Public Health Nutrition: 17(7), 1531–1537 doi:10.1017/S1368980013001249

Dietary patterns and the risk of non-Hodgkin lymphoma

Nicholas J Ollberding[1,2,*], Briseis Aschebrook-Kilfoy[2], Donne Bennett D Caces[3], Sonali M Smith[3], Dennis D Weisenburger[4] and Brian C-H Chiu[2,5]

[1]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA: [2]Department of Health Studies, University of Chicago, 5841 South Maryland Avenue, MC 2007, Chicago, IL 60637, USA: [3]Division of Hematology/Oncology, Department of Medicine, University of Chicago, Chicago, IL, USA: [4]Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA: [5]University of Chicago Comprehensive Cancer Center, Chicago, IL, USA

Submitted 17 October 2012: Final revision received 19 February 2013: Accepted 4 April 2013: First published online 9 May 2013

Abstract

Objective: Previous studies examining the role of single foods or nutrients in the aetiology of non-Hodgkin lymphoma (NHL) have produced inconsistent findings. Few studies have examined associations for dietary patterns, which may more accurately reflect patterns of consumption and the complexity of dietary intake. The objective of the present study was to examine whether dietary patterns identified by factor analysis were associated with NHL risk.

Design: Case–control.

Setting: Population-based sample residing in Nebraska from 1999 to 2002.

Subjects: A total of 336 cases and 460 controls.

Results: Factor analysis identified two major dietary patterns: (i) a 'Meat, Fat and Sweets' dietary pattern characterized by high intakes of French fries, red meat, processed meat, pizza, salty snacks, sweets and desserts, and sweetened beverages; and (ii) a 'Fruit, Vegetables and Starch' dietary pattern characterized by high intakes of vegetables, fruit, fish, and cereals and starches. In multivariable logistic regression models, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of overall NHL ($OR_{Q4\,v.\,Q1} = 3{\cdot}6$, 95% CI 1·9, 6·8; $P_{trend} = 0{\cdot}0004$), follicular lymphoma ($OR_{Q4\,v.\,Q1} = 3{\cdot}1$, 95% CI 1·2, 8·0; $P_{trend} = 0{\cdot}01$), diffuse large B-cell lymphoma ($OR_{Q4\,v.\,Q1} = 3{\cdot}2$, 95% CI 1·1, 9·0; $P_{trend} = 0{\cdot}09$) and marginal zone lymphoma ($OR_{Q4\,v.\,Q1} = 8{\cdot}2$, 95% CI 1·3, 51·2; $P_{trend} = 0{\cdot}05$). No association with overall or subtype-specific risk was detected for the 'Fruit, Vegetables and Starch' dietary pattern. No evidence of heterogeneity was detected across strata of age, sex, BMI, smoking status or alcohol consumption.

Conclusions: Our results suggest that a dietary pattern high in meats, fats and sweets may be associated with an increased risk of NHL.

Keywords
Dietary pattern
Non-Hodgkin lymphoma
Diet
Factor analysis
FFQ
Case–control study

The non-Hodgkin lymphomas (NHL) are a heterogeneous group of malignant neoplasms arising from the B, T and natural killer cells of the immune system. NHL is the fifth most commonly diagnosed cancer in the USA[1], with immune dysregulation thought to contribute to lymphoma development[2,3]. However, well-characterized immunodeficiency states only partially account for the rising rates of this disease in recent decades[3,4]. Dietary factors have been examined in relation to the development of NHL given their role in immune system regulation, oxidative stress and hormonal pathways regulating the proliferation of lymphoid tissue[5].

Previous studies examining the association between single food items and the risk of NHL have reported inconsistent findings for items such as red meat and processed meat, fruits and vegetables, and dairy products[5,6]. These inconsistencies may, in part, reflect the difficulty in disentangling the influence of individual food items that, when consumed in combination, may be highly correlated and exert synergistic or antagonistic effects on NHL risk. The examination of dietary patterns, which better reflect actual patterns of consumption and the complexity of dietary intake, has been used to address such limitations[7]. Factor analysis is an approach that can be applied to dietary data to identify underlying patterns of dietary intake based on the intercorrelations of individual foods and food groups.

In the present study, we used factor analysis to identify empirically derived dietary patterns in a population of adults residing in eastern Nebraska and to examine the associations of these dietary patterns with the risk of overall NHL. In secondary analyses, we examined the associations with risk for the common NHL subtypes.

*Corresponding author: Email Nicholas.Ollberding@cchmc.org

Materials and methods

Study population

Detailed information on the present case–control study has been reported previously[8,9]. Briefly, a rapid case ascertainment system was used to identify cases of NHL diagnosed in the sixty-six counties of eastern Nebraska between January 1999 and December 2002. All cases were reviewed and classified by an expert haematopathologist (D.D.W.) in accordance with the WHO classification of NHL[10]. Persons reporting any prior cancer (other than cutaneous squamous cell carcinoma or basal cell carcinoma), HIV infection, were deceased at initial contact or not mentally competent to participate were not included. Of the 529 eligible cases, 387 (73 %) participated in the study. The controls were frequency matched to the cases on gender and 5-year age group and were identified through random digit dialling of the same sixty-six counties in eastern Nebraska. Of the 697 eligible controls, 535 (77 %) participated in the study. Informed consent was obtained from all participants prior to the interview and the study protocol was approved by the Institutional Review Board of the University of Nebraska Medical Center.

Dietary assessment

A structured telephone interview, administered concurrently for cases and for controls, was used to collect information on demographic characteristics, lifestyle factors and environmental exposures. Following the administration of the telephone interview, a modified version of the Block 1995 Revision of the Health Habits and History Questionnaire (HHHQ) was mailed to all participants[11]. This quantitative FFQ included queries on 117 items and assessed the consumption of food items during the previous year. The HHHQ was developed using dietary data from the National Health and Nutrition Examination Survey II[12] and has been validated against dietary records with correlation coefficients in the range of 0·5–0·6 for most nutrients[13]. The FFQ was completed and returned by 348 (90 %) cases and 470 (88 %) controls. The participants who did not return the FFQ were slightly younger, but were similar with respect to BMI, smoking status and educational attainment to participants who returned the FFQ. In addition, excluded from the current analysis were twelve cases and ten controls who reported total energy intake of <3347 or >25 105 kJ/d (<800 or >6000 kcal/d) for men and <2510 or >20 920 kJ/d (<600 or >5000 kcal/d) for women or left more than 20 % of the items blank on the FFQ. In total, 336 cases and 460 controls provided complete dietary data and were included in the present analysis.

Prior to conducting the dietary pattern factor analysis, we aggregated the 117 food items queried by the FFQ into thirty-seven predefined food groups (Table 1). The food groups were classified according to similarities in food type, nutrient content and culinary usage. Food items representing distinct food types (e.g. pizza, French fries, beer, liquor) were retained as individual items for the factor analysis. A factor analysis deriving dietary patterns from all 117 food items was also performed to assess the influence of the classification on the dietary patterns identified.

Statistical analysis

Prior to performing the factor analysis, the factorability of the data was supported by examination of the correlation matrix for the thirty-seven foods and food groups, Bartlett's test of sphericity ($P < 0{\cdot}001$) rejecting the null hypothesis that the correlation matrix is from an identity matrix and the Kaiser–Meyer–Olkin measure of sampling adequacy (MSA = 0·82) indicating sufficient partial correlations among variables. Exploratory factor analysis with principal component factor extraction was used to identify latent factors (dietary patterns) explaining the greatest amount of variance in the correlation matrix. The factors were rotated by orthogonal varimax transformation to obtain independent factors with a simpler structure and greater interpretability. The number of factors chosen to retain was based on the following criteria: eigenvalue >1, visual examination of the scree plot and the interpretability of the factors[14]. Individual factor scores for each dietary pattern were calculated by weighted least squares regression. Foods and food groups with a factor loading ≥0·30 were used to label the factors.

Characteristics of the cases and of the controls were compared using χ^2 tests for categorical variables and t tests for continuous variables. Unconditional logistic regression was used to calculate odds ratios and 95 % confidence intervals for overall NHL. Polytomous logistic regression was used to calculate odds ratios and 95 % confidence intervals for six common subtypes of NHL according to the WHO classification: follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), small lymphocytic lymphoma/chronic lymphocytic leukaemia (SLL/CLL), marginal zone lymphoma (MZL), other miscellaneous B-cell lymphomas and T-cell lymphomas. Dietary pattern scores were examined as quartiles using cut-off points based on the exposure distribution among controls. The lowest exposure group (quartile 1) served as the referent in all models. Linear trends were tested by entering quartiles as ordinal variables in regression models. Models were also constructed to examine the associations between dietary pattern scores entered into regression as continuous variables and NHL risk. Potential non-linear relationships were first assessed by a likelihood ratio test comparing the model with only the linear term to a model containing linear and cubic spline terms. Covariates included in the final multivariable models were age (continuous), sex, education (<12 years, 12–15 years, 16+ years) and total energy intake (continuous). Alcohol consumption, smoking status, BMI, physical activity, farming status, use of hair dye and history of blood transfusion were examined as potential confounders, but were not

Table 1 Classification of food groups used in the dietary pattern factor analysis

Foods or food groups	Food items
Red meat	Beef, pork, lamb, hamburger
Poultry	Chicken, turkey
Processed meat	Bacon, hot dogs, sausage, processed lunch meat
Organ meat	Beef or poultry liver
Fish	Broiled fish, fried fish, canned tuna, shellfish
High-fat dairy products	Cream, cottage cheese, yoghurt, cheese, ice cream, whole milk
Low-fat dairy products	Low-fat cheese, low-fat cottage cheese, low-fat ice cream, low-fat yoghurt, 2% milk, skimmed milk
Cruciferous vegetables	Broccoli, cauliflower, Brussels sprouts, cabbage
Green leafy vegetables	Spinach, green salad, collards/kale
Legumes and soya	Alfalfa sprouts, lentils, peas, string beans, tofu, meat substitutes
Potatoes	Baked/mashed potatoes, sweet potatoes
Tomatoes or tomato juice	Tomatoes or tomato juice
Other vegetables	Rhubarb, beets, radishes, celery, carrots, corn, other vegetables
Vegetable soup	Vegetable soup
Other soups	Other soups
Fruits and fruit juice	Oranges/tangerines, orange juice, apple/grape juice, watermelon, grapefruit, strawberries, prunes, peaches/apricots, apples, bananas, other fruits
Nuts	Peanuts/peanut butter
Cereals and starches	Cornbread, biscuits/muffins, dry cereals, sweetened cereals, high-fibre cereals, bagels/English muffins/buns, dark bread, cooked cereal/grits, white bread, tortillas, pancakes/waffles
Pasta and rice	Pasta salad, macaroni, rice, spaghetti
Pizza	Pizza
French fries	French fries
Salty snacks	Salty snacks, nachos
Sweets and desserts	Cookies/cakes, pies, candy, chocolate, doughnuts/pastry
Eggs	Eggs, egg substitutes
Butter	Butter, whipped butter
Margarine	Margarine
Cooking fat and gravy	Crisco, lard, gravy
Olive oil or vegetable oil	Olive/canola oil, vegetable oil
Mayonnaise or creamy salad dressing	Mayonnaise/salad dressing
Condiments	Salsa/ketchup/taco sauce, sugar in coffee or tea, non-dairy creamer
Meal replacements	Breakfast or diet shakes, breakfast bars/power bars
Wine	Red wine, white wine
Beer	Beer
Liquor	Liquor
Sweetened beverages	Soft drinks, Kool-Aid or similar, Snapple or similar
Coffee or tea	Coffee, tea
Water	Water

Food items were measured in g/d using a modified version of the Block 1995 Revision of the Health Habits and History Questionnaire[11].

included in the final models as they were not found alone, or in combination, to change the risk estimates by more than 10%[15]. In addition, mutual adjustment for the dietary patterns was performed.

The associations for the dietary patterns with overall NHL risk were also examined in analyses stratified by sex, age (<61 years *v.* ≥61 years, median), BMI (<25 kg/m^2, 25–<30 kg/m^2, ≥30 kg/m^2), smoking status (never smoker, former smoker, current smoker), alcohol consumption (<0·3% of energy/d *v.* ≥0·3% of energy/d, median) and NHL subtype. Heterogeneity in the risk estimates by sex, age, BMI, smoking status and alcohol consumption was assessed using a Wald test of the cross-product terms. Heterogeneity in the risk estimates across the six NHL subtypes was assessed using a Wald test of the parameter estimates obtained from unconditional polytomous logistic regression comparing case subgroups. All tests were two-sided with α of $P < 0{\cdot}05$ considered statistically significant. Data analyses were performed using the statistical software package SAS version 9·2.

Results

Selected characteristics of the cases and of the controls are given in Table 2. Cases and controls were similar with respect to age, sex, race, educational attainment, first-degree family history of cancer, smoking status, alcohol consumption, BMI and energy intake. The majority of NHL cases were of B-cell origin (>94%) and classified as SLL/CLL (7·4%), FL (31·3%), DLBCL (26·5%), MZL (8·9%) or other B-cell lymphomas (20·2%), with only a small number of T-cell lymphomas (5·7%).

The principal component factor extraction identified two dietary patterns that were labelled descriptively as a 'Meat, Fat and Sweets' dietary pattern and a 'Fruit, Vegetables and Starch' dietary pattern based on factor score loadings (Table 3). The 'Meat, Fat and Sweets' dietary pattern loaded high on French fries, red meat, processed meat, pizza, salty snacks, sweets and desserts, sweetened beverages, condiments, cooking fat and gravy, margarine, high-fat dairy foods, pasta and rice, and eggs. The 'Fruit,

Table 2 Characteristics of eligible non-Hodgkin lymphoma (NHL) cases and controls, Nebraska, USA, 1999–2002

	Cases (*n* 336)		Controls (*n* 460)		
	Mean or *n*	SD or %	Mean or *n*	SD or %	*P* value*
Age (years)	58·6	12·7	58·0	12·8	0·54
Male†	185	55·1	236	51·3	0·29
White race†	325	96·7	445	96·7	0·99
Education†					
<12 years	15	4·5	9	2·0	
12–15 years	126	37·8	191	41·5	
16+ years	192	57·7	260	56·5	0·09
First-degree family history of cancer†					
None	159	47·3	224	48·7	
Non-haematopoietic	140	41·7	196	42·6	
Haematopoietic	37	11·0	40	8·7	0·55
Smoking status†					
Never	169	52·2	221	50·3	
Former	108	33·3	140	31·9	
Current	47	14·5	78	17·8	0·48
BMI status†					
Normal/underweight (<25 kg/m^2)	103	30·9	160	34·8	
Overweight (25 to <30 kg/m^2)	131	39·3	188	40·9	
Obese (≥30 kg/m^2)	99	29·7	112	24·3	0·21
Percentage of energy from alcohol	2·2	5·2	2·6	5·7	0·31
Energy (kJ/d)	8303	3310	7996	2955	0·18
Energy (kcal/d)	1984·5	791·0	1911·1	706·3	0·18
WHO-defined NHL subtypes†					
SLL/CLL	25	7·4			
FL	105	31·3			
DLBCL	89	26·5			
MZL	30	8·9			
Other B-cell lymphomas	68	20·2			
T-cell lymphomas	19	5·7			

SLL/CLL, small lymphocytic lymphoma/chronic lymphocytic leukaemia; FL, follicular lymphoma; DLBCL, diffuse large B-cell lymphoma; MZL, marginal zone lymphoma.
Numbers may not sum to the total due to missing data.
*P value for the χ^2 test for categorical variables and the t test for continuous variables when comparing the proportions and the means respectively between cases and controls.
†Values presented are *n* and %.

Vegetables and Starch' dietary pattern loaded high on other vegetables, tomatoes and juice, cruciferous vegetables, green leafy vegetables, fruit and fruit juice, legumes and soya, vegetable soup, fish, mayonnaise and creamy salad dressing, cereals and starches, potatoes, other soups, and low-fat dairy foods. Similar factors were identified when deriving dietary patterns from all 117 food items (data not shown).

The odds ratios and 95% confidence intervals for NHL according to quartiles of the factor scores for the two dietary patterns are provided in Table 4. In models adjusting for age, sex, educational attainment and total energy, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of overall NHL ($OR_{Q4\,v.\,Q1} = 3{\cdot}6$, 95% CI 1·9, 6·8; $P_{trend} = 0{\cdot}0004$). The results were similar when the dietary pattern was entered into regression as a continuous variable (the log of the unit change in the odds ratio per unit change in the factor score (β) = 0·54, SE = 0·12; $P < 0{\cdot}0001$). Red meat ($\beta = 0{\cdot}003$, SE = 0·001; $P = 0{\cdot}01$) and pizza ($\beta = 0{\cdot}02$, SE = 0·006; $P = 0{\cdot}004$) were the only individual food groups loading high on the 'Meat, Fat and Sweets' dietary pattern that demonstrated a linear association with the risk of overall NHL. No association with overall NHL risk was detected for the 'Fruit, Vegetables and Starch' dietary pattern when examined as quartiles of exposure ($OR_{Q4\,v.\,Q1} = 0{\cdot}9$, 95% CI 0·6, 1·4; $P_{trend} = 0{\cdot}75$) or as a continuous variable ($\beta = -0{\cdot}05$, SE = 0·09; $P < 0{\cdot}56$). In subtype-specific analyses, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of FL ($OR_{Q4\,v.\,Q1} = 3{\cdot}1$, 95% CI 1·2, 8·0; $P_{trend} = 0{\cdot}01$), DLBCL ($OR_{Q4\,v.\,Q1} = 3{\cdot}2$, 1·1, 9·0; $P_{trend} = 0{\cdot}09$) and MZL ($OR_{Q4\,v.\,Q1} = 8{\cdot}2$, 95% CI 1·3, 51·2; $P_{trend} = 0{\cdot}05$). No associations were detected for the 'Meat, Fat and Sweets' dietary pattern and risk of SLL/CLL, other miscellaneous B-cell lymphomas or T-cell lymphomas (online supplementary material, Supplemental Table 1); however, point estimates for all subtypes were in the direction of an increased risk and no evidence of heterogeneity was detected ($P = 0{\cdot}95$) across the NHL subtypes examined. No associations were detected for the 'Fruit, Vegetables and Starch' dietary pattern and the risk of any of the subtypes examined, nor was there evidence of heterogeneity in the risk estimates ($P = 0{\cdot}72$) across subtypes. Mutual adjustment for the dietary patterns had no effect on the overall or subtype-specific risk estimates (data not shown).

Table 3 Factor loading matrix for the major dietary patterns identified using principal component factor analysis

Food or food group	Factor 1 'Meat, Fat and Sweets' dietary pattern	Factor 2 'Fruit, Vegetables and Starch' dietary pattern
French fries	0·71	–
Red meat	0·70	–
Processed meat	0·65	–
Pizza	0·57	–
Salty snacks	0·57	–
Sweets and desserts	0·55	–
Sweetened beverages	0·48	–
Condiments	0·47	–
Cooking fat and gravy	0·47	–
Margarine	0·45	–
High-fat dairy products	0·39	–
Pasta and rice	0·38	–
Eggs	0·37	–
Other vegetables	–	0·68
Tomatoes and juice	–	0·62
Cruciferous vegetables	–	0·62
Green leafy vegetables	–	0·61
Fruit and fruit juice	–	0·57
Legumes and soya	–	0·56
Vegetable soup	–	0·50
Fish	–	0·47
Mayonnaise or creamy salad dressing	–	0·45
Cereals and starches	–	0·43
Potatoes	–	0·39
Other soups	–	0·33
Low-fat dairy products	–	0·31

The factor score provides an estimate of the degree to which an individual's diet adheres to the dietary pattern, with higher scores denoting greater adherence. Foods or food groups with factor loadings <0·30 on both factors are not presented.

Table 4 Odds ratios and 95 % confidence intervals for non-Hodgkin lymphoma (NHL) according to quartiles of the dietary pattern scores, Nebraska, USA, 1999–2002

		Overall NHL			Follicular lymphoma			Diffuse large B-cell lymphoma			
	Controls	Cases	OR*	95 % CI	Cases	OR*	95 % CI	Cases	OR*	95 % CI	$P_{heterogeneity}$†
'Meat, Fat and Sweets' dietary pattern											
Quartile 1	115	40	1·0	–	15	1·0	–	10	1·0	–	
Quartile 2	115	102	2·9	1·8, 4·6	25	1·9	0·9, 3·8	29	3·1	1·4, 6·8	
Quartile 3	115	92	2·7	1·7, 4·5	34	2·8	1·4, 5·9	21	2·2	0·9, 5·3	
Quartile 4	115	99	3·6	1·9, 6·8	30	3·1	1·2, 8·0	28	3·2	1·1, 9·0	
P_{trend}‡			<0·01			0·01			0·09		0·95
'Fruit, Vegetables and Starch' dietary pattern											
Quartile 1	115	76	1·0	–	27	1·0	–	22	1·0	–	
Quartile 2	115	84	1·1	0·7, 1·6	23	0·8	0·4, 1·5	25	1·1	0·6, 2·0	
Quartile 3	115	97	1·2	0·8, 1·8	31	1·0	0·6, 1·9	19	0·7	0·4, 1·5	
Quartile 4	115	76	0·9	0·6, 1·4	23	0·7	0·4, 1·5	22	0·7	0·4, 1·6	
P_{trend}‡			0·75			0·57			0·30		0·72

*Odds ratios estimated from unconditional logistic regression and adjusted for age (continuous), sex, education (<12 years, 12–15 years, 16+ years) and total energy intake (continuous).
†*P* value for the test of heterogeneity in the parameter estimates across the six NHL subtypes.
‡*P* value for the Wald χ^2 test of H_o: $\beta = 0$ when modelling quartiles as an ordinal variable.

The associations for the dietary patterns with overall NHL risk were similar in models stratified by sex, age, BMI, smoking status and alcohol consumption and no heterogeneity in the risk estimates was detected for the cross-product terms (data not shown).

Discussion

In the present population-based case–control study conducted in Nebraska, we identified a 'Meat Fat, and Sweets' dietary pattern and a 'Fruit, Vegetables and Starch' dietary

pattern using principal component factor extraction. In multivariable models, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of overall NHL. In secondary analyses stratified by NHL subtype, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of FL, DLBCL and MZL. No association with overall or subtype-specific NHL risk was detected for the 'Fruit, Vegetables and Starch' dietary pattern.

To the best of our knowledge, only one prior study has examined dietary patterns in relation to the risk of NHL. In a multiethnic cohort of older adults, Erber *et al.*[16] reported no association with NHL risk for a 'Meat and Fat' dietary pattern characterized by high intakes of discretionary fat, meat and organ meat, processed meat, white potatoes, non-whole grains, eggs and cheese. However, the association for this 'Meat and Fat' dietary pattern was in the direction of an increased risk ($OR_{T3\ v.\ T1} = 1{\cdot}40$, 95 % CI 0·82, 2·41) among Caucasian males. In our study, the association between the 'Meat, Fat and Sweets' dietary pattern and overall NHL risk was similar for men and women. Discrepancies in the findings between these studies may reflect differences in the individual food items aggregated and entered into the factor analysis based, in part, on differences in the location, age distribution and ethnic composition of the study participants. In addition, the 'Meat, Fat and Sweets' dietary pattern identified in our sample also loaded high on sweets and desserts, sweetened beverages, and pasta and rice (rich sources of refined carbohydrates and other nutrients), as well as high-fat dairy products which may have influenced the results. Differences in the study designs, including the potential for recall bias in our case–control study, may have also contributed to the discrepant findings. In addition, to the extent to which unmeasured changes in dietary intake occurred over the follow-up period, misclassification error may have attenuated the associations reported for the Multiethnic Cohort (MEC) study. In line with our results, no associations with overall NHL risk were detected in the MEC study for a dietary pattern characterized by high intakes of fruits and vegetables or for a dietary pattern characterized by high intakes of milk and yoghurt and fruits[16].

In the secondary analyses examining the risk of specific NHL subtypes, the 'Meat, Fat and Sweets' dietary pattern was associated with an increased risk of FL, DLBCL and MZL, with the point estimates for all subtypes in the direction of an increased risk. In the MEC study, the risk of FL was increased fivefold among males, but not females, with the highest factor scores for the 'Meat and Fat' dietary pattern[16]. In our sample, we detected no heterogeneity in the association between the 'Meat, Fat and Sweets' dietary pattern and the risk of FL by sex. In agreement with our results, no consistent association with the risk of FL, DLBCL or SLL/CLL was detected for a dietary pattern characterized by high intakes of fruits and vegetables, or for a dietary pattern characterized by high intakes of milk and yoghurt and fruits, in the MEC study[16].

Dietary pattern analysis addresses several limitations of traditional approaches examining associations of single foods or nutrients in relation to cancer risk. Factor analysis provides an empirically derived measure that better reflects the combination of foods consumed *ad libitum* in the diet of free-living people, accounts for potential interactions between foods and nutrients, aggregates additive effects of single items included in the composite factor, provides a measure more amenable to dietary recommendations, and accounts for intercorrelations between foods and/or nutrients[7]. However, dietary pattern analysis is not without its limitations. Factor analysis requires subjective decisions to be made at several points when deriving dietary patterns[7], may provide little insight into the biological processes underlying the associations with disease risk[17], and empirically derived dietary patterns may not be reproducible across populations.

The current study has several strengths including the confirmation of NHL diagnoses by an expert haematopathologist; the high response rates for cases (73 %) and controls (77 %); the recruitment of randomly selected, population-based controls sampled from the same source population giving rise to the cases; the use of a rapid case ascertainment system to minimize the potential for survival bias; and the use of a validated FFQ. There were also limitations. First, as previously discussed, our factor analysis attempting to identify empirically derived, *a posteriori* dietary patterns required subjective decisions regarding the classification of food groups and the retention of factors. To address this limitation, we adhered closely to previous studies when aggregating food items into predefined groups[7] and conducted a second factor analysis on all 117 food items to assess the influence of the classification on the dietary patterns identified. In addition, we closely followed established guidelines with respect to the retention of factors[14] and identified a 'Meat, Fat and Sweets' dietary pattern and a 'Fruit, Vegetables and Starch' dietary pattern similar to those reported in previous studies[7], providing support as to their reproducibility in other populations. Second, the dietary patterns retained in our analysis explained only a small portion of the total variance (8·6 %). Third, the small number cases limited our ability to detect associations in subtype-specific analyses, as well as our ability to detect heterogeneity in the risk estimates across the NHL subtypes examined. Fourth, we cannot rule out the possibility of residual confounding and recall bias.

Conclusion

Our findings suggest that greater adherence to a 'Meat, Fat and Sweets' dietary pattern may increase the risk of NHL. Additional studies, ideally allowing for prediagnostic exposure assessment, are required to confirm these findings.

In addition, future studies should consider whether associations between dietary patterns and the risk of NHL are consistent across racial or ethnic groups. The associations detected for specific NHL subtypes in our population also highlight the need for pooled analyses of NHL subtypes to address aetiological heterogeneity.

Acknowledgements

Sources of funding: This research was supported by research grant 99B083 from the American Institute for Cancer Research and, in part, by grants CA94770 and CA100555 from the National Cancer Institute. *Conflicts of interest:* The authors have no conflicts of interest to declare. *Authors' contributions:* N.J.O. designed and carried out the statistical analysis, interpreted the results and wrote the manuscript; B.A.-K. provided insight into the analytic strategy and interpretation of results; D.B.C.C. provided insight into the analytic strategy and interpretation of results; S.M.S. provided insight into the analytic strategy and interpretation of results; D.D.W. conducted the pathology review, provided insight into the analytic strategy and interpretation of results; B.C.-H.C. collected the data, originally conceived of the analysis, provided insight into the analytic strategy and interpreted the results. All authors provided feedback on the initial draft of the manuscript and approved the final version. *Acknowledgements:* The authors thank Mr Martin Bast of the Nebraska Lymphoma Study Group for coordinating the case identification and for obtaining the physician consents for the study cases.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1368980013001249

References

1. Siegel R, Ward E, Brawley O *et al.* (2011) Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* **61**, 212–236.
2. Hennessy BT, Hanrahan EO & Daly PA (2004) Non-Hodgkin lymphoma: an update. *Lancet Oncol* **5**, 341–353.
3. Ekstrom-Smedby K (2006) Epidemiology and etiology of non-Hodgkin lymphoma – a review. *Acta Oncol* **45**, 258–271.
4. Alexander DD, Mink PJ, Adami HO *et al.* (2007) The non-Hodgkin lymphomas: a review of the epidemiologic literature. *Int J Cancer* **120**, Suppl. 12, 1–39.
5. Skibola CF (2007) Obesity, diet and risk of non-Hodgkin lymphoma. *Cancer Epidemiol Biomarkers Prev* **16**, 392–395.
6. Cross AJ & Lim U (2006) The role of dietary factors in the epidemiology of non-Hodgkin's lymphoma. *Leuk Lymphoma* **47**, 2477–2487.
7. Newby PK & Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177–203.
8. Chiu BC, Kolar C, Gapstur SM *et al.* (2005) Association of NAT and GST polymorphisms with non-Hodgkin's lymphoma: a population-based case–control study. *Br J Haematol* **128**, 610–615.
9. Chiu BC, Soni L, Gapstur SM *et al.* (2007) Obesity and risk of non-Hodgkin lymphoma (United States). *Cancer Causes Control* **18**, 677–685.
10. Jaffe ES, Harris NL, Stein H *et al.* (editors) (2001) *World Health Organization Classification of Tumors. Pathology and Genetics of Tumors of Hematopoietic and Lymphoid Tissues.* Lyon: IARC Press.
11. Block G, Coyle LM, Hartman AM *et al.* (1994) Revision of dietary analysis software for the Health Habits and History Questionnaire. *Am J Epidemiol* **139**, 1190–1196.
12. Block G, Hartman AM, Dresser CM *et al.* (1986) A data-based approach to diet questionnaire design and testing. *Am J Epidemiol* **124**, 453–469.
13. Block G, Woods M, Potosky A *et al.* (1990) Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol* **43**, 1327–1335.
14. Kim J-O & Mueller CW (1978) *Introduction to Factor Analysis.* Newbury Park, CA: SAGE Publications.
15. Mickey RM & Greenland S (1989) The impact of confounder selection criteria on effect estimation. *Am J Epidemiol* **129**, 125–137.
16. Erber E, Maskarinec G, Gill JK *et al.* (2009) Dietary patterns and the risk of non-Hodgkin lymphoma: the Multiethnic Cohort. *Leuk Lymphoma* **50**, 1269–1275.
17. Hu FB, Rimm EB, Stampfer MJ *et al.* (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr* **72**, 912–921.