Supporting Statement A for:

A multi-center international hospital-based case-control study of lymphoma in Asia (AsiaLymph) (NCI)

OMB No. 0925-0654, Expiration Date: 10/31/2015

Yellow highlights indicate changes to submission from prior approval.

October 14, 2015

Submitted by:

Occupational and Environmental Epidemiology Branch
Epidemiology and Biostatistics Program
Division of Cancer Epidemiology and Genetics (DCEG)
National Cancer Institute
National Institutes of Health
Rockville, Maryland 20852

Contact Person:

Nathaniel Rothman, M.D., M.P.H., M.H.S.
NCI Project Officer
9609 Medical Center Drive MSC 9776 Room 6E134
Bethesda, MD 20892

Phone: (240)-276-7169

e-mail: rothmann@mail.nih.gov

TABLE OF CONTENTS

LIST OF ATTACHMENTS

This is a request for a revision of the "AsiaLymph Study" for an additional 3 years. The project focuses on collecting critically needed information to understand and reduce the cancer burden from lymphoid malignancies in the Asian population, the incidence of which has risen in recent decades. Specifically, environmental exposures to industrial emissions, genetic susceptibility, viral exposures, early life exposures, ultraviolet (UV) radiation exposures, and other risk factors for lymphoma overall and specifically for populations in Asia will continue to be examined.

A. JUSTIFICATION

A.1 Circumstances Making the Collection of Information Necessary

Under Section 411 of the Public Health Service Act (42 USC *§* 285a), the Division of Cancer Epidemiology and Genetics (DCEG) of the National Cancer Institute (NCI) is authorized to collect information to generate and test hypotheses concerning environmental and host determinants of cancer. The mission of the DCEG is to conduct "multidisciplinary research to discover the genetic and environmental determinants of cancer and new approaches to cancer prevention. The Division's research impacts public health policy in the United States and around the world." More specifically, the mission of the Occupational and Environmental Epidemiology Branch is to "conduct studies in the United States and abroad to identify and evaluate environmental and workplace exposures that may be associated with cancer risk." The AsiaLymph study continues to test hypotheses regarding the association of specific environmental, occupational, viral, and other exposures and lymphoid malignancies.

Incidence rates of certain lymphomas have increased in the United States and in many other parts of the world. The contribution of environmental, occupational, and genetic factors to the cause of lymphoma has generated a series of novel findings from epidemiological studies conducted in the United States that have attempted to explain this increase. However, none of the

chemical associations have been conclusively established and the identification of the key, functional alleles in gene regions associated with risk of NHL requires further elucidation. Further, the ability to follow-up, confirm, and extend these observations in the United States is limited by the low prevalence and limited range of several important chemical and viral exposures and the high to complete linkage disequilibrium among key candidate genetic loci in Western populations. To optimize the ability to build on and clarify these findings, it is necessary to investigate populations that differ from those in the West in both exposure patterns and underlying genetic structure. A multidisciplinary case-control study of lymphoma in Asia, where lymphoma rates have also risen, provides an opportunity to replicate and extend recent and novel observations made in studies in the West in a population that is distinctly different with regard to patterns of key risk factors, including range of exposures, prevalence of exposures, correlations between exposures, and variation in gene regions of particular interest. It also improves the ability to understand the causes of certain types of rare lymphoma tumors in the United States that occur at much higher rates in Asia. As such, the overall goals of the AsiaLymph study are to confirm and extend previous findings and yield novel insights into the causes of lymphoma in both Asia and in the United States.

Research suggests that organochlorines (OCs), trichloroethylene (TCE), and benzene may be associated with risk of lymphoma and that benzene and TCE have immunotoxic properties (Stewart 2009; Lan et al. 2004; Lan 2009), that genetic variation in certain loci involved in immunologic regulation (e.g., *TNF/LTA*, *IL10*, and *IL4*) may contribute to risk of lymphoma, and that interactions between these chemicals and genes may exist (Lan et al. 2004; Colt et al. 2009; Wang et al. 2007). However, none of these chemical or genetic associations have been conclusively established, and the underlying biologic plausibility, including

identification of critical functional alleles in genetic studies, requires further elucidation. At the same time, there is a growing appreciation of the critical need for high quality pathology review in etiologic studies of lymphoma, as evidence is increasing that some risk factors may be highly specific to one or more subtypes of lymphoma and therefore the subtypes need to be very accurately characterized (Morton et al. 2008).

Lymphoma in Asia

Although NHL rates historically have been lower in Asia than in the West, there is evidence that rates have been rising in recent decades in Asia from some of the best cancer registries in this region located in Shanghai and Singapore (Jin et al. 1999; Chia et al. 2001). For example, in Shanghai between 1972-3 and 1993-4, NHL rates rose 33% in males and 66% in females, while there was a small drop in incidence rates for leukemia in both sexes. Overall, there was an 11% and 13% decline in the incidence rates of all cancers for males and females, respectively, during this time period (Jin et al. 1999). The distribution of NHL histologic subtypes also differs in Asians and Caucasians. Although diffuse large B-cell lymphoma (DLBCL) is the most common histologic subtype in both Asians and Caucasians, rates of follicular lymphoma are substantially lower in Asians, whereas rates of T-cell lymphomas, particularly nasal type Natural Killer T (NK/T) cell lymphomas, are substantially higher in Asians (Au and Lo 2005; Ng et al. 1986; Gross et al. 2008; Kadin et al. 1983). As a consequence, this study provides a unique opportunity to replicate and extend key findings observed in Caucasians for histologies with characteristics shared by both populations as well as to rigorously study the epidemiology of those tumors that appear to be more common in Asia than in Western populations. See the protocol for a complete discussion of the background and rationale to study environmental exposures to industrial emissions, genetic susceptibility, viral

exposures, early life exposures, ultraviolet (UV) radiation exposures, and other risk factors for lymphoma overall and specifically for populations in Asia (**Attachment 1)**.

A multidisciplinary case-control study of lymphoma in Asia is timely because it provides an opportunity to replicate and extend recent and novel observations made in studies among Caucasians in a population that is distinctly different with regard to patterns of key risk factors. Refer to **Attachment 1** (p. 3-10) for a complete discussion of background, rationale and references. This study involves some of the leading lymphoma clinicians and pathologists in Asia and represents the culmination of a seven year effort to launch (**Attachment 2**). This study represents an opportunity to learn about the etiology of this important tumor in one of the largest populations in the world and to confirm many leads identified from previous studies conducted in the United States and elsewhere in the West.

A.2 Purpose and Use of the Information

The purpose of the study is to evaluate the etiology of lymphoma in Asia. The main focus of the study is on chemical exposures, viral exposures, and genetic susceptibility, with central pathology review to characterize effects by histologic subtype. Specific study goals are as follows:

1) Investigate the role of environmental exposure to organochlorines and occupational exposure to trichloroethylene, benzene, and other chemical solvents as well as to other potential occupational exposures;

2) Investigate the role of family history, high-prior candidate genetic variants (e.g., *TNF/LTA* locus) and emerging findings from genome-wide association studies of lymphoma in Caucasians, and use state-of-the-art genomics to study genetic variants that may be unique to risk of lymphoma in Asia;

3) Investigate the etiologic role of Epstein Bar (EBV), Hepatitis B (HBV) and Hepatitis C (HCV) viruses; evaluate potential novel viral agents in T-cell lymphoma; carry out

studies to understand pathogenetic mechanisms of Natural Killer T (NK/T) cell lymphoma;

4) Study other potential determinants of lymphoma including early childhood exposures including crowding and animal exposures, medical conditions, ultraviolet radiation exposure, and other lifestyle factors;

5) Determine the influence of risk factors for lymphoma overall and by histologic subtype determined by central pathology review.

The study methods include an initial contact script followed by a screener and core questionnaire that are administered by trained interviewers. There are three Chinese-version questionnaires that are used at the different study centers where enrollment is ongoing: Hong Kong, Chengdu, and Tianjin, where the study is occurring. The differences in the three Chinese questionnaires include slight differences in the phrasing of a few questions, reflecting local and regional language usage. For updated versions of the screener and core questionnaire in Chinese that will be used in Hong Kong, Chengdu, and Tianjin, and see **Attachments 3B**, **3D**, and **3F**. English versions of the updated screener and core questionnaires **(Attachments 3A, 3C, and 3E)** are also included. There is a small change to the screener in which potential study subjects are asked if they have had a prior history of myeloid neoplasms or acute leukemia (in addition to being asked if they have had a prior history of lymphoma) and no change to the questionnaire, so content and information being collected is the same and the amount of time required to administer the questionnaire is the same as in the previous request to OMB.

In addition, depending upon the occupational history obtained, the computer-assisted personal interview (CAPI) brings up one or more occupational assessment modules which consist of questions that are targeted to capture additional information regarding chemical exposures (**Attachments 4 and 5**). A blood and buccal cell sample are collected from each subject to measure chemical and viral exposures and genetic factors. Samples of tumor tissue are

collected from pathology blocks from cases after clinical diagnostic tests are completed and used for confirmatory tests, viral studies, and molecular diagnostic tests to be conducted by the study pathology center.

The study is the largest molecular epidemiology study of lymphoma ever carried out anywhere in the world and the results and information from this study will be used in a number of different ways including to:

- Contribute substantial scientific contributions to the literature concerning the etiology of lymphoma in Asia and world-wide focusing on the risk factors being studied;
- Confirm previous occupational, environmental, viral, genetic and other hypotheses identified in epidemiologic studies conducted among in the West;
- Assist senior leadership of DCEG to determine how best to allocate resources to study lymphoma in the future and;
- Contribute to a better understanding of the etiology of lymphoma world-wide and thus may eventually help to understand the reason for the increase in the incidence of this cancer in Asia and elsewhere and identify approaches that can be used for its prevention.

The study has been conducted successfully in the four study centers (Hong Kong, Chengdu and Tianjin, China, and Taiwan). We have enrolled 2,962 lymphoma cases and their 2,962 individually-matched controls. We will need approximately 3-4 months to complete enrollment of the previously approved 3,300 cases and 3,300 individually-matched controls. The information collected to date has been used to monitor study quality. No statistical analyses will be conducted until the full study, including the extension below, is completed.

In addition, we have requested and received NCI scientific and IRB approval (**Attachments 6A, 6B**) to increase the study sample size to an additional 900 lymphoma cases and 900 individually-matched controls, and to extend the study to include enrollment of 2,000 myeloid neoplasm (leukemia) cases, all using the same study procedures and instruments. We are therefore requesting OMB approval to increase the sample size of the study by 2,900 cases and

900 controls, which we plan to enroll over the next three years in three study centers (Hong Kong, Chengdu, and Tianjin, China). We have completed enrollment in the Taiwan study center.

Justification for increasing the sample size to a total of 4,200 lymphoma cases and 4,200 matched controls: There is recent recognition of the critical importance of studying lymphoma subtype-specific effects for both environmental and genetic risk factors, which often do not overlap across subtypes. In order to have adequate statistical power to study risk factors for the less common lymphoma subtypes, we need to increase the study sample size. In addition, we also need to increase the study sample size to compensate for the loss of case-control pairs because of study subjects who declined to provide a blood sample.

Justification for extending the study to include 2,000 myeloid neoplasm (leukemia) cases: In the ongoing AsiaLymph study, we have enrolled cases with both solid lymphoma tumors and blood-borne lymphoma tumors (including acute lymphocytic leukemia and chronic lymphocytic leukemia, which derive from lymphoid stem cells). We have been approved to extend the study to include enrollment of cases with myeloid leukemia (including acute and chronic myeloid leukemia, which derive from myeloid stem cells) because there is overlap in risk factors identified to date for lymphoid and myeloid leukemias and because the cause of most cases of both classes of leukemia have not been identified. AsiaLymph provides an outstanding and efficient opportunity to compare and contrast risk factors for the lymphoid and myeloid leukemias as well as to make new observations about the etiology of myeloid leukemias.

Only a small percentage of the etiology of myeloid leukemia and related disorders has been explained. Although there are a few proven associations (e.g., ionizing radiation, benzene, certain drugs, and weak effects of tobacco use), the definitive identification of other suspected exposures as well as comprehensive identification of genetic risk factors has been challenging

and has not taken place to date. There is now an excellent opportunity for a comprehensive study of myeloid leukemias because there have been major advances in understanding myeloid neoplasms at the molecular level, and because the WHO classification of these neoplasms has evolved to the point that tumors can be classified into standard subtypes that can be studied with regard to important occupational, environmental, lifestyle, and genetic risk factors. The ongoing AsiaLymph study provides a special opportunity and efficient platform to study myeloid neoplasms including acute myeloid leukemia and chronic myeloid leukemia. Key exposures to be evaluated are occupational exposure to solvents and formaldehyde and environmental exposure to organochlorines. In addition, we will conduct a comprehensive assessment of genetic susceptibility. The same methods, questionnaire, biological sample protocol, personnel and hospitals used in the AsiaLymph study in the three participating study centers will be used for enrolling myeloid neoplasm cases. Further, controls already being enrolled in the AsiaLymph study will be frequency-matched to the myeloid neoplasm cases, so that we will not need to enroll additional controls for these cases.

A.3 Use of Improved Information Technology and Burden Reduction

AsiaLymph employs cutting edge methodologies to efficiently collect detailed information about occupational exposures using a new software tool (OccIDEAS) that is integrated into the computer assisted personal interview (CAPI), which has never been used in a previous study of lymphoma. Further, each interviewer uses a touch-tablet (Lenovo X 220) to administer the questionnaire.

After consultation with the NCI Privacy Act Coordinator it was determined that a Privacy Impact Assessment (PIA) is needed. Based on the coordinator's recommendation, a PIA has been undertaken and is currently in at the NCI level.

A.4 Efforts to Identify Duplication and Use of Similar Information

There are no other sources of similar information that have been collected that yield the same results as this effort in Asia. First, almost every epidemiologic study of lymphoma has been conducted among primarily Caucasians in Western countries and no large scale multi-center study of lymphoma has ever been conducted in Asia. Secondly, no large-scale epidemiologic study of lymphoma in any population has ever conducted central immunophenotyping and pathology review, which is critical to understand risk factors for specific subtypes of lymphoma. Third, this is the largest multidisciplinary study of lymphoma ever conducted anywhere in the world. As such, it has better statistical power than any previous study to evaluate risk factors for histological subtypes, all collected and characterized using the same uniform methods.

A.5 Impact on Small Businesses or Other Small Entities

There is no impact on small businesses or other small entities.

A.6 Consequences of Collecting the Information Less Frequently

Patients are asked to participate once for the interview and biological specimen collection.

A.7 Special Circumstances Relating to the Guidelines of 5 CFR 1320.5

The study fully complies with the Guidelines.

A.8 Comments in Response to the Federal Register Notice and Efforts to Consult Outside the Agency

A 60-day Federal Register was published on August 28, 2015 Vol.80, P. 52325 soliciting additional comments from the public. One comment was received. Many people were consulted on all aspects of this study including previously existing information, identification of study centers and hospitals, study design, and study methods over the last three years. This includes some of the world's leading lymphoma pathologists, clinicians, and epidemiologists, and occupational and environmental exposure assessment experts. These individuals are located in the United States, Europe, Australia, and Asia (**Attachment 2**). All issues were successfully resolved through these consultations.

A.9 Explanation of Any Payment or Gift to Respondent

Study subjects receive $22.50 to participate for their time, effort, and invasive nature of the biological sample. The dollar amount was arrived at after consultation with clinicians and epidemiologists at each of the four study centers and the Queen Elizabeth hospital and considered a reasonable amount that is needed in order to obtain high participation rates. This amount of remuneration has been approved by the NCI Special Studies Institutional Review Board (SSIRB) and the NIH IRB (**Attachments 6A,B**).

A.10 Assurance of Confidentiality Provided to Respondents

Procedures have been developed to protect the security of all information collected from study subjects and approved by the NCI SSIRB and NIH IRB (**Attachments 6A,6B**). All information will be kept private to the extent allowable under the law. All files sent from interviewers to the study coordinating center (University of Hong Kong School of Public Health) by internet are fully encrypted. All original questionnaire responses and other study documents with personal identifiers are kept securely at the study coordinating center. All scanned medical records have the subject name blocked out before scanning and be identified by only a subject ID number. All biological samples are identified only by a subject ID and not have any personal identifiers, and these samples when sent from the NCI biorepository to laboratories for analysis are characterized by identifier numbers only. Biological sample analytic results are sent to for addition to the study database. All personal identifiers are stripped from data analysis files prepared by the study coordinating center for study investigators. Finally, no individual results are presented in publications or other reports. See **Attachment 1** for additional details.

The study is conducted by administering contact, screening, informed consent and questionnaire forms [**Attachments 7 (initial contact), 3 (screening), 8 and 9 (consents), 3 (core questionnaire),** and **5 (occupational modules)].** The study design was approved by the NCI SSIRB on July 11, 2014 (**Attachment 6A**) and NIH IRB on July 21, 2014 (**Attachment 6B**); the IRB at University of Hong Kong on August 29, 2014 (**Attachment 6C**), which is the overall study coordinating center and the study center for Hong Kong; Sichuan University Huaxi (West China) Hospital on September 10, 2014 (**Attachment 6D1,2**), which is the coordinating center for Chengdu; and Tianjin Medical University Tumor Hospital on July 11, 2014 (**Attachment 6E1,2**). The AsiaLymph study organization is shown in **Attachment 12**.

The Privacy Act does not apply to an international population.

A.11 Justification for Sensitive Questions

Personally identifiable information (PII) are collected. Most questions asked in the study questionnaire (**Attachment 1,3**) are typically not considered sensitive. Some questions, such as those about alcohol consumption, medical history, family history, reproductive health, and family income may seem sensitive to some respondents. However, these are important factors to evaluate as possible confounders and are known or suspected risk factors in and of themselves for lymphoma and represent questions that are common to health studies. National ID numbers are important to collect since they are often used as a way of identifying patients and their medical records in hospital databases and are unique. Respondents are informed that their responses are kept secure and they have the right to skip any questions even if they consent to the interview as a whole (**Attachments 8,9**).

A.12 Estimates of Annualized Burden Hours and Costs

Based on our plans for continued enrollment over the next 3 years, the annual burden for the respondents, the pathologists, and the interviewers is estimated to be 3,262 hours which amount to a total of 9,786 hours over a three-year data collection period (Table A.12-1). It is estimated that 2,110 additional patients per year will be screened to identify 1,801 potentially eligible study subjects, and of those approximately 967 cases and 300 controls per year will be consented for participation in the study over the next three years. Controls are drawn from patients seen at the same hospital for diseases/conditions that are unlikely to be associated with risk factors under study, such as injuries and selected diseases of the circulatory, digestive,

genitourinary, and central nervous system (**Attachment 1**). The category of respondents is considered "Individuals."

Based on patients that have already been enrolled in the study, the average participation time has been: 75 minutes to complete the screener and core questionnaire, 15 minutes for the patient-specific computer-triggered occupational exposure assessment modules, and 15 minutes to collect the biospecimens, for a total time of one hour and 45 minutes (105 minutes) per participant (**Attachments 3 and 5**). About 50% of the study subjects do not have any job module asked and the most of the remaining 50% have about 1-3 job modules asked, and the overall estimated time burden is an average of 15 minutes.

In addition to the patients, the study interviewers and the hospital pathologists complete administrative forms related to the collection of the biospecimens (**Attachments 10 and 11**). The 10 pathologists fill out the form in Attachment 10 for 967 cases, for an average of 97 forms per pathologist. The 15 interviewers fill out the forms in Attachment 11 for 1267 subjects (967 cases and 300 control patients) for an average of 85 set of forms per interviewer.

Table A.12-1. Estimates of Annual Burden Hours

Types of Respondents	Instrument	Number of Respondents	Frequency of Response	Time per Response (Hours)	Annual Burden Hours
Potential Study Subjects	Screening Questions (Attachments 3A to 3F)	2,110	1	5/60	176
Eligible Potential Study Subjects	Consent Form	1,801	1	5/60	150
Consented Patient Cases	Core Questionnaire & Occupational Job Module (Attachments 3A to 3F, 5)	967	1	105/60	1,692
Consented Patient Controls	Core Questionnaire & Occupational Job Module (Attachments 3A to 3F, and 5)	300	1	105/60	525
Study Pathologists	Pathology sample request and tracking form (Attachment 10)	10	97	5/60	81

Interviewers	Tracking forms (Attachment 11)	15	85	30/60	638
Total			7,423		3,262

Based on a weighted median hourly wage rate of $7 for study subjects (Attachment 13) using data from Hong Kong (http://www.statistics.gov.hk/pub/B10500092015QQ01B0100.pdf) (Attachment 14), Chengdu (http://www.cdstats.chengdu.gov.cn/detail.asp?ID=85814&ClassID=0205) (Attachment 15), and Tianjin (http://www.gov.cn/xinwen/2014-12/29/content_2797836.htm) (Attachment 15), $42 for pathologists and $15.50 for interviewers (Westat Contract #HHSN261201200075C) (Attachments 16 and 17), the annualized average cost is $31,092 which amounts to $93,276 over a three-year period (Table A.12-2).

Table A.12-2. Annualized Cost to Respondents

Type of Respondent	Instrument	Annual Burden Hours	Hourly Wage Rate	Respondent Cost
Potential Study Subjects	Screening Questions	176	$7.00	$1,232.00
Eligible Potential Study Subjects	Consent Form	150	$7.00	$1,050
Consented Patient Cases	Core Questionnaire & Occupational Job Module	1,692	$7.00	$ 11,844.00
Consented Patient Controls	Core Questionnaire & Occupational Job Module	525	$7.00	$3,675.00
Study Pathologists	Pathology sample request and tracking form	81	$42.00	$3,402.00
Interviewers	Tracking forms	638	$15.50	$9,889.00
Total		3,263		$ 31,092.00

A.13 Estimates of Other Total Annual Cost Burden to Respondents and Record Keepers

There are no direct costs to the patients, pathologists or interviewers other than their time to participate in the study.

A.14 Annualized Costs to the Federal Government

The annualized cost to the government is $736,119. This includes contractor costs, NCI staff time, and biological collection, biological transport, biological sample analysis and data analysis. The contractors for this study are epidemiologists who have extensive experience conducting case-control epidemiological studies, including providing quality control oversight and data management and analysis. The federal NCI personnel include two Senior Investigators involved in all aspects of study oversight and operations, two Staff Scientists, and one Post-Doctoral Fellow who are involved in data management, study operations, and quality control. The total costs to the Federal Government over the five-year period of data collection and analysis are estimated to be $3,680,596 (see Table A.14-1).

Table A.14-1. Annual Cost to the Federal Government

	5 YEAR TOTAL	ANNUAL AVERAGE
Contractor Costs	**$2,065,521**	**$413,104**
NCI Personnel Subtotal	**$510,075**	**$102,015**
Senior Investigator, Captain/O-6$245,000 (10% Effort)	$122,500	$24,500
Senior Investigator, $170,000 (20% Effort)	$170,000	$34,000
Staff Scientist$93,500 (25% Effort)	$116,875	$23,375
Staff Scientist, Title 42, $140,000 (10% Effort)	$70,000	$14,000
Post-doctoral CRTA Fellow, $61,400 (10% Effort)	$30,700	$6,140
Biologic sample and data analysis	**$545,000**	**$109,000**

Pathology Reagents	**$320,000**	**$64,000**
Travel Costs for NCI Staff for Study Management	**$240,000**	**$48,000**
Total	**$3,680,596**	**$736,119**

A.15 Explanation for Program Changes and Adjustments

This is a request for a revision. We are requesting OMB approval to increase the sample size of the study by 2,900 cases and 900 controls, who we plan to enroll over the next three years in three study centers (Hong Kong, Chengdu, and Tianjin, China). The consent forms have been updated to reflect the proposed increase in total subjects to be enrolled in the study. There is also a small change to the screener in which potential study subjects are asked if they have had a prior history of myeloid neoplasms or acute leukemia (in addition to being asked if they have had a prior history of lymphoma) and no change to the questionnaire, so content and information being collected is the same and the amount of time required to administer the questionnaire is the same as in the previous request to OMB.

A.16 Plans for Tabulation and Publication and Project Time Schedule

Full-scale data collection, cleaning and analyses are followed by publication in peer-reviewed, scientific journals. Our project time schedule is given in Table A.16-1.

TABLE A.16-1. PROJECT SCHEDULE FOR PHASE III	
Component	**Time after OMB approval**
Data collection	1 day to 36 months after approval
Data editing	1 week to 40 months after approval
Data analysis	40 to 60 months after approval
Publication	42 to 60 months after approval

Since lymphoma and myeloid neoplasms comprise a group of related yet heterogeneous diseases, each characterized by the malignant transformation of lymphoid or myeloid cells but with distinctive morphologic, immunophenotypic, genetic, and clinical features, we will analyze risks by lymphoid and myeloid subtype as well as larger subgroups. For analyses by subtype, odds ratios (ORs) and 95% confidence intervals (CIs) are derived for each risk factor from polytomous unconditional logistic regression models adjusting for matching variables. P values for the linear trend will be computed for continuous variables and using ordinal variables. To evaluate heterogeneity among lymphoma subtypes, we will use 2 statistical approaches. First, we will conduct a homogeneity test in the polytomous model, testing the null hypothesis that the regression coefficient for each risk factor was the same for all subtypes. Values of P less than .05 will be considered to provide evidence of heterogeneity. The test for homogeneity has the greatest power to detect risk differences when the risks for the subtypes all vary slightly from one another. Second, we will analyze all possible case-case pairwise comparisons using dichotomous logistic regression models (Morton et al. 2008). We will compute test the null hypothesis that the particular risk factor does not discriminate between the 2 disease groups modeled. To account for the pairwise analysis, we will apply a Bonferroni correction. In contrast to the test for homogeneity, the pairwise analysis has the greatest power to detect risk differences when the risk for one disease group is distinct from the other(s). For risk factors with more than

2 categories, we will use the ordinal variable for the homogeneity test and pairwise analysis. Analyses will also be conducted for larger lymphoma subgroups including NHL and B-cell lymphomas, all myeloid neoplasms and acute myeloid leukemia using unconditional logistic regression models in order to utilize all controls, adjusting for the matching factors. Analyses will also be conducted for the entire lymphoma study population using conditional logistic regression models.

For genetic analyses, standard methods are used to test the effect of each SNP. We will also use a new powerful and flexible subset-based approach to the combined analysis of heterogeneous traits, which is an approach that agnostically explores subsets of the traits to identify the strongest association signal and then evaluates the significance of the detected association using efficient adjustment for multiple correlated tests involved (N. Chatterjee, Chief DCEG Biostatistics Branch, personal communication). Data obtained from the study are analyzed using the STATA and SAS package of statistical programs. Data are cross-classified by age and sex, but analyses by sex-specific groups will also be performed.

Initial analyses are conducted for lifestyle risk factors, occupational exposures, environmental exposures, viral exposures, and genetic main effects. Exploratory gene-environment interaction analyses will also be conducted. We will also conduct genetic pathway analysis to evaluate whether the set of genes in a well-defined pathway (e.g., Th1/Th2 pathway) are associated with the disease risk. This type of analysis is particularly helpful in situations when the pathway is enriched with multiple SNPs with small effects. All models are adjusted for sex, age, study center, and date of enrollment (the control matching factors) and education. Additional potential confounders are selected based on initial analyses of the study data set and through identification of well-established risk factors in the literature. Essentially the entire study

population are comprised of Han Chinese so race-specific analyses will not be conducted. The 2012 memo for justification for limited ethnic/racial categories is shown in **Attachment 18**. Analyses by sex-specific groups are performed. Adjustments for potential confounding factors will depend upon the exposures under consideration.

Separate questionnaire-based publications will include the relationship between lymphoma, myeloid neoplasms and occupational exposure to trichloroethylene, occupational exposure to benzene, occupational exposure to chemical solvents, occupational exposure to formaldehyde, farming, residential proximity to traffic, residential proximity to incinerators, animal exposures, indoor air pollution from heating and cooking, family history, medical conditions, childhood crowding, early childhood exposures, UV exposure, physical activity, alcohol use, tobacco use, selected dietary factors (e.g., green tea, soybean products), hair dye use, reproductive history, and socioeconomic status. As indicated in **Attachment 1**, there is support in the literature for a relationship between each of these types of exposures and lymphoma. Further, descriptive papers on the distribution of lymphoma subtypes will be published. In addition, chemical and molecular laboratory analyses of the collected biological samples including pathology samples will result in separate publications on the relationship between lymphoma and organochlorine plasma levels, exposure to Hepatitis A, B, and C, Epstein Barr Virus, and genetic susceptibility using cutting-edge technology at the NCI. Additional publications will be produced on potential gene- and viral-chemical interactions.

A.17 Reason(s) Display of OMB Expiration Date is Inappropriate

There are no reasons to preclude display of the OMB expiration date on the questionnaires.

A.18 Exceptions to Certification for Paperwork Reduction Act Submissions

There are no exceptions to the certification statement.