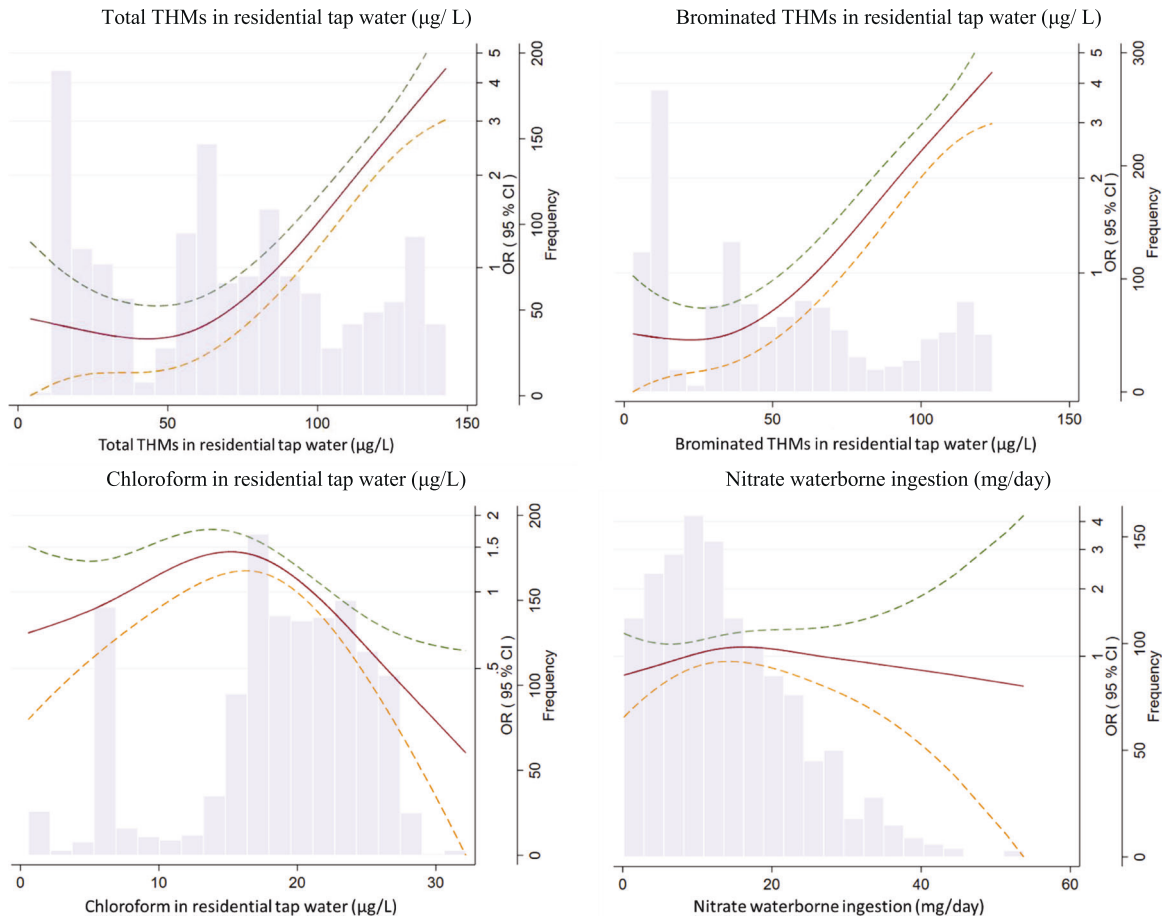
<sup>c</sup>Model further mutually adjusted for the other corresponding components, i.e., total THMs (nitrate model), nitrate (THMs model), chloroform and nitrate (brominated THMs model), brominated THMs and nitrate (chloroform model).

and mutagenic when tested under a wide range of experimental conditions (including human cells) [28, 29]. On the other hand, some animal data have observed that chloroform may have an inhibitory effect on certain tumors [30, 31]. In our sample, chloroform levels showed limited variability in the range where was most of the population. Associations should be cautiously interpreted and remain to be confirmed.

Among the 700 DBP identified, THMs and haloacetic acids are the most prevalent chlorination by-products, and have consequently been used in epidemiological studies as surrogates of total DBPs. However, there are other DBPs more toxic [32] such as haloacetonitriles, haloacetamides, or halonitromethanes [28, 33, 34]. Thus, a possible synergism or additive effect among the different DBPs in drinking water can be suspected and future studies should evaluate more completely the potential effects of combined exposures.

Our results on swimming pool attendance suggest that DBP exposure over the years in swimmers –which occurs through inhalation and dermal absorption, leading to a higher dose and longer duration in the bloodstream relative to oral exposure [8, 35, 36]—might be linked to CLL risk. Swimming pool water frequently exceeds DBP content relative drinking water due to

the constant input of organic matter from swimmers and application of disinfectant [37]. The body fluids of the swimmers (such as urine, sweat, hair, skin cells) and personal care products (such as sunscreen and body lotions) together with other contaminants from leaves, dust or rainwater in the case of outdoor pools, react with the disinfectants used to prevent microbial proliferation –mostly chlorine, and to a lesser extent bromine or ozone– and forms unwanted DBPs. Hundreds of DBPs have also been identified in pools, including brominated THMs [38]. The DBP formation in swimming pools depends on several factors such as number of people in the pool, water temperature, and the amount of total organic carbon [39]. DBPs are reported to be significantly higher in the outdoor and recreational pools than in the indoors [40]. Outdoor and recreational pools are generally smaller in size and depth, with a greater influx of users, many of whom are children, which translates into a higher organic load and, therefore, a higher concentration of DBPs [41]. We, however, did not observe a greater risk of CLL associated with attendance to pools during summer (31% reported to be indoors, 67% outdoors and 2% mixed) than during the rest of the year (98.5% was indoors, and 1.5% was outdoors).



**Fig. 2** Exposure–response relationship of trihalomethanes in residential tap water, nitrate waterborne ingestion, and lifetime total swimming pool hours with chronic lymphocytic leukemia (CLL); expressed as odds ratios (ORs) and 95% confidence intervals (CIs). Smoothed spline with three degrees of freedom from general additive models adjusted for region, age, sex, and educational level, family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture, alcohol consumption. The dashed lines represent the 95% confidence intervals (CIs).

In a previous experimental study, in which we collected blood, urine, and exhaled air samples from adult volunteers before and after they swam for 40 min, we observed genotoxicity of swimming pool water and in swimmers. This genotoxicity was derived mainly from brominated THM but not chloroform [42]. The importance of brominated DBPs in the mutagenicity of recreational waters has been highlighted by other studies [43, 44]. Likewise, the presence of nitrogen-containing organic matter from bathers leads to the formation of nitrogenated by-products such as chloramines (monochloramine, dichloramine, trichloramine), haloacetonitriles and carcinogenic nitrosamines, also important drivers of toxicity [45, 46].

The main strengths of our study include the relatively large sample size that allowed us to analyze exposures using an updated and rigorous classification of CLL, and the long-term exposure approach, from 18 years of age to 2 years before the study interview. We were able to collect detailed individual information and to statistically adjust for a number of potential confounding factors. Residual confounding cannot be ruled out. Finally, the multicentric nature of the study allowed a wide geographic variability of THM and nitrate levels in residential tap water.

The limited historical measurements (particularly before 1980) and assumptions to model historical concentrations could reduce the accuracy of exposure estimates. However, to minimize exposure measurement error, we only included subjects with known exposures for at least 70% of the exposure window.

Inability to account for exposures outside home and use of domestic filtration systems may have introduced nondifferential misclassification in the waterborne ingested estimates. However, the reported amount of water consumed at work (mean  $\pm$  SD:  $0.2 \pm 0.3$  L/day) and other places ( $0.01 \pm 0.05$  L/day) was smaller than that consumed at home ( $1.2 \pm 0.7$  L/day), and minor bias is expected. As for the use of domestic filters, a reduction of THMs levels has been reported [47]. There were no statistics on the use of domestic filters in Spain for the study exposure window, but expert knowledge suggests that the use of household water filters was probably uncommon. Overall, the expected effect on the associations from exposure measurement error is attenuation towards the null [48], as has been shown for other residence-based exposures [49].

In order to minimize potential recall bias in assessing exposure to swimming pools, we focused solely on individual average frequency of pool attendance, distinguishing between summer and the rest of the year. Attempting to obtain more precise information, such as the average hours spent in the pool during in life, would likely be prone to reporting errors. However, despite our efforts, we acknowledge that some degree of recall bias and exposure misclassification may be present. Notwithstanding, it is plausible that any such biases would be nondifferential, thereby tending to bias the association towards a null.

Personal information was collected retrospectively after diagnosis, and differential recall between cases and controls may not be totally ruled out. However, the questionnaire was administered

**Table 3.** Stratified analyses by chronic lymphocytic leukemia (CLL) rai stage and sex, and sensitivity analyses.

Stratification by:	CLL rai stage		Sex	
	Stage 0	Stage I-IV	Men	Women
Brominated THMs in residential tap water, OR <sup>a</sup> (95% CI) Per 10 µg/L	1.21 (1.11, 1.31), (N = 1035)	1.23 (1.10, 1.37), (N = 1190)	1.26 (1.15, 1.39), (N = 736)	1.17 (1.04, 1.32), (N = 536)
Chloroform in residential tap water, OR <sup>b</sup> (95% CI) Per 10 µg/L	0.57 (0.32, 1.03), (N = 1035)	0.38 (0.19, 0.77), (N = 1190)	0.46 (0.25, 0.85), (N = 736)	0.65 (0.31, 1.35), (N = 536)
Nitrate waterborne ingestion, OR <sup>c</sup> (95% CI) Per 5 mg/day	0.95 (0.82, 1.11), (N = 1027)	0.87 (0.71, 1.07), (N = 1163)	0.93 (0.79, 1.10), (N = 724)	0.94 (0.76, 1.16), (N = 521)
Average frequency of pool attendance, OR <sup>d</sup> (95% CI) Per 25 times/year increase <sup>e</sup>	1.06 (0.99, 1.14), (N = 1180)	1.00 (0.90, 1.11), (N = 1139)	1.07 (0.96, 1.18), (N = 521)	1.05 (0.97, 1.13), (N = 719)
<b>Excluding:</b>	<b>Family history of hematologic malignancy</b>	<b>Prevalent CLL cases<sup>f</sup></b>		
Brominated THMs in residential tap water, OR <sup>a</sup> (95% CI) Per 10 µg/L	1.22 (1.14, 1.32), (N = 1118)	1.21 (1.08, 1.37), (N = 1006)		
Chloroform in residential tap water, OR <sup>b</sup> (95% CI) Per 10 µg/L	0.41 (0.25, 0.66), (N = 1118)	0.73 (0.31, 1.68), (N = 1006)		
Nitrate waterborne ingestion, OR <sup>c</sup> (95% CI) Per 5 mg/day	0.91 (0.80, 1.05), (N = 1093)	0.76 (0.61, 0.95), (N = 998)		
Average frequency of pool attendance, OR <sup>d</sup> (95% CI) Per 25 times/year increase <sup>e</sup>	1.06 (0.99, 1.13), (N = 1098)	1.05 (0.96, 1.16), (N = 914)		

Brominated THMs includes bromodichloromethane, dibromochloromethane, and bromoform.

ORs Odds ratios, 95% CIs 95% confidence intervals.

<sup>a</sup>Model adjusted for region, age, sex, and educational level, family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture, alcohol consumption, chloroform residential tap water and nitrate waterborne ingestion.

<sup>b</sup>Model adjusted for region, age, sex, and educational level, family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture, alcohol consumption, brominated THM residential tap water and nitrate waterborne ingestion.

<sup>c</sup>Model adjusted for region, age, sex, and educational level, family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture, alcohol consumption, total THM residential tap water.

<sup>d</sup>Model adjusted for region, age, sex, and educational level, family history of hematologic malignancy, body mass index, smoking, physical activity, ever worked in farming or agriculture and alcohol consumption.

<sup>e</sup>Corresponds to the 25th percentile among users.

<sup>f</sup>Only including newly diagnosed patients during the study period.



**Table 4.** Association of average frequency of pool attendance with chronic lymphocytic leukemia (CLL).

Exposure	Controls	Cases	OR (95% CI) Multivariable adjusted <sup>a</sup>	OR (95% CI) Multivariable adjusted <sup>b</sup>	OR (95% CI) Multivariable adjusted <sup>c</sup>
Lifetime swimming pool attendance					
Non-users (<10 times in life)	658	65	Ref. [1]	Ref. [1]	Ref. [1]
Users (≥10 times in life)	425	89	2.08 (1.43, 3.03)	2.23 (1.52–3.27)	2.38 (1.61–3.52)
Average annual frequency					
Tertile 1 (non-users)	658	68	Ref. [1]	Ref. [1]	Ref. [1]
Tertile 2 (1 to <78 times/year)	228	47	2.07 (1.33, 3.23)	2.23 (1.42, 3.51)	2.36 (1.49, 3.72)
Tertile 3 (>78 times/year)	197	42	2.09 (1.34, 3.26)	2.23 (1.42, 3.50)	2.40 (1.51, 3.83)
P-trend	1083	157	0.004	0.002	0.001
Per 25 times/year increase <sup>d</sup>	1083	157	1.04 (0.98, 1.10)	1.04 (0.98, 1.11)	1.05 (0.99, 1.12)
Per 15 times/year increase during summer <sup>d</sup>	1015	149	1.08 (0.98, 1.18)	1.08 (0.98, 1.19)	1.09 (1.00, 1.20)
Per 40 times/year increase during the rest of the year <sup>d</sup>	872	104	1.14 (1.00, 1.30)	1.16 (1.02, 1.32)	1.19 (1.04, 1.36)

Odds ratios (OR) and 95% confidence intervals (CI) were calculated using mixed models with residential area (Barcelona, Asturias, Cantabria) as random effect. Odds ratios (ORs) and 95% confidence intervals (CIs)  $N = 1240$ .

<sup>a</sup>Model adjusted for age, sex, and educational level.

<sup>b</sup>Model further adjusted for family history of hematologic malignancy, body mass index, smoking, ever worked in farming or agriculture and alcohol consumption.

<sup>c</sup>Model further adjusted for physical activity.

<sup>d</sup>Corresponds to the 25th percentile among users.

by trained personnel in a face-to-face interview, and there is not an obvious link between the water questions and CLL that could motivate different responses between cases and controls. Moreover, the interviewers rated the quality of the interview and unreliable or inconsistent interviews were excluded from our analyses.

Selection bias arising from control sampling might be of concern since response rates were moderate, especially among controls, partly explained by the population-based source as opposed to hospital-based cases. However, there were no differences in the main characteristics between cases and controls and the probability of participation can be assumed to be independent from the exposure, thus, nondifferential bias is expected, if any. It was not possible to perform analyses by recruitment area due to the limited within area variability, but we conducted mixed models with area as random effect, which adjusts for environmental factors with potential different geographic distribution. Nevertheless, since correlation of these factors with our main exposures is unlikely, the expected impact on the results is minimal. Finally, the inclusion of prevalent cases might be another cause for attention since patients who survived might have a different etiology than those who died rapidly after diagnosis or might have modified their lifestyle, presumably towards healthier-like habits. However, sensitivity analysis suggested that the use of prevalent cases might not have introduced selective survival bias or reverse causality.

Our findings suggest that long-term exposure to brominated THMs in drinking water, as well as swimming pool attendance may have a role in the CLL etiology. Our unprecedented findings are highly relevant since no previous study has explored these associations. However, this is a first piece of evidence that needs to be confirmed in other studies and the biologic basis for the associations observed remains to be elucidated.

#### DATA AVAILABILITY

Additional data that may be of interest to readers are available from the corresponding author on reasonable request.

#### REFERENCES

1. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127:2375–90.
2. NIH. The Surveillance Epidemiology and End Results (SEER) Program of the National Cancer Institute Cancer fact sheets: chronic lymphocytic leukemia (CLL) <https://seer.cancer.gov/statfacts/html/clyl.html> (Accessed 13th September 2023).
3. Yang SM, Li JY, Gale RP, Huang XJ. The mystery of chronic lymphocytic leukemia (CLL): why is it absent in Asians and what does this tell us about etiology, pathogenesis and biology? *Blood Rev*. 2015;29:205–13.
4. Benavente Y, Casabonne D, Costas L, Robles C, Alonso E, de la Banda E, et al. Established and suggested exposures on CLL/SLL etiology: results from the CLL-MCC-Spain study. *Cancer Epidemiol*. 2018;52:106–11.
5. Slager SL, Benavente Y, Blair A, Vermeulen R, Cerhan JR, Costantini AS, et al. Medical history, lifestyle, family history, and occupational risk factors for chronic lymphocytic leukemia/small lymphocytic lymphoma: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014;2014:41–51.
6. Richardson S, Thruston A, Caughran T, Chen P, Collette T, Schenck K, et al. Identification of new drinking water disinfection by-products from ozone, chlorine dioxide, chloramine, and chlorine. *Water Air Soil Pollut*. 2000;123:95–102.
7. International Agency for Research on Cancer. IARC, monographs on the evaluation of carcinogenic risk to humans. In: Trichloroethylene, tetrachloroethylene, and some other chlorinated agents: Lyon, France: IARC; 2014.
8. Zare Afifi M, Blatchley ER 3rd. Seasonal dynamics of water and air chemistry in an indoor chlorinated swimming pool. *Water Res*. 2015;68:771–83.
9. Ward MH, Jones RR, Brender JD, de Kok TM, Weyer PJ, Nolan BT, et al. Drinking water nitrate and human health: an updated review. *Int J Environ Res Public Health*. 2018;15:1557.
10. Espejo-Herrera N, Gracia-Lavedan E, Pollan M, Aragonés N, Boldo E, Perez-Gomez B, et al. Ingested nitrate and breast cancer in the Spanish Multicase-Control Study on Cancer (MCC-Spain). *Environ Health Perspect*. 2016;124:1042–9.
11. Grosse Y, Baan R, Straif K, Secretan B, El Ghissassi F, Cogliano V, et al. Carcinogenicity of nitrate, nitrite, and cyanobacterial peptide toxins. *Lancet Oncol*. 2006;7:628–9.
12. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. IARC monographs on the evaluation of carcinogenic risks to humans. Ingested nitrate and nitrite, and cyanobacterial peptide toxins. *IARC Monogr Eval Carcinog Risks Hum*. 2010;94:1–412.
13. Castano-Vinyals G, Aragonés N, Perez-Gomez B, Martin V, Llorca J, Moreno V, et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *Gac Sanit*. 2015;29:308–15.
14. Martin-Moreno JM, Boyle P, Gorgojo L, Maisonneuve P, Fernandez-Rodriguez JC, Salvini S, et al. Development and validation of a food frequency questionnaire in Spain. *Int J Epidemiol*. 1993;22:512–9.

15. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;111:5446–56.
16. Croghan C, Egeghy PP. Methods of dealing with values below the limit of detection using SAS. St. Petersburg, FL: Southern SAS User Group; 2003.
17. Espejo-Herrera N, Kogevinas M, Castano-Vinyals G, Aragonés N, Boldo E, Ardanaz E, et al. Nitrate and trace elements in municipal and bottled water in Spain. *Gac Sanit*. 2013;27:156–60.
18. Font-Ribera L, Kogevinas M, Nieuwenhuijsen MJ, Grimalt JO, Villanueva CM. Patterns of water use and exposure to trihalomethanes among children in Spain. *Environ Res*. 2010;110:571–9.
19. Villanueva CM, Kogevinas M, Grimalt JO. Haloacetic acids and trihalomethanes in finished drinking waters from heterogeneous sources. *Water Res*. 2003;37:953–8.
20. International Agency for Research on Cancer, et al. IARC, Monographs on the Evaluation of Carcinogenic Risk to Humans. Some chemicals present in industrial and consumer products. Food and Drinking-Water Lyon (France). 2016;101:349–74.
21. Villanueva CM, Cordier S, Font-Ribera L, Salas LA, Levallois P. Overview of disinfection by-products and associated health effects. *Curr Environ Health Rep*. 2015;2:107–15.
22. Costet N, Villanueva CM, Jaakkola JJ, Kogevinas M, Cantor KP, King WD, et al. Water disinfection by-products and bladder cancer: is there a European specificity? A pooled and meta-analysis of European case-control studies. *Occup Environ Med*. 2011;68:379–85.
23. Save-Soderbergh M, Toljander J, Donat-Vargas C, Akesson A. Drinking water disinfection by-products and congenital malformations: a nationwide register-based prospective study. *Environ Health Perspect*. 2021;129:97012.
24. Nieuwenhuijsen MJ, Martinez D, Grellier J, Bennett J, Best N, Izzatt N, et al. Chlorination disinfection by-products in drinking water and congenital anomalies: review and meta-analyses. *Environ Health Perspect*. 2009;117:1486–93.
25. Villanueva CM, Font-Ribera L. Health impact of disinfection by-products in swimming pools. *Ann Ist Super Sanita*. 2012;48:387–96.
26. Font-Ribera L, Marco E, Grimalt JO, Pastor S, Marcos R, Abramsson-Zetterberg L, et al. Exposure to disinfection by-products in swimming pools and biomarkers of genotoxicity and respiratory damage - the PISCINA2 Study. *Environ Int*. 2019;131:104988.
27. Cortes C, Marcos R. Genotoxicity of disinfection byproducts and disinfected waters: a review of recent literature. *Mutat Res Genet Toxicol Environ Mutagen*. 2018;831:1–12.
28. Richardson SD, Plewa MJ, Wagner ED, Schoeny R, Demarini DM. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. *Mutat Res*. 2007;636:178–242.
29. de Castro Medeiros L, de Alencar FLS, Navoni JA, de Araujo ALC, do Amaral VS. Toxicological aspects of trihalomethanes: a systematic review. *Environ Sci Pollut Res Int*. 2019;26:5316–32.
30. Daniel FB, Reddy TV, Stober JA, Olson GR. Site-specific modulation of carcinogen-induced gastrointestinal tract nuclear anomalies in B6C3F1 mice by chloroform. *Anticancer Res*. 1991;11:665–70.
31. Zhang L, Cai Q, Lin J, Fang Y, Zhan Y, Shen A, et al. Chloroform fraction of *Scutellaria barbata* D. Don promotes apoptosis and suppresses proliferation in human colon cancer cells. *Mol Med Rep*. 2014;9:701–6.
32. Plewa MJ, Muellner MG, Richardson SD, Fasano F, Buettner KM, Woo YT, et al. Occurrence, synthesis, and mammalian cell cytotoxicity and genotoxicity of haloacetamides: an emerging class of nitrogenous drinking water disinfection byproducts. *Environ Sci Technol*. 2008;42:955–61.
33. Ding S, Chu W, Krasner SW, Yu Y, Fang C, Xu B, et al. The stability of chlorinated, brominated, and iodinated haloacetamides in drinking water. *Water Res*. 2018;142:490–500.
34. Allen JM, Plewa MJ, Wagner ED, Wei X, Bokenkamp K, Hur K, et al. Drivers of disinfection byproduct cytotoxicity in U.S. Drinking Water: Should Other DBPs be considered for regulation? *Environ Sci Technol*. 2022;56:392–402.
35. Ashley DL, Blount BC, Singer PC, Depaz E, Wilkes C, Gordon S, et al. Changes in blood trihalomethane concentrations resulting from differences in water quality and water use activities. *Arch Environ Occup Health*. 2005;60:7–15.
36. Leavens TL, Blount BC, DeMarini DM, Madden MC, Valentine JL, Case MW, et al. Disposition of bromodichloromethane in humans following oral and dermal exposure. *Toxicol Sci*. 2007;99:432–45.
37. Carter RAA, Allard S, Croue JP, Joll CA. Occurrence of disinfection by-products in swimming pools and the estimated resulting cytotoxicity. *Sci Total Environ*. 2019;664:851–64.
38. Richardson SD, DeMarini DM, Kogevinas M, Fernandez P, Marco E, Lourencetti C, et al. What's in the pool? A comprehensive identification of disinfection by-products and assessment of mutagenicity of chlorinated and brominated swimming pool water. *Environ Health Perspect*. 2010;118:1523–30.
39. Chu H, Nieuwenhuijsen MJ. Distribution and determinants of trihalomethane concentrations in indoor swimming pools. *Occup Environ Med*. 2002;59:243–7.
40. Font-Ribera L, Esplugues A, Ballester F, Martínez-Arguelles B, Tardon A, Freire C, et al. [Trihalomethanes in swimming pool water in four areas of Spain participating in the INMA project]. *Gac Sanit*. 2010;24:483–6.
41. Carter RAA, Joll CA. Occurrence and formation of disinfection by-products in the swimming pool environment: A critical review. *J Environ Sci*. 2017;58:19–50.
42. Kogevinas M, Villanueva CM, Font-Ribera L, Liviá D, Bustamante M, Espinoza F, et al. Genotoxic effects in swimmers exposed to disinfection by-products in indoor swimming pools. *Environ Health Perspect*. 2010;118:1531–7.
43. Weaver WA, Li J, Wen Y, Johnston J, Blatchley MR, Blatchley ER 3rd. Volatile disinfection by-product analysis from chlorinated indoor swimming pools. *Water Res*. 2009;43:3308–18.
44. Liberatore HK, Daiber EJ, Ravuri SA, Schmid JE, Richardson SD, DeMarini DM. Disinfection byproducts in chlorinated or brominated swimming pools and spas: role of brominated DBPs and association with mutagenicity. *J Environ Sci*. 2022;117:253–63.
45. Jurado-Sanchez B, Ballesteros E, Gallego M. Screening of N-nitrosamines in tap and swimming pool waters using fast gas chromatography. *J Sep Sci*. 2010;33:610–6.
46. Walse SS, Mitch WA. Nitrosamine carcinogens also swim in chlorinated pools. *Environ Sci Technol*. 2008;42:1032–7.
47. Egorov AI, Tereschenko AA, Altshul LM, Vartiainen T, Samsonov D, LaBrecque B, et al. Exposures to drinking water chlorination by-products in a Russian city. *Int J Hyg Environ Health*. 2003;206:539–51.
48. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Vol. 3. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
49. Wei Y, Qiu X, Yazdi MD, Shtein A, Shi L, Yang J, et al. The impact of exposure measurement error on the estimated concentration-response relationship between long-term exposure to PM2.5 and mortality. *Environ Health Perspect*. 2022;130:77006.

## ACKNOWLEDGEMENTS

We would like to acknowledge all the research assistants and interviewers in the study centers, the staff of all participating hospitals, and, most of all the study participants.

## AUTHOR CONTRIBUTIONS

CDV conceptualized the study and designed the methodology. CDV carried out data analysis and drafted the manuscript, which was reviewed by all co-authors. LC conducted the exposure modeling. YB contributed to the statistical methods. EC, GCV, GFT, JLL, IGA, NA, coordinated the fieldwork for data collection from study subjects. MK, MP, and DC administered the project and acquired the financial support.

## FUNDING

The study was partially funded by the “Acción Transversal del Cáncer”, approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PS09/00773, PS09/01286, PS09/01903, PS09/02078, PS09/01662, PI11/01889, PI11/00226, PI12/01270, PI12/00715, PI14/0613, PI15/00914, PI17CIII/00034), by the Fundación Marqués de Valdecilla (API 10/09), by the Conselleria de Sanitat of the Generalitat Valenciana (AP\_061/10), by the European Commission grants FOOD-CT-2006-036224-HIWATE, by the Spanish Association Against Cancer (AECC) Scientific Foundation, by the Catalan Government-Agency for Management of University and Research Grants (AGAUR) grants 2017SGR723, 2021SGR01354 and 2014SGR850, by the Fundación Caja de Ahorros de Asturias and by the University of Oviedo. ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the “Centro de Excelencia Severo Ochoa 2019-2023” Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

## COMPETING INTERESTS

The authors declare no competing interests.

## ETHICAL APPROVAL

The MCC-Spain study followed national and international directives on ethics and data protection (i.e., declaration of Helsinki and Spanish law on confidentiality of data, Ley Organica 15/1999 de 13 Diciembre de Protección de Datos de carácter personal-LOPD; approval number 2008/3121/I). All subjects were informed about the

study and signed an informed consent form for study recruitment, collection and testing of biological samples and verification of clinical information.

#### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41370-023-00600-7>.

**Correspondence** and requests for materials should be addressed to Cristina M. Villanueva.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023