DOI:10.1093/jncimonographs/lgu013

Etiologic Heterogeneity Among Non-Hodgkin Lymphoma Subtypes: The InterLymph Non-Hodgkin Lymphoma Subtypes Project

Lindsay M. Morton*, Susan L. Slager*, James R. Cerhan*, Sophia S. Wang, Claire M. Vajdic, Christine F. Skibola, Paige M. Bracci, Silvia de Sanjosé, Karin E. Smedby, Brian C. H. Chiu, Yawei Zhang, Sam M. Mbulaiteye, Alain Monnereau, Jennifer J. Turner, Jacqueline Clavel, Hans-Olov Adami, Ellen T. Chang, Bengt Glimelius, Henrik Hjalgrim, Mads Melbye, Paolo Crosignani, Simonetta di Lollo, Lucia Miligi, Oriana Nanni, Valerio Ramazzotti, Stefania Rodella, Adele Seniori Costantini, Emanuele Stagnaro, Rosario Tumino, Carla Vindigni, Paolo Vineis, Nikolaus Becker, Yolanda Benavente, Paolo Boffetta, Paul Brennan, Pierluigi Cocco, Lenka Foretova, Marc Maynadié, Alexandra Nieters, Anthony Staines, Joanne S. Colt, Wendy Cozen, Scott Davis, Anneclaire J. de Roos, Patricia Hartge, Nathaniel Rothman, Richard K. Severson, Elizabeth A. Holly, Timothy G. Call, Andrew L. Feldman, Thomas M. Habermann, Mark Liebow, Aaron Blair, Kenneth P. Cantor, Eleanor V. Kane, Tracy Lightfoot, Eve Roman, Alex Smith, Angela Brooks-Wilson, Joseph M. Connors, Randy D. Gascoyne, John J. Spinelli, Bruce K. Armstrong, Anne Kricker, Theodore R. Holford, Qing Lan, Tongzhang Zheng, Laurent Orsi, Luigino Dal Maso, Silvia Franceschi, Carlo La Vecchia, Eva Negri, Diego Serraino, Leslie Bernstein, Alexandra Levine, Jonathan W. Friedberg, Jennifer L. Kelly, Sonja I. Berndt, Brenda M. Birmann, Christina A. Clarke, Christopher R. Flowers, James M. Foran, Marshall E. Kadin, Ora Paltiel, Dennis D. Weisenburger*, Martha S. Linet*, Joshua N. Sampson*

*These authors contributed equally to this work.

Correspondence to: Lindsay M. Morton, PhD, Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 9609 Medical Center Drive, Room 7E-454, Bethesda, MD 20892-9778 (e-mail: mortonli@mail.nih.gov).

Background Non-Hodgkin lymphoma (NHL) comprises biologically and clinically heterogeneous subtypes. Previously, study size has limited the ability to compare and contrast the risk factor profiles among these heterogeneous subtypes.

Methods We pooled individual-level data from 17 471 NHL cases and 23 096 controls in 20 case–control studies from the International Lymphoma Epidemiology Consortium (InterLymph). We estimated the associations, measured as odds ratios, between each of 11 NHL subtypes and self-reported medical history, family history of hematologic malignancy, lifestyle factors, and occupation. We then assessed the heterogeneity of associations by evaluating the variability (Q value) of the estimated odds ratios for a given exposure among subtypes. Finally, we organized the subtypes into a hierarchical tree to identify groups that had similar risk factor profiles. Statistical significance of tree partitions was estimated by permutation-based P values (P_{NODE}).

Results Risks differed statistically significantly among NHL subtypes for medical history factors (autoimmune diseases, hepatitis C virus seropositivity, eczema, and blood transfusion), family history of leukemia and multiple myeloma, alcohol consumption, cigarette smoking, and certain occupations, whereas generally homogeneous risks among subtypes were observed for family history of NHL, recreational sun exposure, hay fever, allergy, and socioeconomic status. Overall, the greatest difference in risk factors occurred between T-cell and B-cell lymphomas ($P_{NODE} < 1.0 \times 10^{-4}$), with increased risks generally restricted to T-cell lymphomas for eczema, T-cell-activating autoimmune diseases, family history of multiple myeloma, and occupation as a painter. We further observed substantial heterogeneity among B-cell lymphomas ($P_{NODE} < 1.0 \times 10^{-4}$). Increased risks for B-cell-activating autoimmune disease and hepatitis C virus seropositivity and decreased risks for alcohol consumption and occupation as a teacher generally were restricted to marginal zone lymphoma, Burkitt/Burkitt-like lymphoma/leukemia, diffuse large B-cell lymphoma, and/or lymphoplasmacytic lymphoma/Waldenström macroglobulinemia.

Conclusions Using a novel approach to investigate etiologic heterogeneity among NHL subtypes, we identified risk factors that were common among subtypes as well as risk factors that appeared to be distinct among individual or a few subtypes, suggesting both subtype-specific and shared underlying mechanisms. Further research is needed to test putative mechanisms, investigate other risk factors (eg, other infections, environmental exposures, and diet), and evaluate potential joint effects with genetic susceptibility.

J Natl Cancer Inst Monogr 2014;48:130–144

Non-Hodgkin lymphoma (NHL) is the most common hematologic malignancy and the fifth most common type of cancer in more developed regions of the world (1). Numerous NHL subtypes with distinct combinations of morphologic, immunophenotypic, genetic, and clinical features are currently recognized (2,3). The incidence of NHL subtypes varies substantially by age, sex, and race/ethnicity (4–7). However, the etiological implications of this biological, clinical, and epidemiological diversity are incompletely understood.

The importance of investigating etiology by NHL subtype is clearly supported by research on immunosuppression, infections, and autoimmune diseases, which are the strongest and most established risk factors for NHL. Studies of solid organ transplant recipients and individuals infected with HIV demonstrate that risks are markedly increased for several—but not all—NHL subtypes (8–13). Some infections and autoimmune diseases are associated with a single specific subtype [eg, human T-cell lymphotropic virus, type I (HTLV-I) with adult T-cell leukemia/lymphoma (14), celiac disease with enteropathy-type peripheral T-cell lymphoma (PTCL) (15–17)], whereas others [eg, Epstein–Barr virus, hepatitis C virus (HCV), Sjögren's syndrome (18–21)] have been associated with multiple subtypes.

In the last two decades, reports from individual epidemiological studies of NHL have suggested differences in risks among NHL subtypes for a wide range of risk factors, but most studies have lacked the statistical power to assess any differences quantitatively and have not systematically evaluated combinations of subtypes. One study assessed multiple risk factors and found support for both etiologic commonality and heterogeneity for NHL subtypes, with risk factor patterns suggesting that immune dysfunction is of greater etiologic importance for diffuse large B-cell lymphoma (DLBCL) and marginal zone lymphoma than for chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL) and follicular lymphoma (22). However, that analysis was limited to approximately 1300 NHL cases and considered only the four most common NHL subtypes. Pooling data from multiple studies through the International Lymphoma Epidemiology Consortium (InterLymph) have provided substantial insight into associations between specific risk factors and NHL subtypes, with evidence that family history of hematologic malignancy, autoimmune diseases, atopic conditions, lifestyle factors (smoking, alcohol, anthropometric measures, and hair dye use), and sun exposure are associated with NHL risk (19,21,23–32). However, no previous study has compared patterns of risk for a range of exposures for both common and rarer NHL subtypes.

We undertook the InterLymph NHL Subtypes Project, a pooled analysis of 20 case–control studies including 17 471 NHL cases and 23 096 controls, to advance understanding of NHL etiology by investigating NHL subtype-specific risks associated with medical history, family history of hematologic malignancy, lifestyle factors, and occupation. The detailed risk factor profiles for each of 11 NHL subtypes appear in this issue (15–17,33–40). In this report, we assess risk factor heterogeneity among the NHL subtypes and identify subtypes that have similar risk factor profiles.

Methods

Study Population and Data Harmonization

Detailed methodology for the InterLymph NHL Subtypes Project is provided elsewhere in this issue (41). Briefly, the 20 studies included in this pooled analysis fulfilled the following criteria: 1) case–control design with incident, histologically confirmed cases of NHL and 2) availability of individual-level data by December 31, 2011. Contributing studies were approved by local ethics review committees, and all participants provided informed consent before interview.

NHL subtypes were defined according to the World Health Organization (WHO) classification (2,3), and guidelines from the InterLymph Pathology Working Group were used to harmonize NHL subtypes classified using other methods (42,43). Consistent with the WHO, lymphoid leukemias were included in this analysis; however, plasma cell neoplasms were excluded because few studies collected data for these cases. Overall, 70% of cases were originally classified using the WHO classification, with the percentage ranging from 54% for Burkitt/Burkitt-like lymphoma/leukemia (BL) to 100% for marginal zone lymphoma, mantle cell lymphoma, and lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM; Table 1).

Each study collected data in a standardized, structured format by in-person or telephone interviews and/or self-administered questionnaires. In some studies, participants also provided a venous blood sample at the time of interview. We centrally harmonized individual-level, de-identified data for medical history, family history of hematologic malignancy, lifestyle factors, and occupation from each study when data on that factor were available from at least four studies. All of these risk factors were included in this analysis regardless of the subtype-specific results presented elsewhere (15–17,33–40).

Statistical Analysis

We first assessed the overall association between each exposure and NHL using odds ratios (ORs) from unconditional fixed effects logistic regression, adjusting for age, race/ethnicity, sex, and study. Because studies selectively focused on specific NHL subtypes and the resulting distribution of cases was not representative of NHL in the general population, our analysis weighted subtypes (using the R function svyglm) to reflect their prevalence among US adults, which is approximately comparable to NHL subtype distributions in Europe and Australia (Supplementary Table 1, available online). For all analyses, categorical and ordinal variables were transformed into a single continuous covariate by ordering the categories and assigning them to equally spaced values between 0 and 1, as listed in Supplementary Table 2 (available online). Therefore, for binary exposures the OR is the increase in the odds of cancer among exposed individuals, while for categorical and ordinal variables, OR is a summary value approximating the increase in odds among individuals in the highest category, compared to those in the lowest category.

We then assessed the association between each exposure and each NHL subtype, estimating ORs from fixed effects logistic regression, adjusting for age, race/ethnicity, sex, and study. The estimated ORs are presented in a colored array (Figure 1) for statistically significantly associated exposures (described below) and in Supplementary Table 2 (available online) for all exposures. We used these estimated ORs to 1) assess whether the exposure was associated with at least one NHL subtype, 2) evaluate risk factor heterogeneity among NHL subtypes, and 3) cluster the subtypes into groups with similar risk factor profiles.

Table 1. Characteristics of 17 471 non-Hodgkin lymphoma cases and 23 096 controls included in the InterLymph NHL Subtypes Project*

Characteristics	Controls	Total NHL cases	Specified NHL subtypes†										
			DLBCL	FL	CLL/SLL	MZL	PTCL	MCL	LPL/WM	MF/SS	BL	HCL	ALL
Total No.	23 096	17 471	4667	3530	2440	1052	584	557	374	324	295	154	152
No. contributing studies	20	20	19	19	13	13	15	13	11	14	18	5	16
Population-based design, %	77.3	80.2	81.4	82.4	67.9	80.5	80.8	78.1	77.8	86.4	83.7	70.8	68.4
By region, %													
North America	49.6	45.9	44.1	52.5	36.1	53.3	40.6	45.4	41.7	61.4	62.4	0.0	36.8
Northern Europe	28.3	31.6	34.8	31.2	45.7	32.6	41.6	47.4	41.7	20.1	19.7	72.7	38.8
Southern Europe	19.0	18.4	16.2	9.2	17.0	8.3	15.1	3.2	9.4	17.3	16.6	20.8	21.1
Australia	3.0	4.0	4.9	7.1	1.2	5.8	2.7	3.9	7.2	1.2	1.4	6.5	3.3
Cases classified by WHO, %	N/A	68.6	71.1	73.2	81.2	100	90.4	100	100	77.8	53.9	80.5	60.5
Male, %	58.4	57.4	55.2	50.6	66.1	46.8	59.4	74.0	60.7	56.8	70.2	78.6	60.5
Non-Hispanic white, %	93.4	91.5	90.4	91.5	95.7	87.2	87.7	93.9	94.1	83.6	84.4	96.1	86.8
Median age, y‡ (range)	59 (16–98)	60 (17–96)	59 (18–96)	58 (18–91)	64 (28–93)	61 (19–91)	56 (18–88)	62 (22–88)	64 (27–89)	56 (22–84)	53 (18–84)	55 (29–79)	41 (18–91)

* ALL = acute lymphoblastic leukemia/lymphoma; BL = Burkitt/Burkitt-like lymphoma/leukemia; CLL/SLL = chronic lymphocytic leukemia/small lymphocytic lymphoma; DLBCL = diffuse large B-cell lymphoma; FL = follicular lymphoma; HCL = hairy cell leukemia; InterLymph = International Lymphoma Epidemiology Consortium; LPL/WM = lymphoplasmacytic lymphoma/Waldenström macroglobulinemia; MCL = mantle cell lymphoma; MF/SS = mycosis fungoides/Sézary syndrome; MZL = marginal zone lymphoma; N/A = not applicable; NHL = non-Hodgkin lymphoma; PTCL = peripheral T-cell lymphoma; WHO = World Health Organization.

† We grouped cases into NHL subtypes according to the WHO classification (2,3) using guidelines from the InterLymph Pathology Working Group (42,43). Total also includes rare subtypes with less than 100 cases (N = 50) and poorly specified subtypes (N = 3292). Most studies had some form of centralized pathology review by at least one expert hematopathologist to confirm the diagnoses. All NHL subtypes were not included in each study, either by design or because that subtype could not be reliably identified based on the available pathology data.

‡ Median age at diagnosis (cases) or interview (controls).

Exposure Category[A]	Specific Exposure	Prevalence (%) Cases	Cntls	P_{ASSET}	P_H	Overall NHL OR (95% CI)
Family history of hematologic malignancy[B]	Any	9.1	5.2	1.6×10^{-22}	3.5×10^{-2}	1.72 (1.54 - 1.93)
	NHL	4.0	2.0	1.7×10^{-13}	5.2×10^{-1}	1.79 (1.51 - 2.13)
	Leukemia	4.2	2.8	1.3×10^{-11}	3.9×10^{-5}	1.51 (1.29 - 1.77)
	Multiple myeloma	0.7	0.4	7.5×10^{-4}	2.2×10^{-2}	1.77 (1.15 - 2.72)
	Hodgkin lymphoma	1.1	0.6	2.0×10^{-3}	4.7×10^{-1}	1.65 (1.18 - 2.29)
Autoimmune disease[C]	Any B-cell activating disease	0.9	0.8	3.8×10^{-22}	9.8×10^{-10}	1.96 (1.60 - 2.40)
	Sjögren's syndrome	0.6	0.1	6.3×10^{-18}	7.3×10^{-9}	7.52 (3.68 - 15.4)
	Systemic lupus erythematosus	0.5	0.2	1.9×10^{-8}	1.8×10^{-1}	2.83 (1.82 - 4.41)
	Any T-cell activating disease	3.4	3.3	5.3×10^{-3}	1.2×10^{-2}	1.07 (0.95 - 1.21)
	Celiac disease	0.4	0.2	5.2×10^{-11}	5.1×10^{-8}	1.77 (1.05 - 2.99)
	Systemic sclerosis/scleroderma	0.1	0.1	5.1×10^{-3}	6.5×10^{-2}	1.03 (0.41 - 2.58)
HCV seropositivity[D]		2.3	2.2	2.3×10^{-8}	2.1×10^{-3}	1.81 (1.39 - 2.37)
Atopic disease[E]	Hay fever	18.2	20.1	9.1×10^{-9}	1.2×10^{-1}	0.82 (0.77 - 0.88)
	Eczema	9.8	9.8	5.0×10^{-5}	2.6×10^{-5}	1.01 (0.93 - 1.10)
	Allergy	22.0	24.4	5.9×10^{-5}	2.4×10^{-1}	0.86 (0.81 - 0.92)
Blood transfusion[F]	Transfusion occurring <1990	14.2	15.5	5.0×10^{-5}	1.3×10^{-2}	0.76 (0.67 - 0.87)
Anthropometric factors[G]	Body mass index as a young adult	21.1	17.9	4.2×10^{-9}	2.8×10^{-1}	1.95 (1.51 - 2.53)
	Height	53.2	52.0	1.7×10^{-3}	2.4×10^{-2}	1.20 (1.08 - 1.32)
Alcohol consumption (≥1 drink per month)	Any alcohol	69.3	72.1	8.9×10^{-8}	6.2×10^{-2}	0.87 (0.81 - 0.93)
	Wine	56.8	57.5	4.9×10^{-9}	1.4×10^{-2}	0.85 (0.79 - 0.91)
	Liquor	37.0	39.9	4.1×10^{-6}	6.6×10^{-1}	0.84 (0.78 - 0.91)
	Beer	44.9	47.2	9.3×10^{-4}	1.4×10^{-1}	0.90 (0.84 - 0.97)
Cigarette smoking[G]	Duration of smoking	57.0	56.7	2.2×10^{-5}	3.2×10^{-9}	1.06 (0.99 - 1.14)
Recreational sun exposure[G]		49.9	53.0	2.7×10^{-6}	7.9×10^{-1}	0.74 (0.66 - 0.83)
Socioeconomic status[G]		43.8	41.1	3.4×10^{-5}	6.1×10^{-2}	0.88 (0.83 - 0.93)
Occupational history[H]	Teacher	8.6	10.0	5.6×10^{-4}	6.2×10^{-3}	0.86 (0.77 - 0.95)
	Painter	2.0	1.8	4.8×10^{-3}	8.6×10^{-2}	1.22 (0.99 - 1.51)
	General farm worker	4.3	3.4	8.2×10^{-3}	3.4×10^{-1}	1.28 (1.10 - 1.50)

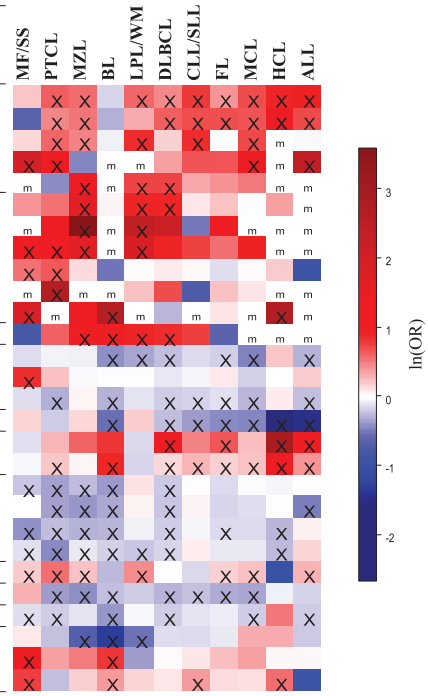


Figure 1. The table lists the overall odds ratio (OR) (95% confidence interval) for all risk factors affecting one or more non-Hodgkin lymphoma NHL subtypes (P_{ASSET} < 0.01), adjusting for age, race/ethnicity, sex, and study. For binary variables, OR compares exposed vs unexposed, and for ordinal variables[G], OR compares highest vs lowest category. The columns list the exposure category, specific exposure, prevalence (all variables dichotomized) in cases and controls, p-value for association (P_{ASSET}), p-value for effect homogeneity (PH), and the OR. The colored grid indicates the log odds ratio associated with the exposure for each subtype separately. **Red (blue)** indicates the exposure increases (decreases) risk. **X** indicates ASSET analysis identified a statistically significant association, whereas **m** indicates missing due to lack of data. For groups of highly correlated exposures (e.g., duration, pack-years smoking), only a single representative variable is listed here. Results for all risk factors are available in Supplementary Table 2 (available online). Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL). [A] In total, the family history category included 5 variables; autoimmune disease – 16; atopic disease – 5; blood transfusion – 5; anthropometric factors – 5; alcohol – 19, smoking – 7, sun – 2, occupation – 33; hair-dye – 8; reproductive and hormone – 5. [B] Type of hematologic malignancy was coded according to International Classification of Diseases (ICD) as non-Hodgkin lymphoma (NHL) (ICD-9: 200, 202.0-202.2, 202.8-202.9; ICD-10: C82-C85, C96.3), Hodgkin lymphoma (ICD-9: 201, ICD-10: C81), leukemia (ICD-9: 202.4, 203.1, 204-208; ICD-10: C90.1, C91-C95), or multiple myeloma (ICD-9: 203, ICD-10: C90.0, C90.2)). Note that leukemia includes both lymphoid and myeloid leukemias, and lymphoid leukemias and plasma cell neoplasms are not considered part of NHL in ICD, in contrast to the World Health Organization (WHO) classification (2,3) and InterLymph guidelines (42,43). [C] Includes self-reported history of specific autoimmune diseases occurring ≥2 years prior to diagnosis/interview (except the New South Wales study, which did not ascertain date of onset). Autoimmune diseases were classified according to whether they are primary mediated by B-cell or T-cell responses (21,54-57). B-cell activating diseases include Hashimoto thyroiditis, hemolytic anemia, myasthenia gravis, pernicious anemia, rheumatoid arthritis, Sjögren's syndrome, and systemic lupus erythematosus. T-cell activating disease include celiac disease, immune thrombocytopenic purpura, inflammatory bowel disorder (Crohn's disease, ulcerative colitis), multiple sclerosis, polymyositis or dermatomyositis, psoriasis, sarcoidosis, systemic sclerosis or scleroderma, and type 1 diabetes. [D] Serum antibodies to HCV were evaluated using a third generation enzyme-linked immunosorbent assay (58). [E] Includes self-reported history of atopic conditions occurring ≥2 years prior to diagnosis/interview. Any allergy included plant, food, animal, dust, insect, or mold, but excluded drug allergies. [F] Includes self-reported history of blood transfusions occurring ≥1 year prior to diagnosis/interview. [G] OR represents risk per increasing category of an ordinal variable with categories assigned to equally spaced values between 0 and 1 for body-mass index as a young adult (<18.5, 18.5-22.4, 22.5-24.9, 25.0-29.9, ≥30 kg/m2), height (sex-specific quartiles, males: <172.0, 172.0-177.7, 177.8-181.9, ≥182.0 cm; females: <159.0, 159.0-162.9, 163.0-167.9, ≥168.0 cm), duration of cigarette smoking (0, 1-19, 20-29, 30-39, ≥40 years), recreational sun exposure (hours per week, study-specific quartiles available upon request), and socioeconomic status (low, medium, high; measured by years of education for studies in North America or by dividing measures of education or socioeconomic status into tertiles for studies in Europe or Australia). [H] Occupations (ascertained by complete work history in 8 studies and longest held occupation in 2 studies) were coded according to the International Standard Classification of Occupations (ISCO), Revised Edition 1968 (59).

Specifically, we tested whether the exposure was associated with at least one subtype using ASSET, a subset-based statistical approach (44). ASSET is designed for studies evaluating exposures with multiple related outcomes, such as multiple NHL subtypes. The method has increased statistical power when the exposure is only associated with a subset of outcomes. ASSET gains this advantage by testing for an association with each subset of outcomes. For a given exposure, our first step in this analysis was to collect the Z-statistics ($\hat{\beta}_k / \sqrt{\hat{\sigma}_k^2}$) from the logistic regressions performed separately for each NHL subtype. We then calculated $Z_M = \max_S \left(\left| \sum_{k \in S} w_k Z_k \right| \right)$, where the weights ($w_k$) depended on the number of subjects and S was a set of subtypes. We identified those subtypes in S^*, where $S^* = \text{argmax}_S \left(\left| \sum_{k \in S} w_k Z_k \right| \right)$, as being putatively associated with the exposure and then calculated a P value, P_{ASSET}, for the significance of Z_M by permutation. Exposures with P_{ASSET} < .01 are included in Figure 1, and the NHL subtype(s) putatively associated with the exposure are marked with an "X".

We then measured the variability in the ORs among NHL subtypes by the Q value (45), $Q = \sum_k w_k (\hat{\beta}_k - \bar{\beta})^2$, where $\hat{\beta}_k$ and $\hat{\sigma}_k^2$ were the estimates of the log(OR) and its variance for subtype k, $\bar{\beta} = \sum_\kappa w_\kappa \hat{\beta}_\kappa$, and $w_k = \left(\sum_k (1/\hat{\sigma}_k^2) \right)^{-1} (1/\hat{\sigma}_k^2)$. We obtained a P value, $P_{\text{HOMOGENEITY}}$, by comparing Q to a χ^2 distribution with $K - 1$ degrees of freedom, where K was the number of studies measuring that exposure.

Finally, we clustered subtypes into groups that shared similar associations with each putative risk factor, or with the total collection of risk factors, using a divisive or "top-down" hierarchical clustering method specifically designed for this study. Again, let S be a set of subtypes and S^c be its complement. Let $Y_s = 1$ and $Y_s = 0$ if a case was diagnosed with a subtype in sets S and S^c, respectively, with Y_s set to missing for all controls. Let p_s be the P value from a case-only logistic regression of Y_s on the risk factor of interest, adjusting for age, race/ethnicity, sex, and study. Let $P_M = \min_S(p_s)$. Then we clustered the subtypes into two groups, S^* and S^{c*}, where $S^* = \text{argmin}_S(p_s)$. We defined P_{NODE} to be the probability that P_M was below the observed value under the null hypothesis and calculated it by 10 000 permutations of subtype assignment. We repeated this clustering procedure on S^* and S^{c*} to continue building the tree. Because the rare subtypes ALL and hairy cell leukemia (N ~ 150 cases) could not be assigned reliably to clusters, we omitted them from this analysis.

When clustering subtypes according to all risk factors (Figure 6), we used a different method for calculating p_s. Had each study included all NHL subtypes and exposures, we could have used the P value from a Wald statistic produced by a single logistic regression. Instead, we used a pseudo-Wald statistic where the log(OR) for each exposure was estimated from a separate analysis. Let $\hat{\beta}_{skj}$ be the parameter from a logistic regression of Y_S on exposure j (adjusting for age, race/ethnicity, sex, and study) in study k, $\hat{\sigma}^2_{skij}$ estimate the covariance between $\hat{\beta}_{ski}$ and $\hat{\beta}_{skj}$, and $\delta_{kj} = 1$ if study k includes exposure j. Then we defined

$$\hat{\beta}_{sj} = \sum_{k:\delta_{kj}=1} w_{skj} \hat{\beta}_{skj}$$

where the weights (w_{skj}) were inversely proportional to the estimated variance

$$w_{skj} = \frac{1/\hat{\sigma}^2_{skjj}}{\sum_{k:\delta_{kj}=1} 1/\hat{\sigma}^2_{skjj}}$$

and we estimated the covariance of $\hat{\beta}^t_s = \{\hat{\beta}_{s1}, \ldots, \hat{\beta}_{sN}\}$ by the $N \times N$ matrix $\hat{\Sigma}$ with the i,jth entry defined as

$$\hat{\Sigma}_s[i,j] = \sum_{k:\delta_{skj}\cdot\delta_{ski}=1} w_{skj} w_{ski} \hat{\sigma}^2_{skij}$$

The resulting test statistic, $\hat{\beta}^t_s \hat{\Sigma}^{-1}_s \hat{\beta}_s$, was our pseudo-Wald statistic, which was compared to a χ^2 distribution with N degrees of freedom to obtain p_s.

Results

The pooled study population included 17 471 NHL cases and 23 096 controls derived from 14 population-based and six hospital/clinic-based case–control studies. The study population was predominantly male (58%) and non-Hispanic white (93%, Table 1). DLBCL (N = 4667) was the most common and acute lymphoblastic leukemia/lymphoma (ALL, N = 152) was the least common NHL subtype included in this analysis. Hairy cell leukemia cases had the most striking male predominance (79%), whereas marginal zone lymphoma cases had the least (47%). The median age at diagnosis ranged from 41 years for ALL cases to 64 years for CLL/SLL and LPL/WM cases.

Risk Factors for One or More NHL Subtypes

We identified family history, medical history, lifestyle, and occupational risk factors that were associated with one or more NHL subtypes ($P_{\text{ASSET}} < .01$, Figure 1; Supplementary Table 2, available online, contains results for all risk factors). For highly correlated variables ($r > 0.8$; eg, duration and pack-years of smoking), we selected the variable with the smaller P_{ASSET}. The total number of variables we analyzed and the correlation among variables within each risk factor category are provided in Figure 1.

Family history of any hematologic malignancy in a first-degree relative was the most statistically significant risk factor ($P_{\text{ASSET}} = 1.6 \times 10^{-22}$), with associations observed for family history of NHL ($P_{\text{ASSET}} = 1.7 \times 10^{-13}$), leukemia ($P_{\text{ASSET}} = 1.3 \times 10^{-11}$), multiple myeloma ($P_{\text{ASSET}} = 7.5 \times 10^{-4}$), and Hodgkin lymphoma ($P_{\text{ASSET}} = .0020$). Some autoimmune diseases also were strongly associated with one or more NHL subtypes. The association for B-cell-activating autoimmune disease ($P_{\text{ASSET}} = 3.8 \times 10^{-22}$) was driven by Sjögren's syndrome ($P_{\text{ASSET}} = 6.3 \times 10^{-18}$) and systemic lupus erythematosus ($P_{\text{ASSET}} = 1.9 \times 10^{-8}$), whereas the association for T-cell-activating autoimmune disease ($P_{\text{ASSET}} = .0053$) was driven mainly by celiac disease ($P_{\text{ASSET}} = 5.2 \times 10^{-11}$) and also by systemic sclerosis/scleroderma ($P_{\text{ASSET}} = .0051$). Other medical history factors associated with one or more NHL subtypes included HCV seropositivity ($P_{\text{ASSET}} = 2.3 \times 10^{-8}$), hay fever ($P_{\text{ASSET}} = 9.1 \times 10^{-9}$), eczema ($P_{\text{ASSET}} = 5.0 \times 10^{-5}$), allergy ($P_{\text{ASSET}} = 5.9 \times 10^{-5}$), and blood transfusion before 1990 ($P_{\text{ASSET}} = 5.0 \times 10^{-5}$).

Among the lifestyle factors we examined, associations with one or more NHL subtypes were observed for body mass index as a young adult ($P_{\text{ASSET}} = 4.2 \times 10^{-9}$); height ($P_{\text{ASSET}} = .0017$); alcohol consumption ($P_{\text{ASSET}} = 8.9 \times 10^{-8}$), including wine ($P_{\text{ASSET}} = 4.9 \times 10^{-9}$), liquor ($P_{\text{ASSET}} = 4.1 \times 10^{-6}$), and beer ($P_{\text{ASSET}} = 9.3 \times 10^{-4}$); duration of cigarette smoking ($P_{\text{ASSET}} = 2.2 \times 10^{-5}$); recreational sun exposure ($P_{\text{ASSET}} = 2.7 \times 10^{-6}$); and socioeconomic status ($P_{\text{ASSET}} = 3.4 \times 10^{-5}$). Certain occupations also were associated with one or more NHL subtypes, specifically occupation as a teacher ($P_{\text{ASSET}} = 5.6 \times 10^{-4}$), painter ($P_{\text{ASSET}} = .0048$), or general farm worker ($P_{\text{ASSET}} = .0082$).

Effect of Heterogeneity Among NHL Subtypes for Specific Risk Factors

Among family history variables, the greatest heterogeneity among NHL subtypes was observed for family history of leukemia ($P_{\text{HOMOGENEITY}} = 3.9 \times 10^{-5}$), which increased risk 2.41-fold for CLL/SLL, 2.19 for LPL/WM, 1.98 for mantle cell, and 1.84 for PTCL

(P_{NODE} = 4.0×10^{-4}), versus weaker (OR = 1.66 for marginal zone lymphoma) or null associations for the other subtypes (Figure 2A). Risk associated with family history of multiple myeloma also was statistically significantly different among NHL subtypes ($P_{HOMOGENEITY}$ = .022), with particularly elevated risks for MF/SS (OR = 6.11, P_{NODE} = .027) compared with weaker or null associations (OR ≤ 3.10) for the other subtypes that were not statistically significantly heterogeneous (Figure 2B). In contrast, family history of NHL or HL increased risk for NHL overall by 1.79- and 1.65-fold, respectively, with no statistically significant heterogeneity in risks among NHL subtypes (NHL: $P_{HOMOGENEITY}$ = .52, P_{NODE} = .94; HL: $P_{HOMOGENEITY}$ = .47, P_{NODE} = .74; Supplementary Table 3, available online, provides the results of the clustering analysis for all risk factors with P_{ASSET} < .01 as listed in Figure 1).

Autoimmune diseases were relatively rare but were associated with the highest ORs for specific NHL subtypes. B-cell-activating autoimmune disease ($P_{HOMOGENEITY}$ = 9.8×10^{-10}) increased risk 5.46-fold for marginal zone lymphoma (P_{NODE} = 1.0×10^{-4}) and 2.61- and 2.45-fold for LPL/WM and DLBCL, respectively (P_{NODE} = .011, Figure 3A). Analyses of specific B-cell-activating autoimmune diseases revealed strikingly increased risk for marginal zone lymphoma associated with Sjögren's syndrome (OR = 38.07, $P_{HOMOGENEITY}$ = 7.3×10^{-9}, P_{NODE} < 1.0×10^{-4}), with weaker associations for LPL/WM (OR = 12.14) and the other subtypes (Figure 3B). ORs for systemic lupus erythematosus ranged from 1.81 to 8.41, but these differences did not reach statistical significance ($P_{HOMOGENEITY}$ = .18, P_{NODE} = .24). T-cell-activating autoimmune disease increased risk for PTCL and MF/SS (OR = 1.95 and 1.66, respectively, $P_{HOMOGENEITY}$ = .012, P_{NODE} = .0054, Figure 3C), with particularly elevated risk for PTCL associated with celiac disease (OR = 14.82, $P_{HOMOGENEITY}$ = 5.1×10^{-8}, P_{NODE} < 1.0×10^{-4}, Figure 3D). ORs for systemic sclerosis/scleroderma ranged from 0.71 to 20.16, but these differences did not reach statistical significance ($P_{HOMOGENEITY}$ = .065, P_{NODE} = .28).

Among the other medical history factors we evaluated, HCV-associated risks differed by NHL subtype ($P_{HOMOGENEITY}$ = .0021), with 3.05-fold increased risk for BL, 3.04 for marginal zone lymphoma, 2.70 for LPL/WM, and 2.33 for DLBCL (P_{NODE} = .010); 2.08-fold increased risk for CLL/SLL (P_{NODE} = .032); and no associations for other subtypes (Figure 4A). Eczema was associated with statistically significantly increased risk for MF/SS (OR = 2.31, $P_{HOMOGENEITY}$ = 2.6×10^{-5}, P_{NODE} < 1.0×10^{-4}) but no other NHL

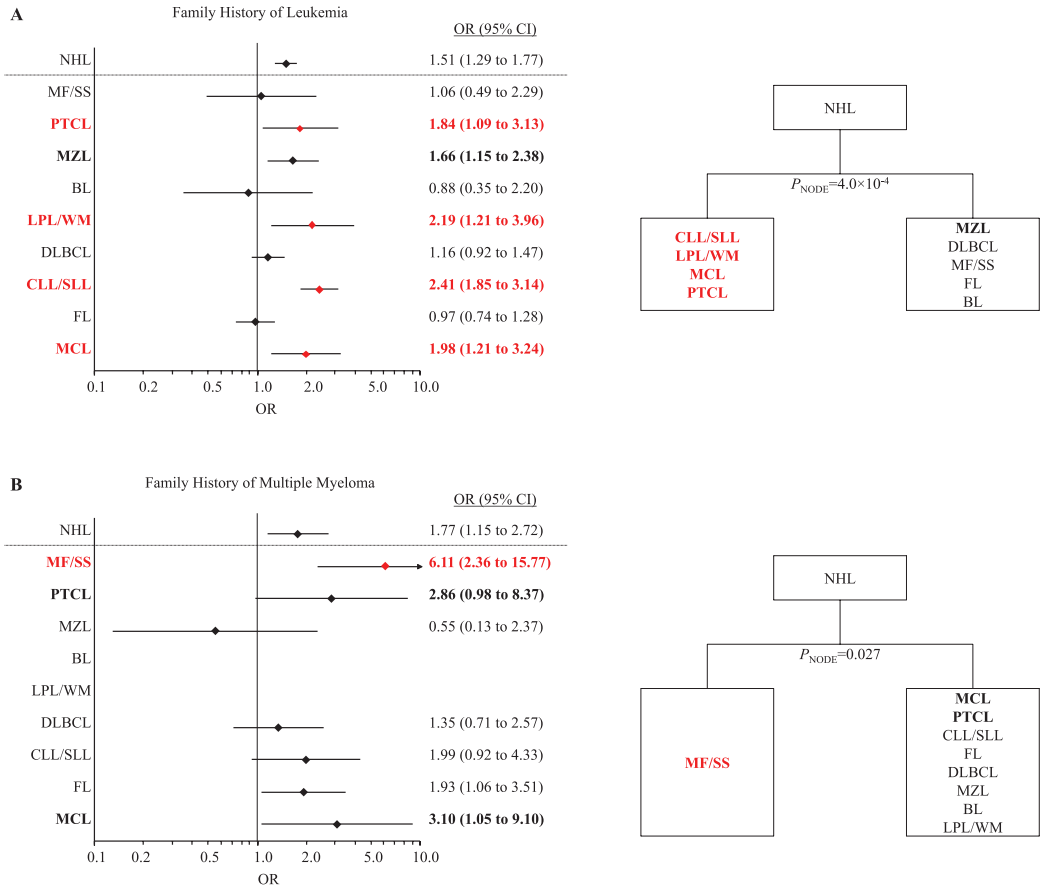


Figure 2. Forest plots list the odds ratio (OR) and 95% confidence interval (CI) for being diagnosed with non-Hodgkin lymphoma (NHL), or its specific subtypes, for individuals with a **(A)** family history of leukemia or **(B)** family history of multiple myeloma, compared to individuals without a family history. ORs were adjusted for age, ethnicity, sex, and study. **Bold font** indicates associated subtypes in ASSET and **colors** represent distinct tree nodes. The trees on the right of the figure split the NHL subtypes into groups of subtypes that were similarly affected by the given exposure. Hairy cell leukemia (HCL) and acute lymphoblastic leukemia/lymphoma (ALL) were excluded from trees because small sample sizes prevented reliable clustering. P_{NODE} is the *P*-value for creation of that node during hierarchical clustering. Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL).

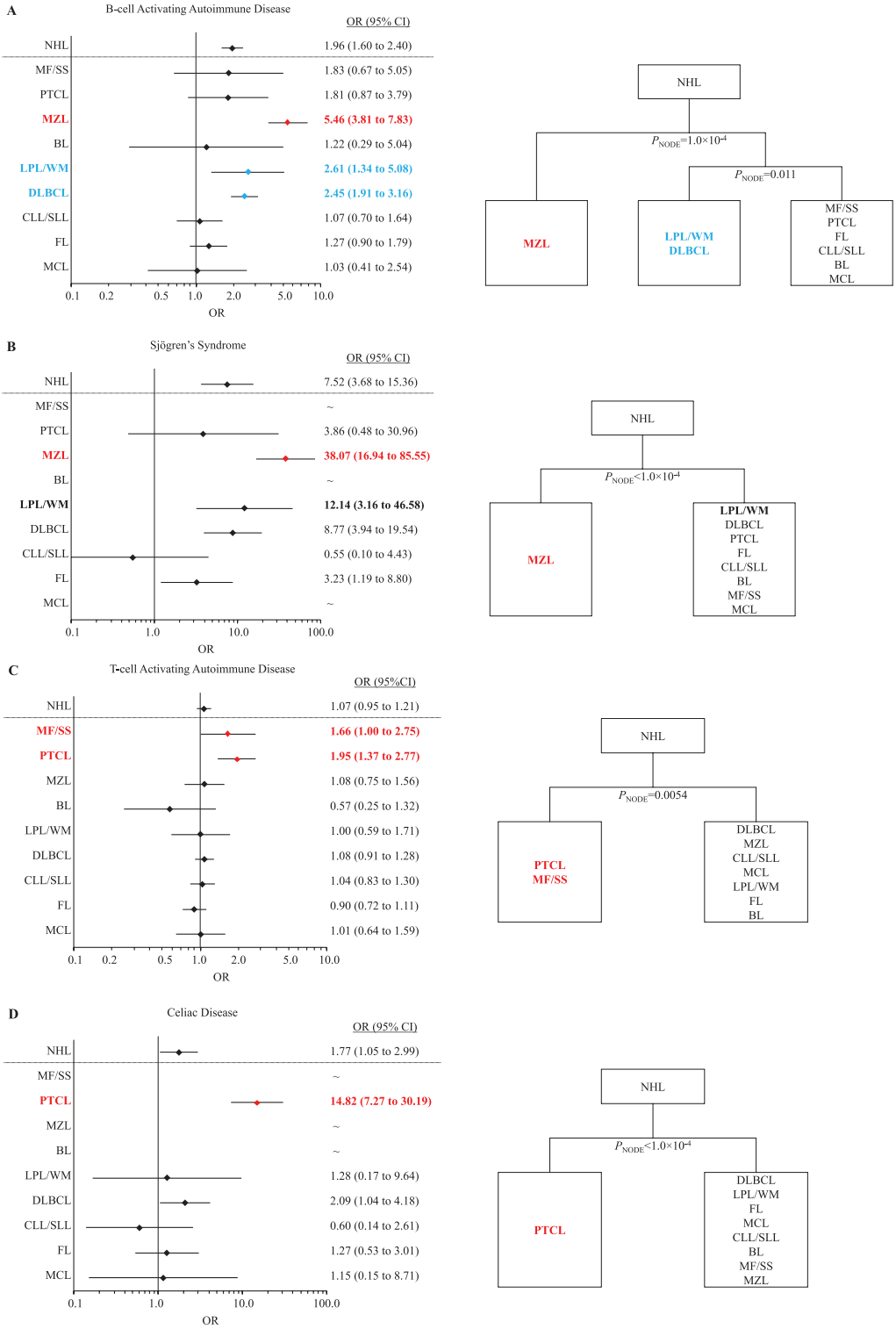


Figure 3. Forest plots list the odds ratio (OR) and 95% confidence interval (CI) for being diagnosed with non-Hodgkin lymphoma (NHL), or its specific subtypes, for individuals with a history of **(A)** B-cell-activating autoimmune disease, **(B)** Sjögren's syndrome, **(C)** T-cell-activating autoimmune disease, and **(D)** celiac disease, compared to individuals without a family history. ORs were adjusted for age, ethnicity, sex, and study. **Bold font** indicates associated subtypes in ASSET and **colors** represent distinct tree nodes. The trees on the right of the figure split the NHL subtypes into groups of subtypes that were similarly affected by the given exposure. Hairy cell leukemia (HCL) and acute lymphoblastic leukemia/lymphoma (ALL) were excluded from trees because small sample sizes prevented reliable clustering. P_{NODE} is the *P*-value for creation of that node during hierarchical clustering. Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL).

subtype (Figure 4B). The ORs for receipt of a blood transfusion before 1990 ranged from 0.57 to 0.84 for PTCL, DLBCL, CLL/SLL, follicular lymphoma, mantle cell lymphoma, and BL ($P_{HOMOGENEITY}$ = .013, P_{NODE} = .025), whereas the OR was nonsignificantly greater than 1 for MF/SS, LPL/WM, and marginal zone lymphoma (Figure 4C). In contrast, the inverse associations observed for NHL overall did not differ statistically significantly among NHL subtypes for hay fever (OR = 0.82, $P_{HOMOGENEITY}$ = .12, P_{NODE} = .36) and allergy (OR = 0.86, $P_{HOMOGENEITY}$ = .24, P_{NODE} = .084). In analyses of other putative medical history risk factors for NHL, peptic ulcer did not reach the threshold for significance in ASSET but demonstrated evidence for heterogeneity, with risk statistically significantly increased 1.55-fold for marginal zone lymphoma and no association observed for any other NHL subtype (P_{ASSET} = .058, $P_{HOMOGENEITY}$ = .034, P_{NODE} = .0057).

Lifestyle factors and occupations generally exhibited smaller ORs and less heterogeneity among NHL subtypes than medical history and family history factors although some differences were observed. The inverse association between alcohol consumption and NHL showed weak evidence of heterogeneity, with slightly stronger associations for DLBCL, BL, PTCL, and marginal zone lymphoma than other subtypes, particularly for wine consumption

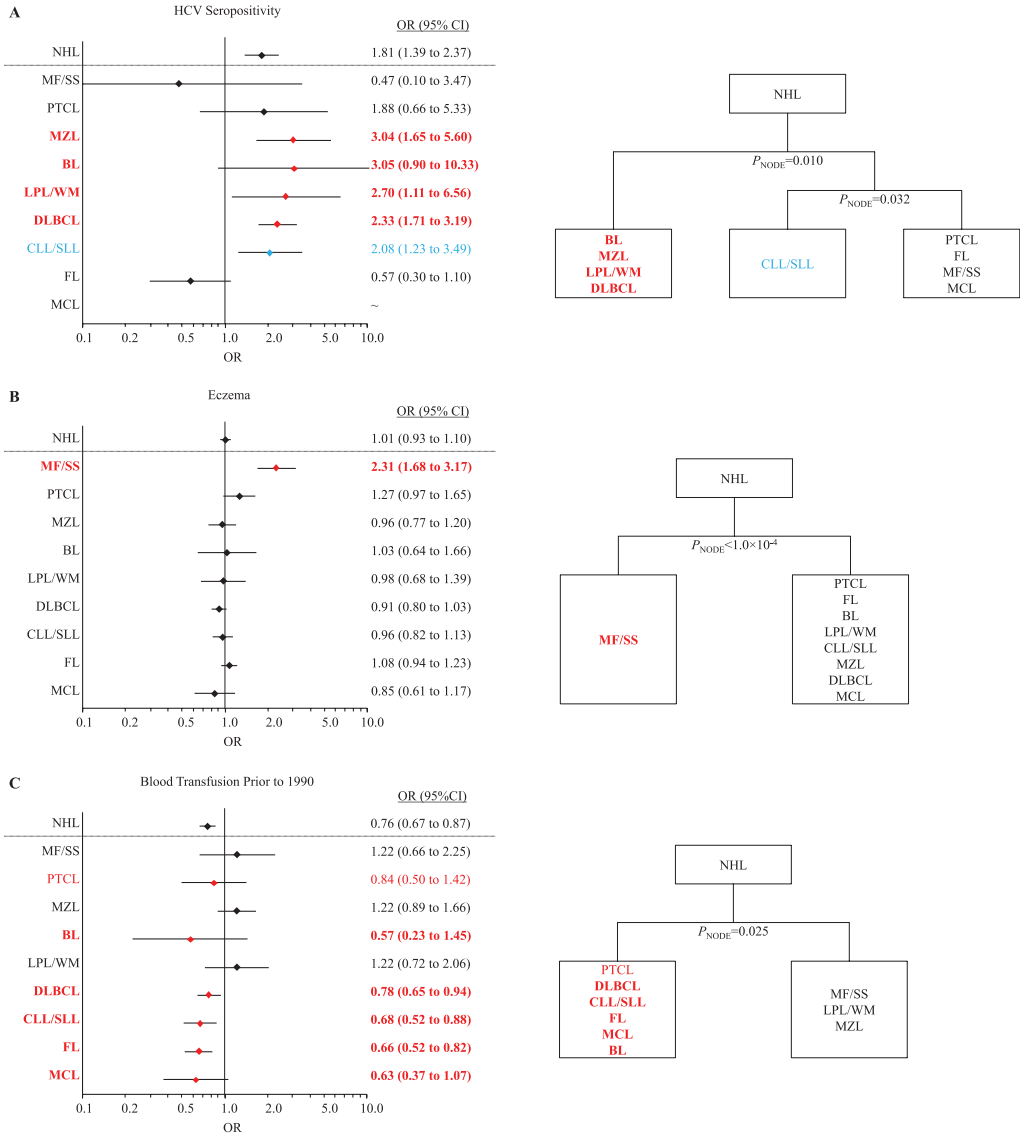


Figure 4. Forest plots list the odds ratio (OR) for being diagnosed with non-Hodgkin lymphoma (NHL), or its specific subtypes, for individuals with **(A)** hepatitis c virus (HCV) seropositivity, **(B)** eczema, and **(C)** blood transfusion prior to 1990, compared to individuals without that condition. ORs were adjusted for age, ethnicity, sex, and study. **Bold font** indicates associated subtypes in ASSET and **colors** represent distinct tree nodes. The trees on the right of the figure split the NHL subtypes into groups of subtypes that were similarly affected by the given exposure. Hairy cell leukemia (HCL) and acute lymphoblastic leukemia/lymphoma (ALL) were excluded from trees because small sample sizes prevented reliable clustering. P_{NODE} is the *P*-value for creation of that node during hierarchical clustering. Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL).

(ORs = 0.64–0.81, $P_{HOMOGENEITY}$ = .014, P_{NODE} = .098, Figure 5A). Increased duration of cigarette smoking was associated with the greatest increased risk for PTCL and LPL/WM (OR = 1.75 and 1.50, respectively, per increasing category of duration) and more modest increases for marginal zone lymphoma, mantle cell lymphoma, MF/SS, and follicular lymphoma (ORs = 1.19–1.27, $P_{HOMOGENEITY} = 3.2\times10^{-9}$, $P_{NODE} = 1.0\times10^{-4}$), whereas the OR was 1.02 for DLBCL, 0.84 for CLL/SLL, and 0.77 for BL (Figure 5B). Occupation as a teacher was inversely associated with LPL/WM, marginal zone lymphoma, and BL (ORs = 0.27–0.59, $P_{HOMOGENEITY}$ = .0062, P_{NODE} = .035, Figure 5C) but not other subtypes, whereas occupation as a painter increased risk for MF/SS and BL (ORs = 3.42 and 2.28, respectively, $P_{HOMOGENEITY}$ = .085, P_{NODE} = .023, Figure 5D). Usual adult body mass index did not reach the threshold for significance in ASSET but demonstrated some evidence for heterogeneity, with risk statistically significantly increased 1.95- and 1.32-fold for MF/SS and DLBCL, respectively, per increasing WHO category (P_{ASSET} = .018, $P_{HOMOGENEITY} = 3.1\times10^{-4}$, P_{NODE} = .015). For height, risks were statistically nonsignificantly higher for BL than other subtypes (OR = 2.43 per increasing sex-specific quartile versus OR = 1.20 for overall NHL, $P_{HOMOGENEITY}$ = .024, P_{NODE} = .26). In contrast, statistically significant variability among NHL subtypes was not observed for the positive associations for body mass index as a young adult (OR = 1.95 per increasing category of body mass index, $P_{HOMOGENEITY}$ = .28, P_{NODE} = .15) and occupation as a general farm worker (OR = 1.28, $P_{HOMOGENEITY}$ = .085, P_{NODE} = .20) or for the negative associations for recreational sun exposure (OR = 0.74 per increasing quartile of hours per week, $P_{HOMOGENEITY}$ = .79, P_{NODE} = .70) and socioeconomic status (OR = 0.88 per increasing tertile, $P_{HOMOGENEITY}$ = .061, P_{NODE} = .45).

Other putative NHL risk factors that we evaluated, including measures of history of living and/or working on a farm, personal, and/or occupational exposure to hair dye, hormonal/reproductive factors, and occupations other than those listed above, did not reach the threshold for significance in ASSET (P_{ASSET} < .01) and showed no clear evidence of heterogeneity among the NHL subtypes (Supplementary Table 2, available online).

Overall Risk Factor Pattern Among NHL Subtypes

Although the specific patterns of association among NHL subtypes varied by exposure, when all risk factors were taken into account, we observed statistically significant clustering among subtypes. The greatest difference in risk factor patterns was between T-cell and B-cell lymphomas ($P_{NODE} < 1.0\times10^{-4}$, Figure 6). Eczema, occupation as a painter, T-cell-activating autoimmune diseases, family history of multiple myeloma, and cigarette smoking were all more strongly associated with risk for T-cell than B-cell lymphomas although some of these factors were not exclusively associated with T-cell lymphomas. MF/SS and PTCL also were different from one another due to the striking association of eczema with MF/SS (P_{NODE} = .058). Additionally, substantial heterogeneity was observed among B-cell lymphomas for the risk factors that we evaluated, with the tree first separating marginal zone lymphoma and BL ($P_{NODE} < 1.0\times10^{-4}$), then follicular lymphoma and mantle cell lymphoma (P_{NODE} = .017), and finally DLBCL and LPL/WM suggestively separating from CLL/SLL (P_{NODE} = .062). Key risk factors differentiating B-cell NHL subtypes included B-cell-activating autoimmune diseases, hay fever, allergy, alcohol consumption, HCV seropositivity, cigarette smoking, and occupation as a teacher or general farm worker.

Discussion

In this large-scale, international collaborative study, we provide the first comprehensive effort to quantitatively compare similarities and differences in postulated risk factors among both common and rarer NHL subtypes. Based on a novel methodological approach to cluster NHL subtypes according to a broad spectrum of risk factors, the majority of risk factors showed differences in risk among NHL subtypes, whereas fewer factors showed consistent risks among subtypes. Overall, this approach most strongly distinguished T-cell from B-cell lymphomas, with additional heterogeneity among specific types of B-cell lymphoma, although the patterns of effect heterogeneity varied substantially for the different risk factors. These results synthesize the highly detailed analyses of risk factors for individual subtypes discussed elsewhere in this issue (15–17,33–40) and expand previous InterLymph pooled analyses by including data from additional studies and/or reporting risks for rarer NHL subtypes (19,21,24–32).

Our clustering results support the relatively greater importance of immune perturbation in the etiologies of PTCL, marginal zone lymphoma, BL, DLBCL, and LPL/WM compared with MF/SS, CLL/SLL, follicular lymphoma, and mantle cell lymphoma. We found that HCV, autoimmune diseases, and peptic ulcer (a proxy for *Helicobacter pylori* infection), which have previously been reported as NHL risk factors and are thought to increase lymphoma risk through chronic antigenic stimulation (18,46,47), were predominantly associated with PTCL, marginal zone lymphoma, BL, DLBCL, and/or LPL/WM. The importance of immune perturbation is further supported by 1) the patterns of association for autoimmune diseases, whereby B-cell-activating autoimmune diseases were most strongly associated with certain B-cell NHLs and T-cell-activating autoimmune diseases with T-cell NHLs and 2) the particularly elevated site-specific risks associated with autoimmune diseases localized to specific organs, as reported in the analyses for marginal zone lymphoma, PTCL, and DLBCL [eg, celiac disease with enteropathy-type PTCL (15–17)]. Intriguingly, our finding that alcohol consumption and occupation as a teacher were most closely associated with some of these same NHL subtypes raises the hypothesis that these factors also may influence lymphoma risk via an immune-related mechanism. Our observations are consistent with the NHL subtype-specific risks observed in solid organ transplant recipients and individuals with HIV/AIDS, where lymphoma risk is thought to be related to reduced control of lymphomagenic viruses such as Epstein–Barr virus, decreased immunosurveillance capability, and immune activation (8–13,48–51). However, variability in the specific immune-related risk factor associations within this group of NHL subtypes suggests that further research is needed to better understand the specific immune perturbations that contribute to each subtype.

Other risk factors that we evaluated—including family history of leukemia or multiple myeloma, cigarette smoking, some anthropometric measures, blood transfusions, and certain

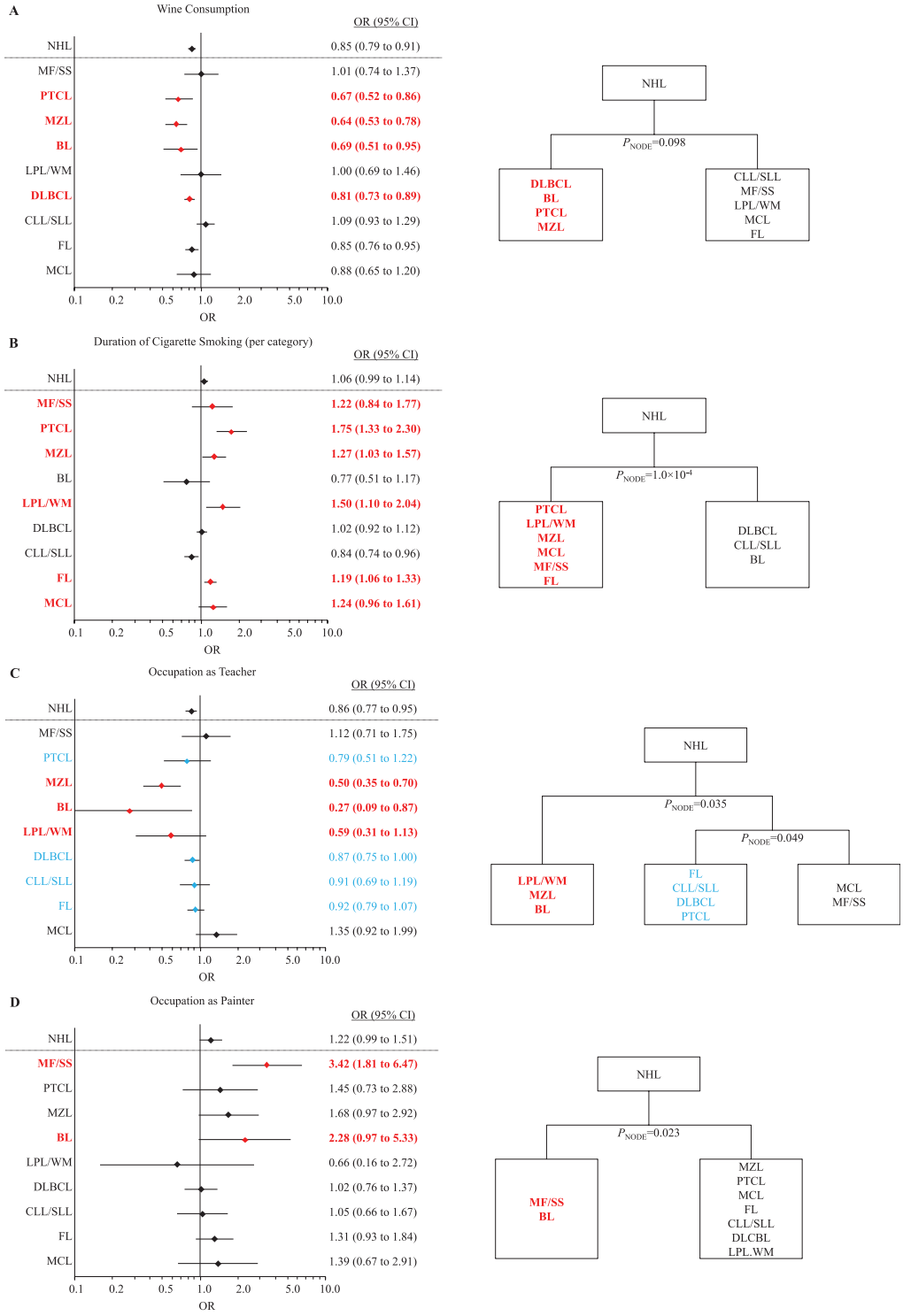
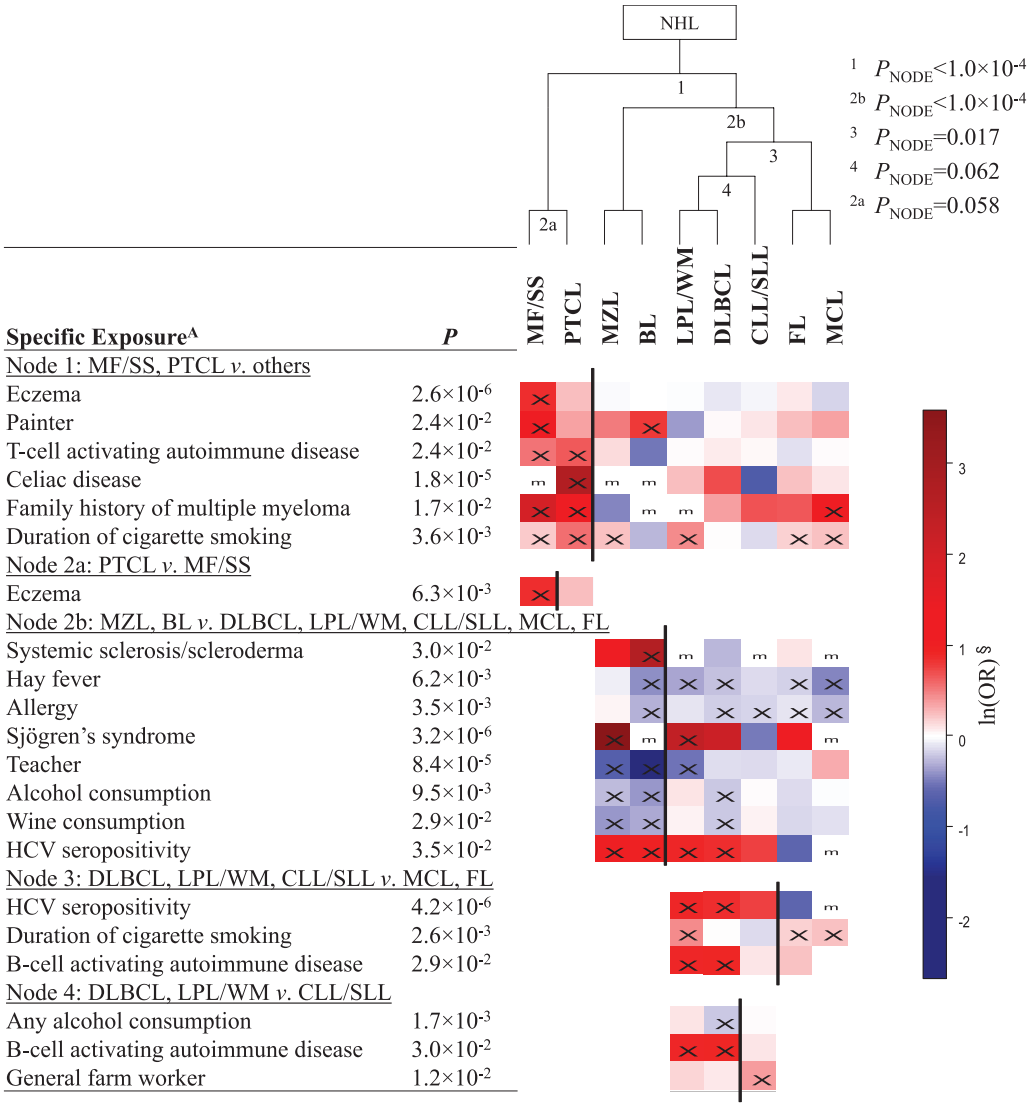


Figure 5. Forest plots list the odds ratio (OR) for being diagnosed with non-Hodgkin lymphoma (NHL), or its specific subtypes, for individuals **(A)** consuming ≥ 1 serving of wine/month; **(B)** smoking longer, smoking duration categorized into groupings of 0, 1–19, 20–29, 30–39, and ≥40 years, with assigned values of 0, 1/4, 2/4, 3/4, and 1 for calculating OR; **(C)** occupation as teacher; and **(D)** occupation as Painter. ORs were adjusted for age, ethnicity, sex, and study. **Bold font** indicates associated subtypes in ASSET and **colors** represent distinct tree nodes. The trees on the right of the figure split the NHL subtypes into groups of subtypes that were similarly affected by the given exposure. Hairy cell leukemia (HCL) and acute lymphoblastic leukemia/lymphoma (ALL) were excluded from trees because small sample sizes prevented reliable clustering. P_{NODE} is the *P*-value for creation of that node during hierarchical clustering. Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL).



Specific Exposure[A]	*P*
Node 1: MF/SS, PTCL *v.* others	
Eczema	2.6×10-6
Painter	2.4×10-2
T-cell activating autoimmune disease	2.4×10-2
Celiac disease	1.8×10-5
Family history of multiple myeloma	1.7×10-2
Duration of cigarette smoking	3.6×10-3
Node 2a: PTCL *v.* MF/SS	
Eczema	6.3×10-3
Node 2b: MZL, BL *v.* DLBCL, LPL/WM, CLL/SLL, MCL, FL	
Systemic sclerosis/scleroderma	3.0×10-2
Hay fever	6.2×10-3
Allergy	3.5×10-3
Sjögren's syndrome	3.2×10-6
Teacher	8.4×10-5
Alcohol consumption	9.5×10-3
Wine consumption	2.9×10-2
HCV seropositivity	3.5×10-2
Node 3: DLBCL, LPL/WM, CLL/SLL *v.* MCL, FL	
HCV seropositivity	4.2×10-6
Duration of cigarette smoking	2.6×10-3
B-cell activating autoimmune disease	2.9×10-2
Node 4: DLBCL, LPL/WM *v.* CLL/SLL	
Any alcohol consumption	1.7×10-3
B-cell activating autoimmune disease	3.0×10-2
General farm worker	1.2×10-2

Figure 6. Top-down heierarchical clustering identified groups of subtypes that had similar risk profiles among significant exposures (P_{ASSET} < 0.01). The tree at the top of the figure illustrates that the first split separated MF/SS and PTCL from the remaining seven subtypes, the second split further divided that larger group, separating MZL and BL from the remaining five subtypes, and so forth. For each split, the table lists the risk factors that distinguish the subtypes in the two resulting nodes at a statistically significant level (p < .05) and the **colored grid** (similar to Figure 1) indicates the odds ratios for the relevant subtype/risk factor pairings. P_{NODE} is the *P*-value for creation of that node during hierarchical clustering. Hairy cell leukemia (HCL) and acute lymphoblastic leukemia/lymphoma (ALL) were excluded from the tree because small sample sizes prevented reliable clustering. Subtypes include Burkitt/Burkitt-like lymphoma/leukemia (BL); chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL); diffuse large B-cell lymphoma (DLBCL); follicular lymphoma (FL); lymphoplasmacytic lymphoma/Waldenström macroglobulinemia (LPL/WM); mantle cell lymphoma (MCL); marginal zone lymphoma (MZL); mycosis fungoides/Sézary syndrome (MF/SS); peripheral T-cell lymphoma (PTCL). [A] Details regarding specific risk factors are provided in the footnote for Figure 1.

occupations—demonstrated heterogeneity among NHL subtypes but no consistent patterns emerged. Detailed consideration of these observed associations and potential biological mechanisms are presented in the NHL subtype-specific analyses in this issue (15–17,33–40). By conducting this analysis among subtypes, two key observations arose. First, our results clearly demonstrated that there is etiologic heterogeneity among NHL subtypes for numerous, but not all, risk factors. However, the inconsistency of some of the patterns suggests that further research is needed to identify the characteristics that may lead to shared etiology among NHL subtypes defined by the WHO classification. Investigation of molecular characteristics is a particularly promising avenue. Molecular characterization of lymphomas has revealed distinct subtypes within existing entities (eg, activated vs germinal center B-cell DLBCL), as well as certain molecular characteristics that may cut across existing entities [eg, Epstein–Barr virus infection, t(14;18) translocations, double-hit lymphomas (52)]. Future research on NHL etiology should explore the potential for relating specific exposures to molecular subtypes of disease. Second, we observed relatively modest associations for many of the risk factors evaluated herein, particularly for lifestyle factors and occupation. Future studies should refine exposure assessment, such as considering relevant periods of exposure, gene–environment interaction, and biomarkers rather than self-reported exposures, and expand research to include other factors not assessed here, such as dietary factors or specific chemicals.

This analysis exemplifies the benefits of international consortial collaboration. Inclusion of more than 17 000 NHL cases provided sufficient statistical power to investigate the etiology of common and rarer NHL subtypes. Across the broad range of exposures we considered in this analysis, we provide the strongest evidence to date of the importance of family history of hematologic malignancy and certain medical conditions, environmental and lifestyle factors, and occupations in lymphoma etiology. Centralized data harmonization with rigorous quality control ensured standardized NHL subtype definitions and exposure variables among studies. Three complementary statistical approaches were used to identify risk factors that were robustly associated with one or more NHL subtypes, quantify the magnitude of the associations, and identify NHL subtypes with similar risk factor patterns. These approaches accounted for the complex pattern of missing data among studies and different sample sizes among NHL subtypes and used permutation-based *P* values to reduce the chance of false positive results. Subtype-specific reports published elsewhere in this issue (15–17,33–40) demonstrate that individual risk factors associated with each subtype generally were independent of one another and that, on the whole, interstudy heterogeneity in risks was not evident despite some differences in exposure prevalence among studies (Supplementary Table 4, available online).

Several key limitations of this project should be considered in the interpretation of our results. It was not feasible to centrally review original pathology reports and materials for all cases, and 30% of the cases were not originally classified according to the WHO. However, each participating study's pathology review procedures, rules for NHL subtype classification, and NHL subtype distribution were reviewed by an interdisciplinary team of pathologists and epidemiologists to ensure that subtype definitions were as consistent as possible among studies and with the WHO classification. Also, the subtype-specific reports confirmed that findings were consistent when restricted to cases classified by the WHO. Despite the large sample size, risk estimates were still unstable for rarer exposures, and the numbers of cases for HCL and ALL were too small to include in the clustering analysis. As with all pooled analyses, data harmonization necessitated broadening of certain exposure categorizations and reduced ability to evaluate detailed exposure characteristics, which might have attenuated risk estimates, and we only considered potential risk factors that were available in at least four contributing studies. Additionally, widely varying sample size among exposures because of variability in data availability among studies may have affected our ability to detect heterogeneity for certain risk factors. Additional limitations inherent to case–control studies include potential for biased risk estimates due to biased study population selection, inaccurate recall of exposures and/or differential recall by cases and controls (53), and reverse causality because exposures were ascertained after disease onset.

In conclusion, we have demonstrated that the etiology of NHL is complex and multifactorial, with substantial heterogeneity among NHL subtypes. Of the risk factors considered in this analysis, most were associated with several subtypes, some were associated with nearly all subtypes, and very few were associated with only a single subtype. Our analysis supports the importance of pooling carefully harmonized data as well as utilizing novel statistical methods to assess risks for specific disease subtypes. Additional research is needed to investigate potential associations with other factors not included in these analyses, such as infectious agents other than HCV, specific environmental and occupational exposures, dietary factors, medications, and genetic susceptibility, particularly for CLL/SLL, follicular lymphoma, and mantle cell lymphoma, which were associated with relatively few risk factors in this analysis. The insights provided by the risk factor patterns that we observed should motivate future research into mechanisms of lymphomagenesis, particularly in understanding the specific immune perturbations that lead to risk of marginal zone lymphoma, BL, LPL/WM, DLBCL, and PTCL. Replication of our results in prospective studies will provide support for the causality of the associations we identified. Further research also is needed to evaluate potential differences in risks for population subgroups, such as by sex or race/ethnicity, and to consider heterogeneity within NHL subtypes, such as by anatomical site or molecular subtype, which is particularly important as our understanding of NHL subtypes continues to evolve. Finally, it will be important to evaluate potential joint effects of risk factors with genetic susceptibility.

References

1. Ferlay J, Soerjomataram L, Ervik M, et al. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]*. Lyon, France: International Agency for Research on Cancer; 2013. http://globocan.iarc.fr. Accessed February 22, 2014.
2. Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press; 2001.
3. Swerdlow SH, Campo E, Harris NL, et al., eds. *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed. Lyon, France: IARC Press; 2008.
4. Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992–2001. *Blood*. 2006;107(1):265–276.
5. Smith A, Howell D, Patmore R, Jack A, Roman E. Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. *Br J Cancer*. 2011;105(11):1684–1692.
6. Sant M, Allemani C, Tereanu C, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood*. 2010;116(19):3724–3734.
7. van Leeuwen MT, Turner JJ, Joske DJ, et al. Lymphoid neoplasm incidence by WHO subtype in Australia 1982–2006 [published online ahead of print March 18, 2014]. *Int J Cancer*. 2014; doi:10.1002/ijc.28849.
8. Coté TR, Biggar RJ, Rosenberg PS, et al. Non-Hodgkin's lymphoma among people with AIDS: incidence, presentation and public health burden. AIDS/Cancer Study Group. *Int J Cancer*. 1997;73(5):645–650.
9. Biggar RJ, Engels EA, Frisch M, Goedert JJ; AIDS Cancer Match Registry Study Group. Risk of T-cell lymphomas in persons with AIDS. *J Acquir Immune Defic Syndr*. 2001;26(4):371–376.
10. Dal Maso L, Franceschi S. Epidemiology of non-Hodgkin lymphomas and other haemolymphopoietic neoplasms in people with AIDS. *Lancet Oncol*. 2003;4(2):110–119.
11. Quinlan SC, Morton LM, Pfeiffer RM, et al. Increased risk for lymphoid and myeloid neoplasms in elderly solid-organ transplant recipients. *Cancer Epidemiol Biomarkers Prev*. 2010;19(5):1229–1237.
12. Vajdic CM, van Leeuwen MT, Turner JJ, et al. No excess risk of follicular lymphoma in kidney transplant and HIV-related immunodeficiency. *Int J Cancer*. 2010;127(11):2732–2735.
13. Clarke CA, Morton LM, Lynch C, et al. Risk of lymphoma subtypes after solid organ transplantation in the United States. *Br J Cancer*. 2013;109(1):280–288.
14. Manns A, Hisada M, La Grenade L. Human T-lymphotropic virus type I infection. *Lancet*. 1999;353(9168):1951–1958.

15. Bracci PM, Benavente Y, Turner JJ, et al. Medical history, lifestyle, family history, and occupational risk factors for marginal zone lymphoma: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:52–65.
16. Cerhan JR, Kricker A, Paltiel O, et al. Medical history, lifestyle, family history, and occupational risk factors for diffuse large B-cell lymphoma: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:15-25.
17. Wang SS, Flowers CR, Kadin ME, et al. Medical history, lifestyle, family history, and occupational risk factors for peripheral T-cell lymphoma: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:66-75.
18. Peveling-Oberhag J, Arcaini L, Hansmann ML, Zeuzem S. Hepatitis C-associated B-cell non-Hodgkin lymphomas. Epidemiology, molecular signature and clinical management. *J Hepatol*. 2013;59(1):169–177.
19. de Sanjose S, Benavente Y, Vajdic CM, et al. Hepatitis C and non-Hodgkin lymphoma among 4784 cases and 6269 controls from the International Lymphoma Epidemiology Consortium. *Clin Gastroenterol Hepatol*. 2008;6(4):451–458.
20. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. *Nat Rev Cancer*. 2004;4(10):757–768.
21. Smedby KE, Vajdic CM, Falster M, et al. Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium. *Blood*. 2008;111(8):4029–4038.
22. Morton LM, Wang SS, Cozen W, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes. *Blood*. 2008;112(13):5150–5160.
23. Boffetta P, Armstrong B, Linet M, Kasten C, Cozen W, Hartge P. Consortia in cancer epidemiology: lessons from InterLymph. *Cancer Epidemiol Biomarkers Prev*. 2007;16(2):197–199.
24. Morton LM, Hartge P, Holford TR, et al. Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (InterLymph). *Cancer Epidemiol Biomarkers Prev*. 2005;14(4):925–933.
25. Morton LM, Zheng T, Holford TR, et al. Alcohol consumption and risk of non-Hodgkin lymphoma: a pooled analysis. *Lancet Oncol*. 2005;6(7):469–476.
26. Wang SS, Slager SL, Brennan P, et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). *Blood*. 2007;109(8):3479–3488.
27. Kricker A, Armstrong BK, Hughes AM, et al. Personal sun exposure and risk of non Hodgkin lymphoma: a pooled analysis from the Interlymph Consortium. *Int J Cancer*. 2008;122(1):144–154.
28. Willett EV, Morton LM, Hartge P, et al. Non-Hodgkin lymphoma and obesity: a pooled analysis from the InterLymph Consortium. *Int J Cancer*. 2008;122(9):2062–2070.
29. Vajdic CM, Falster MO, de Sanjose S, et al. Atopic disease and risk of non-Hodgkin lymphoma: an InterLymph pooled analysis. *Cancer Res*. 2009;69(16):6482–6489.
30. Kane EV, Roman E, Becker N, et al. Menstrual and reproductive factors, and hormonal contraception use: associations with non-Hodgkin lymphoma in a pooled analysis of InterLymph case-control studies. *Ann Oncol*. 2012;23(9):2362–2374.
31. Kane EV, Bernstein L, Bracci PM, et al. Postmenopausal hormone therapy and non-Hodgkin lymphoma: a pooled analysis of InterLymph case-control studies. *Ann Oncol*. 2013;24(2):433–441.
32. Zhang Y, de Sanjosé S, Bracci PM, et al. Personal use of hair dye and the risk of certain subtypes of non-Hodgkin lymphoma. *Am J Epidemiol*. 2008;167(11):1321–1331.
33. Linet MS, Vajdic CM, Morton LM, et al. Medical history, lifestyle, family history, and occupational risk factors for follicular lymphoma: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:26–40.
34. Slager SL, Benavente Y, Blair A, 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; 48:41–51.
35. Smedby KE, Sampson JN, Turner JJ, et al. Medical history, lifestyle, family history, and occupational risk factors for mantle cell lymphoma: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:76–86.
36. Vajdic CM, Landgren O, McMaster ML, et al. Medical history, lifestyle, family history, and occupational risk factors for lymphoplasmacytic lymphoma/Waldenström's macroglobulinemia: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:87–97.
37. Aschebrook-Kilfoy B, Cocco P, La Vecchia C, et al. Medical history, lifestyle, family history, and occupational risk factors for mycosis fungoides and Sezary syndrome: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:98–105.
38. Mbulaiteye SM, Morton LM, Sampson JN, et al. Medical history, lifestyle, family history, and occupational risk factors for sporadic Burkitt lymphoma/leukemia: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; (48):106–114.
39. Monnereau A, Slager SL, Hughes AM, et al. Medical history, lifestyle, and occupational risk factors for hairy cell leukemia: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:115–124.
40. Skibola CF, Slager SL, Berndt SI, et al. Medical history, lifestyle, family history, and occupational risk factors for adult acute lymphocytic leukemia: The InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:125–129.
41. Morton LM, Sampson JN, Cerhan JR, et al. Rationale and design of the International Lymphoma Epidemiology Consortium (InterLymph) Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014; 48:1–14.
42. Morton LM, Turner JJ, Cerhan JR, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the International Lymphoma Epidemiology Consortium (InterLymph). *Blood*. 2007;110(2):695–708.
43. Turner JJ, Morton LM, Linet MS, et al. InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. *Blood*. 2010;116(20):e90–e98.
44. Bhattacharjee S, Rajaraman P, Jacobs KB, et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am J Hum Genet*. 2012;90(5):821–835.
45. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–1558.
46. Anderson LA, Gadalla S, Morton LM, et al. Population-based study of autoimmune conditions and the risk of specific lymphoid malignancies. *Int J Cancer*. 2009;125(2):398–405.
47. Suarez F, Lortholary O, Hermine O, Lecuit M. Infection-associated lymphomas derived from marginal zone B cells: a model of antigen-driven lymphoproliferation. *Blood*. 2006;107(8):3034–3044.
48. Grulich AE, van Leeuwen MT, Falster MO, Vajdic CM. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. *Lancet*. 2007;370(9581):59–67.
49. Biggar RJ, Chaturvedi AK, Goedert JJ, Engels EA; HIV/AIDS Cancer Match Study. AIDS-related cancer and severity of immunosuppression in persons with AIDS. *J Natl Cancer Inst*. 2007;99(12):962–972.
50. Engels EA, Biggar RJ, Hall HI, et al. Cancer risk in people infected with human immunodeficiency virus in the United States. *Int J Cancer*. 2008;123(1):187–194.
51. Mbulaiteye SM, Clarke CA, Morton LM, et al. Burkitt lymphoma risk in U.S. solid organ transplant recipients. *Am J Hematol*. 2013;88(4):245–250.
52. Aukema SM, Siebert R, Schuuring E, et al. Double-hit B-cell lymphomas. *Blood*. 2011;117(8):2319–2331.
53. Chang ET, Smedby KE, Hjalgrim H, Glimelius B, Adami HO. Reliability of self-reported family history of cancer in a large case-control study of lymphoma. *J Natl Cancer Inst*. 2006;98(1):61–68.
54. Sweet RA, Cullen JL, Shlomchik MJ. Rheumatoid factor B cell memory leads to rapid, switched antibody-forming cell responses. *J Immunol*. 2013;190(5):1974–1981.
55. Ballotti S, Chiarelli F, de Martino M. Autoimmunity: basic mechanisms and implications in endocrine diseases. Part II. *Horm Res*. 2006;66(3):142–152.

56. Zhang X, Ing S, Fraser A, et al. Follicular helper T cells: new insights into mechanisms of autoimmune diseases. *Ochsner J*. 2013;13(1):131–139.
57. Porakishvili N, Mageed R, Jamin C, et al. Recent progress in the understanding of B-cell functions in autoimmunity. *Scand J Immunol*. 2001;54(1–2):30–38.
58. Colin C, Lanoir D, Touzet S, et al. Sensitivity and specificity of third-generation hepatitis C virus antibody detection assays: an analysis of the literature. *J Viral Hepat*. 2001;8(2):87–95.
59. International Labour Office. *International Standard Classification of Occupations, Revised Edition 1968*. Geneva (Switzerland): International Labour Office; 1969.

Funding

This pooled analysis was supported by the Intramural Research Program of the National Cancer Institute/National Institutes of Health and National Cancer Institute/National Institutes of Health (R01 CA14690, U01 CA118444, and R01 CA92153-S1).

InterLymph annual meetings during 2010–2013 were supported by the Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute/National Institutes of Health (2010–2013); Lymphoma Coalition (2010–2013); National Institutes of Health Office of Rare Diseases Research (2010); National Cancer Institute/National Institutes of Health (R13 CA159842 01) (2011); University of Cagliari, Provincial Administration of Cagliari, Banca di Credito Sardo, and Consorzio Industriale Sardo, Italy (2011); Intramural Research Program of the National Cancer Institute/National Institutes of Health (2012); and Faculté de Médecine de Dijon, Institut de Veille Sanitaire, Registre des hémopathies malignes de Côte d'Or, INSERM, Institut National du Cancer, Université de Bourgogne, Groupe Ouest Est d'Etude des Leucémies et Autres Maladies du Sang (GOELAMS), l'Institut Bergonié, The Lymphoma Study Association (LYSA), Registre Régional des Hémopathies de Basse Normandie, and the City of Dijon, France (2013). Meeting space at the 2013 Annual Meeting of the American Association for Cancer Research (AACR) was provided by the Molecular Epidemiology Group (MEG) of the AACR. Pooling of the occupation data was supported by the National Cancer Institute/National Institutes of Health (R03CA125831).

Individual studies were supported by the Canadian Institutes for Health Research (CIHR), Canadian Cancer Society, and Michael Smith Foundation for Health Research (British Columbia); Intramural Research Program of the National Cancer Institute/National Institutes of Health (Iowa/Minnesota); National Cancer Institute/National Institutes of Health (N01-CP-ES-11027) (Kansas); National Cancer Institute/National Institutes of Health (R01 CA50850) and the City of Hope Comprehensive Cancer Center (P30 CA033572) (Los Angeles); National Cancer Institute/National Institutes of Health (R01 CA92153 and P50 CA97274), Lymphoma Research Foundation (164738), and the Henry J. Predolin Foundation (Mayo Clinic); Intramural Research Program of the National Cancer Institute/National Institutes of Health and Public Health Service (contracts N01-PC-65064, N01-PC-67008, N01-PC-67009, N01-PC-67010, and N02-PC-71105) (NCI-SEER); National Cancer Institute/National Institutes of Health (R01CA100555 and R03CA132153) and American Institute for Cancer Research (99B083) (Nebraska [newer]); National Cancer Institute/National Institutes of Health (N01-CP-95618) and State of Nebraska Department of Health (LB-506) (Nebraska [older]); National Cancer Institute/National Institutes of Health (R01CA45614, RO1CA154643-01A1, and R01CA104682) (UCSF1); National Cancer Institute/National Institutes of Health (CA143947, CA150037, R01CA087014, R01CA104682, RO1CA122663, and RO1CA154643-01A1) (UCSF2); National Heart Lung and Blood Institute/National Institutes of Health (hematology training grant award T32 HL007152), National Center for Research Resources/National Institutes of Health (UL 1 RR024160), and National Cancer Institute/National Institutes of Health (K23 CA102216 and P50 CA130805) (University of Rochester); National Cancer Institute/National Institutes of Health (CA62006 and CA165923) (Yale); Association pour la Recherche contre le Cancer, Fondation de France, AFSSET, and a donation from Faberge employees (Engela); European Commission (QLK4-CT-2000-00422 and FOOD-CT-2006–023103), Spanish Ministry of Health (CIBERESP, PI11/01810, RCESP C03/09, RTICESP C03/10, and RTIC RD06/0020/0095), Rio Hortega (CM13/00232), Agència de Gestió d'Ajuts Universitaris i de Recerca–Generalitat de Catalunya (Catalonian Government, 2009SGR1465), National Institutes of Health (contract NO1-CO-12400), Italian Ministry of Education, University and Research (PRIN 2007 prot.2007WEJLZB, PRIN 2009 prot. 20092ZELR2), Italian Association for Cancer Research (IG grant 11855/2011), Federal Office for Radiation Protection (StSch4261 and StSch4420), José Carreras Leukemia Foundation (DJCLS-R04/08), German Federal Ministry for Education and Research (BMBF-01-EO-1303), Health Research Board, Ireland and Cancer Research Ireland, and Czech Republic MH CZ - DRO (MMCI, 00209805) (EpiLymph); National Cancer Institute/National Institutes of Health (CA51086), European Community (Europe Against Cancer Programme), and Italian Alliance Against Cancer (Lega Italiana per la Lotta contro i Tumori) (Italy, multicenter); Italian Association for Cancer Research (IG 10068) (Italy, Aviano-Milan); Italian Association for Cancer Research (Italy, Aviano-Naples); Swedish Cancer Society (2009/659), Stockholm County Council (20110209), Strategic Research Program in Epidemiology at Karolinska Institut, Swedish Cancer Society (02 6661), Danish Cancer Research Foundation, Lundbeck Foundation (R19-A2364), Danish Cancer Society (DP 08-155), National Cancer Institute/National Institutes of Health (5R01 CA69669-02), and Plan Denmark (SCALE); Leukaemia & Lymphoma Research (United Kingdom); and Australian National Health and Medical Research Council (ID990920), Cancer Council NSW, and University of Sydney Faculty of Medicine (New South Wales).

Note

We thank the following individuals for their substantial contributions to this project: Aaron D. Norman, Dennis P. Robinson, and Priya Ramar (Mayo Clinic College of Medicine) for their work at the InterLymph Data Coordinating Center in organizing, collating, harmonizing, and documenting of the data from the participating studies in the InterLymph Consortium; Michael Spriggs, Peter Hui, and Bill Wheeler (Information Management Services, Inc) for their programming support; and Noelle Richa Siegfried and Emily Smith (RTI International) for project coordination.

Affiliations of authors: Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD (LMM, SMM, JSC, PH, NR, AB, KPC, QL, SIB, MSL, JNS); Department of Health Sciences Research (SLS, JRC, TMH), Division of Hematology (TGC), and Divison of General Internal Medicine (ML), College of Medicine, Mayo Clinic, Rochester, MN; Department of Cancer Etiology, City of Hope Beckman Research Institute, Duarte, CA (SSW, LB, AL); Prince of Wales Clinical School, University of New South Wales, Sydney, Australia (CMV); Department of Epidemiology, Comprehensive Cancer Center, University of Alabama, Birmingham, AL (CFS); Department of Epidemiology and Biostatistics, School of Medicine, University of California San Francisco, San Francisco, CA (PMB, EAH); Unit of Infections and Cancer (UNIC), Cancer Epidemiology Research Programme, Institut Català d' Oncologia, IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain, CIBER de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain (SdS, YB); Unit of Clinical Epidemiology, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden (KES); Department of Health Studies, University of Chicago, Chicago, IL (BCHC); Department of Environmental Health Sciences (YZ, TZ) and Department of Biostatistics (TRH), Yale School of Public Health, New Haven, CT; INSERM, Centre for Research in Epidemiology and Population Health (CESP), U1018, Environmental Epidemiology of Cancer Group, Villejuif, France, Univ Paris Sud, UMRS 1018, Villejuif, France (AM, JC, LO); Registry of Hematological Malignancies in Gironde, Bergonié Institute, 33076 Bordeaux, France (AM); Department of Histopathology, Douglass Hanly Moir Pathology, Macquarie Park, Australia, The Australian School of Advanced Medicine, Macquarie University, Sydney, Australia (JJT); Department of Medical Epidemiology and Biostatistics (H-OA) and Department of Oncology and Pathology (BG), Karolinska Institutet, Stockholm, Sweden; Department of Epidemiology, Harvard School of Public Health, Boston, MA (H-OA); Health Sciences Practice, Exponent, Inc., Menlo Park, CA, Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA (ETC); Department of Radiology, Oncology and Radiation

Science, Uppsala University, Uppsala, Sweden (BG); Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark (HH, MM); Epidemiology Unit, Fondazione IRCCS Istituto dei Tumori, Milan, Italy (PC); Department of Surgery and Translational Medicine, University of Florence, Florence, Italy (SdL); Unit of Occupational and Environmental Epidemiology, Cancer Prevention and Research Institute ISPO, Florence, Italy (LM, ASC); Biostatistics and Clinical Trials Unit—IRCCS IRST, Meldola, Italy (ON); Epidemiology Unit, Regina Elena National Cancer Institute, Rome, Italy (VR); Healthcare Development and Evaluation Unit, Agency for Health and Social Care, Bologna, Italy (SR); Unit of Epidemiology, Biostatistics and Clinical Trials, IRCCS AOU San Martino-IST, Genoa, Italy (ES); Cancer Registry and Histopathology Unit, "Civile - M.P. Arezzo" Hospital, ASP Ragusa, Ragusa, Italy (RT); Pathology Unit, Azienda Ospedaliera Universitaria Senese, Siena, Italy (CV); Cancer Epidemiology Unit, Imperial College London, London, UK (PV); Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany (NB); Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY (PBo); International Agency for Research on Cancer, Lyon, France (PBr, SF); Department of Public Health, Clinical and Molecular Medicine, Occupational Health Section, University of Cagliari, Cagliari, Italy (PC); Cancer Epidemiology & Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic (LF); Biological Hematology Unit, CRB Ferdinand Cabanne, Universitary Hospital of Dijon, Dijon, France, EA4184, University of Burgundy, Dijon, France (MM); Center for Chronic Immunodeficiency (CCI), University Medical Center Freiburg, Freiburg, Germany (AN); School of Nursing and Human Sciences, Dublin City University, Dublin, Ireland (ASt); Department of Preventive Medicine and Pathology and Norris Comprehensive Cancer Center, USC Keck School of Medicine, University of Southern California, Los Angeles, CA (WC); Fred Hutchinson Cancer Research Center, Seattle, WA, School of Public Health University of Washington, Seattle, WA (SD); Department of Environmental and Occupational Health, Drexel University School of Public Health, Philadelphia, PA (AJdR); Department of Family Medicine and Public Health Sciences, Wayne State University, Detroit, MI (RKS); Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN (ALF); Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York, York, UK (EVK, TL, ER, ASm); Genome Sciences Centre (AB-W), Centre for Lymphoid Cancer (JMC, RDG), and Cancer Control Research (JJS), BC Cancer Agency, Vancouver, BC, Canada; Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada (AB-W); Department of Pathology and Laboratory Medicine, (RDG) and School of Population and Public Health (JJS), University of British Columbia, Vancouver, BC, Canada; Sydney School of Public Health, University of Sydney, Sydney, Australia (BKA, AK); Unit of Epidemiology and Biostatistics, Centro di Riferimento Oncologico, IRCCS, Aviano, Italy (LDM, DS); Department of Epidemiology, IRCCS, Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy (CLV, EN); Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy (CLV); James P. Wilmot Cancer Center, University of Rochester, Rochester, NY (JWF, JLK); Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA (BMB); Cancer Prevention Institute of California, Fremont, CA (CAC); Winship Cancer Institute, Emory University, Atlanta, GA (CRF); Department of Cancer Biology, Mayo Clinic Cancer Center, Jacksonville, FL (JMF); Department of Dermatology, Boston University, Boston, MA, Roger Williams Medical Center, Providence, RI (MEK); Hadassah-Hebrew University, Jerusalem, Israel (OP); Department of Pathology, City of Hope National Medical Center, Duarte, CA (DDW).