Ashland Nyanza Health Study

Final Report

April 2006

Massachusetts Department of Public Health
Center for Environmental Health

ACKNOWLEDGEMENT

This study could not have been completed without the support and assistance of the Ashland community and particularly the Ashland Nyanza Health Study Community Advisory Committee. They are a true testament to the value of community partnership in public health investigations. We also acknowledge Ashland High School administrative staff, Ashland public health officials and the Ashland Cedar Street Fire Department for allowing frequent use of their facility for study meetings, as well as Representative John Stefanini and Senator David Magnani for their efforts in this important public health study.

In addition, the MDPH gratefully acknowledges all of the individuals who devoted their time to participate in the Ashland Nyanza Health Study. We would specifically like to acknowledge the efforts of Kevin Kane who in the face of his illness never faltered in his search to find answers. We would also like to thank the entire Kane family and many others in the community for their unending dedication and collaboration in this study.

This report was supported in part by funds from the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) trust fund provided to the Massachusetts Department of Public Health from the Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services.

Table of Contents

Figures

Lists

Tables

Appendices

I. Background/Introduction

The Town of Ashland, Massachusetts is located approximately 22 miles west of Boston and has a population of approximately 14,674 people (Ashland 2004). Although Ashland is predominantly a residential community, historically it was also home to some industry including Nyanza, Inc., which operated a dye manufacturing facility in the town from 1965 to 1978. However, the Nyanza site has a lengthy history in the town. During 1917 and 1965, a number of companies operated at the Nyanza site location. The Nyanza Inc. Company was one of the first and largest dye manufacturers in the United States. During the early 1980s several hydrogeologic and environmental studies at the Nyanza property documented widespread chemical contamination at and around the site. Consequently, in 1983 the U.S. Environmental Protection Agency (USEPA) placed the Nyanza property on the National Priority List as the Nyanza Chemical Waste Dump. In the past, liquid wastes were discharged from the Nyanza site into the environment in several ways including into an underground vault, unlined lagoons, and nearby brooks and wetlands. More than 100 different chemicals including volatile organic compounds (VOCs), semi-volatile organic compounds (SVOCs), dye manufacturing compounds, and metals were detected on the approximately 35-acre site (USEPA 2004).

A Public Health Assessment for the Nyanza Chemical Waste Dump completed by the Massachusetts Department of Public Health (MDPH) in 1994 identified a number of exposure pathways from the Nyanza site (MDPH 1994). This assessment demonstrated that opportunity for human exposure to Nyanza site contaminants in the past was high and included exposures to children playing in the soils and lagoons on-site as well as in the Chemical Brook adjacent to the site. The Public Health Assessment concluded that the Nyanza Chemical Dump is a public health hazard because humans have probably been exposed to hazardous substances at concentrations that may result in adverse human health effects. Human exposure in the past occurred due to: 1) ambient air emissions from the site via inhalation, 2) contaminated on-site media via dermal contact, incidental ingestion and inhalation and 3) contaminated fish, surface water and sediments.

Although the 1994 Public Health Assessment determined that exposure of adults and children to contaminated media in or around the Nyanza site occurred in the past, information regarding the community's health status was not available at that time to determine if exposure could be related to adverse health effects, such as cancer among Ashland residents. However, as part of the Public Health Assessment for the Nyanza site, two related health studies were conducted. The first health study was a descriptive study of bladder and kidney cancer incidence and mortality in the town of Ashland (MDPH 1988). Bladder and kidney cancer were targeted because these cancers have been associated with occupational exposure to azo dye manufacture using benzidine and 2-naphthylamine, two contaminants detected at the Nyanza site. The study found that elevations existed in the two cancer types. Specifically, two areas of the town were identified where bladder and kidney cancer cases appeared to be geographically concentrated. The second health study, also conducted by the MDPH, was a case series investigation of the individuals diagnosed with bladder and kidney cancer (MDPH 1990). This second study attempted to address questions about risk factors and possible environmental exposures that could be related to the pattern of bladder and kidney cancer in Ashland but were not evaluated in the first health study. The study concluded that the number and geographic distribution of the bladder and kidney cancer cases observed in Ashland during the study period was not an atypical finding and did not show a pattern that appeared associated with the Nyanza Chemical site (MDPH 1990).

Upon completion of the second health study, the MDPH convened an expert panel to review all of the available environmental and health information obtained during the two health studies. Although the panel concluded that the pattern of bladder and kidney cancer did not appear atypical, it also determined that potential exposures to environmental contaminants from the Nyanza site to Ashland residents may have existed in the past and these potential exposures could have contributed to a variety of disease outcomes. The panel recommended to local health officials that, if possible, information on individuals who resided in Ashland during the 1960s be identified (MDPH 1990). The panel believed the 1960s to be a critical time period of interest based on possible population changes in the town during the 1970s and 1980s and the latency period of

diseases, such as cancer, that could be associated with exposures from the Nyanza site. Consequently, in the conclusion of the Public Health Assessment, the Health Activities Recommendation Panel (HARP) recommended that should the pattern of cancer among Ashland residents change, a retrospective epidemiological study be conducted to evaluate increased adverse health outcomes to residents of the town (MDPH 1994).

In 1998, residents of Ashland reported to the MDPH that children who lived in the town and played on the Nyanza site in the past developed similar types of rare cancers as young adults. Of greatest concern was the report of five men in their early 20's that developed various types of soft tissue sarcoma. These individuals were in the general age range identified in the Public Health Assessment as the population with the greatest opportunities for exposure to Nyanza site contaminants. Further, the reported individuals were school-aged residents of Ashland during the time period that was described by the expert panel as a critical period of interest in terms of possible exposures from the Nyanza site. Although the previous health studies of bladder and kidney cancer in Ashland concluded that the pattern of these two cancer types did not appear associated with the Nyanza Chemical site, the studies reviewed the pattern of bladder and kidney cancer among current Ashland residents during the period 1982 through 1986. The age range at the time of diagnosis for the majority of individuals in the investigations was between 50 and 85. Only one individual who was included in the previous studies of bladder and kidney cancer was younger than age 45. Further, none of the individuals had attended school in Ashland.

An additional review of cancer incidence data for the town of Ashland available from the Massachusetts Cancer Registry at the time the suspected cluster of soft tissue sarcoma was reported indicated that there was an overall increased incidence of cancer among younger residents of the town (i.e., those under the age of 40 years). The increase occurred among individuals who were residents of Ashland at the time of diagnosis during the years 1982 to 1994 and was for all cancer types combined. The incidence of all cancers combined was statistically significantly elevated among both males and females during this time period. Please refer to *Table 1* for a summary of these data.

While the cancer registry data indicated an increase in the incidence of cancer among younger residents of the town through 1994, there was no way to determine the pattern of cancer incidence among former residents who resided in Ashland during the critical period of interest with respect to opportunities for exposure form the Nyanza site.

Much is still unknown about the potential health effects of exposure to complex chemical mixtures. Identification of the types of compounds detected at the Nyanza site can help in understanding what types of health effects one might expect to see if exposure occurred to only one of these compounds. However, Ashland residents who came in contact with the Nyanza site, particularly children, were likely exposed to a wide range of chemical mixtures resulting from hazardous waste disposed of at the Nyanza Chemical Waste Dump. The toxicology of complex chemical mixtures is a critical field in human health. Childhood physiology, metabolism, diet and chemical specific response may increase a child's sensitivity to chemicals (Schmidt 1999; Hanson et al. 1998). Depending on the chemical, the stage of a child's growth and development may be a critical variable in evaluating the potential toxicity of chemicals at hazardous waste sites. Some epidemiological studies indicate that exposure to complex chemical mixtures deserve consideration as a risk factor for the development of various cancers (Zeliger 2003, 2004). Other studies suggest that an association between chemical exposures and soft tissue sarcomas is biologically plausible (Costani et al. 2000; Ericksson 1988, 1990; Hardell and Erickson 1988; Hoppin et al. 1999; Kogan 1988, Zahm et al. 1988, 1989, 1996; Zahm and Fraumeni 1997). Based on these findings, the more reports of cancer diagnoses among former Ashland residents, and information regarding a pattern of increased cancer incidence among young adults in Ashland, the MDPH initiated the Ashland Nyanza Health Study. This study is a large epidemiologic study aimed at determining whether opportunities for historical exposure to contaminants from the Nyanza site could be related to increased cancer diagnoses among children and young adults who lived in Ashland at the time the site was operational.

Nyanza Site History

The Nyanza site is located in the central section of Ashland in a largely residential area. The closest residences are located approximately 100 to 150 feet north/northeast of the site perimeter on Pleasant Street (*Figure 1*). Chemical sludge generated by the on-site dye manufacturing processes was disposed of in an area of high ground referred to as, "Megunko Hill". Two open waste ponds were located on the top of Megunko Hill. Sludge containing large quantities of heavy metals such as mercury, chromium, lead and cadmium spilled and washed down the sides of Megunko Hill to the wetlands below. Reports from the Ashland community describe occurrences where children playing in the Megunko Hill area would often return home with blistered hands and discolored clothing. Heavy metal sludges and organic solvents were disposed of in the lower industrial area of the site in active lagoons or the underground vault. Solid waste from chemical precipitations and dye pigment processes was put in 55-gallon drums and buried in different areas of the site. Site related wastewater was discharged to Chemical Brook, an intermittent stream located along the northern site boundary, as well as Trolley Brook, located in the northeast corner of the site. The groundwater beneath the site is highly contaminated with VOCs, SVOCs and metals. In 1986, a fish advisory was posted for the Sudbury River in Ashland to its confluence with the Assabet River in Concord, MA. Residential properties located in close proximity to the Nyanza Chemical Waste Dump were also found to contain soils and sediments contaminated with high levels of mercury (MDPH 1994).

The Massachusetts Department of Environmental Quality Engineering (currently the Massachusetts Department of Environmental Protection (MDEP)) restricted site access in 1982 by installing a partial fence. However, portions of the site were accessible to the public until the fence line was expanded in 1985. Until the early 1980's, children used the site as a recreational area. Some of the activities that took place at the site included biking, swimming, wading, skating, and building forts. A walking path transversed the site from Pleasant Street to Ashland High School located on West Union Street. Ashland High School students reported using the path through the site as a shortcut walkway to and from school (MDPH 1994). A baseball field was located 150 feet north of the

northern wetland area of the site and Stone Park and a neighboring playground area are located approximately ¼ mile southeast of the site.

The USEPA developed a remedial plan consisting of five phases to clean up contamination at the Nyanza Chemical Waste Dump. The remedial plan included initial actions at the site and four long-term remedial phases aimed at source control and clean-up of soil, off-site groundwater, wetlands and drainage ways, and the Sudbury River (USEPA 2004). Two of the four long-term remedial phases are complete; source control and soil cleanup as well as clean up of mercury contaminated sediments in on-site wetlands and drainage ways. Currently, risk assessment activities are on-going to determine a remedial solution for cleanup of contaminated groundwater and Sudbury River sediment and fish.

II. Study Design and Methods

Upon initiation of the Ashland Nyanza Health Study, the MDPH established a Community Advisory Committee. The Community Advisory Committee was made up of residents of the town of Ashland as well as the Ashland Board of Health and representatives from both the USEPA and the MDEP. The advisory committee provided input into the study design and approach and played an active role in researching source information for current and former Ashland residents in order to establish the population cohort of interest for the study.

A. Study Hypothesis

Epidemiologic studies involve statistical hypothesis testing which typically focus on the null hypothesis (H_o). The null hypothesis is the hypothesis that there is no association between two factors or variables. The null hypothesis states that the results observed in a study are no different from what might have occurred as a result of chance alone (Last 1988). By conducting statistical analyses and comparisons, researchers determine whether the data collected in the study provides evidence against the null hypothesis. If so, then the null hypothesis can be rejected in favor of the research or study hypothesis

(H_a). Stated as a null hypothesis, the hypothesis for the Ashland Nyanza Health Study is (H_o): exposure to the Nyanza Chemical Waste Dump is not a factor related to the incidence of cancer among children and young adults who resided in Ashland during the period 1965 to 1985.

The primary aim of the Ashland Nyanza Health Study was to examine exposure opportunities associated with the Nyanza Chemical Waste Dump to former residents of Ashland, Massachusetts as a risk factor for development of cancer, particularly sarcoma. The main study hypothesis of this investigation is (H_a): exposure to contamination at the Nyanza Chemical Waste Dump increases the risk of cancer among children and young adults who resided in Ashland during the period 1965 to 1985. As indicated in the 1994 Public Health Assessment, this is the population most likely to have been exposed to Nyanza site contaminants and the time period when the Nyanza site was operational and site access unrestricted therefore offering the greatest opportunity for exposure to the public.

B. Study Design

The Ashland Nyanza Health Study is a retrospective cohort study which attempted to recreate the population of children between the ages of 10 and 18 who resided in Ashland during the years 1965 to 1985. The objective of the study was to determine and compare the incidence of cancer among exposed and non-exposed individuals of the study cohort (i.e., individuals who reported contact with the Nyanza site versus individuals who did not report contact with the site). The cohort was restricted by age based on opportunity for past exposure. That is, by restricting the study to individuals who were aged 10 to 18 during the time period of interest (i.e., 1965 to 1985); the study targets those Ashland residents that likely had the greatest opportunity for exposure to the site. As previously described, opportunity for exposure to site contaminants for children younger than the targeted study population (i.e., less than age 10) and adults (i.e., greater than age 18) in Ashland was considered to be low (MDPH 1994).

C. Recruitment of Study Participants

Typically recruitment of former residents in this type of study design is difficult at best given that individuals from any community are not likely to remain in one residence for the course of their lives. Residents of Ashland are no different than residents of other communities, with the exception of the degree to which the community seeks to maintain contact with former residents. Study participants were identified from public school records and lists of graduates from Ashland High School. Upon the recommendation of a Community Advisory Committee member, the MDPH obtained class lists for all graduates of Ashland High School during the years 1972 to 1992. These are the high school class years that correspond to individuals who would have been between the ages of 10 and 18 during the period 1965 to 1985. With assistance from reunion class representatives and the Community Advisory Committee, current address and contact information was obtained for members of each Ashland High School class. The total annual number of graduates from Ashland High School is between 100 to 150 individuals. Therefore, it is estimated that the maximum number of study subjects that comprised the cohort population was approximately 2,500. Based on updated lists of Ashland High School graduates, 82% of the cohort of interest resided in Massachusetts at the time of the study. Thirty-three percent (33%) of the cohort of interest were current Ashland residents at the time of the study, 49% were former residents of Ashland living in Massachusetts and a smaller portion (approximately 18%) were former Ashland residents residing outside of Massachusetts. In addition, a small portion of Ashland residents had attended private schools in adjacent communities. Therefore, graduate lists from both Marian High School and Keefe Technical High School were obtained to identify Ashland residents that could potentially be part of the cohort.

Recruitment of potential study subjects was accomplished primarily through direct mailings to individuals in the cohort of interest. Using current address information obtained through Ashland High School class listings and high school class reunion lists, each member of the study population was mailed a letter requesting their participation in the Ashland Nyanza Health Study. Therefore, both current and former Ashland residents

including individuals that were not able to be reached through local community outreach efforts had an equal opportunity to participate in the study. However, in order to maximize study participation, a variety of additional recruitment methods were employed. The MDPH established an interactive website for the study. The website provided status reports and updates on the study progress for the Ashland community but also allowed potential study participants to contact MDPH staff directly through electronic mail to enroll in the study. The website was linked to the Nyanza site web page established by the USEPA as well as the Town of Ashland web page. The MDPH also established a national toll free 1-800 number for potential study participants to contact staff regarding study participation. The 1-800 number was staffed by MDPH researchers during regular work hours and directed to an answering service during weekend and evening hours.

Additional outreach methods included announcements in the local press and cable television, community postings and notices, participation in local community events (e.g., Ashland Day). Study information was also included in high school reunion mailings to Ashland High School graduates. With assistance from the study's Community Advisory Committee, MDPH staff also hand delivered study notices requesting participation to each residence in the town of Ashland.

Upon identification of a potentially eligible study subject, a letter was sent to the individual that briefly described the study and requested participation (see Appendix B). The letter included a consent form and a study information sheet requesting contact information (e.g., current address and telephone number, verification of Ashland residence) for each participant. Individuals who did not agree to participate were asked to complete a refusal section of the study information sheet indicating their name, class year and reason for refusal. Refusals were entered into a database by high school class year to provide baseline information to characterize non-participants. This information was used to evaluate reasons for non-participation and to determine whether non-participation varied by class year and potentially exposure opportunity.

After no response to the first mailing, individuals were sent an additional mailing approximately two to three weeks later, again requesting study participation. If no response was received after a second mailing, participants were contacted by certified mail with an additional letter strongly emphasizing the significance of study participation. Individuals who ultimately refused to participate were not further contacted.

D. Data Collection

The following data sources were used to collect data for the study.

- Ashland Health Study Questionnaire – used to collect detailed information on residential history, occupational history, medical and family history of cancer, alcohol and tobacco use as well as demographic information (age, date of birth, etc.).

- Exposure Assessment Questionnaire – used to collect detailed information on Nyanza site contact, play activities, potential site exposures and additional hobbies and exposure activities (i.e., potential pesticide exposure).

- Massachusetts Cancer Registry – for identification and confirmation of incident cases of cancer diagnosed among current Ashland residents and former Ashland residents with a Massachusetts residence.

- Medical records – for identification and confirmation of incident cases of cancer diagnosed among former Ashland residents with non-Massachusetts residence.

E. Ashland Health Study Questionnaire

A standardized, structured questionnaire (see Appendix C) was used to obtain information on both disease status and exposure opportunities to the Nyanza site. Information about other known or suspected risk factors that might confound the association between exposure to the Nyanza site and development of cancer was also included. The questionnaire was peer reviewed and then pilot tested by a group of

volunteers from the Ashland community prior to undergoing final revisions and administration to study participants. The Ashland Health Study Questionnaire was developed using Teleform software (Verity 1999). Teleform is scanning software that integrates both questionnaire form design and scanning reader software. This process allowed the questionnaire to be pre-formatted primarily as close ended questions resulting in the collection of pre-coded categorical data. The process eliminates the potential for out of range responses by study participants and reduces end stage data cleaning.

Both exposure and disease status were determined based on participant responses to the questionnaire. The outcome variable of interest for the study was a cancer diagnosis. Disease status was determined based on self reported responses to questions during the interview. Information about cancer diagnoses reported during the interview was then confirmed through medical record review and review of the Massachusetts Cancer Registry data files. Determination of individuals in exposed and non-exposed groups was based on participant responses to the Exposure Assessment Questionnaire (i.e., Section C of the Ashland Health Study Questionnaire). The exposure assessment portion of the questionnaire included questions regarding activities on and around the Nyanza Chemical Waste Dump. To assist in answering activity and contact questions about the site, each study participant was provided with two maps; a Reference Map (Map A) and an Exposure Map (Map B). These maps are illustrated in *Figures 2 and 3*. Map A, the Reference Map, was distributed to study participants to provide a frame of reference for areas in the Town of Ashland and the general location of the Nyanza Chemical Waste Dump within Ashland. Map B, the Exposure Map, depicts the Nyanza site and delineates nine discrete areas of interest (labeled as Areas A through I) which are located on or near the site property. The defined Areas A through I correspond to areas of contamination on the site and adjacent site properties but not in a manner that would identify the nature and extent of site contamination and could potentially bias participant responses. The map also provides some description of land features within the defined Areas A through I. For example, Area E depicts the western wetlands and a baseball field and Area D depicts the Megunko Hill area and the location of two waste ponds. Areas A, B, D, and E generally

comprise the property bounds of the Nyanza Chemical site. Areas C, F, G, and H, depict off-site areas of the Nyanza site where chemical contamination was detected. The Exposure Map B was developed based on information contained in environmental investigations of the site, historical aerial photographs, as well as information obtained in consultation with the USEPA Remedial Project Manager for the Nyanza site (MDEQE 1980; Fauss 1982; Jaratowicz 1952; NUS 1992; Stokely 1989; USEPA 2000). Members of the Community Advisory Committee also provided input on specific map areas in order to ensure that names or terms commonly used by Ashland residents to describe areas associated with the site were included to aid study participants in recall of map locations. Map B contains a number of cartographic features aimed at focusing the attention of study participants on the Nyanza site and surrounding areas. A standard accompanying interview script allowed for review of these features with descriptions of the specific exposure areas so that all study participants were provided ample orientation to Map B before responding to the Exposure Assessment portion of the questionnaire.

F. Interviews

Once signed consent to participate in the study was received, study participants were contacted by telephone to arrange a convenient interview time. Interviewers trained in standardized non-directive interviewing techniques administered questionnaires by telephone. Interviews were approximately 60 minutes in length and interviewers were blinded to the study hypothesis and both the exposure and disease status of the participants.

Approximately one week prior to the interview appointment, each study subject was mailed a reminder and confirmation notice, and a package including the Reference Map, Exposure Map and a consent form for medical records review. Medical records consent forms were completed during the interview for those individuals who reported a cancer diagnosis and returned to the MDPH by self-addressed stamped envelope. If the study participant was not at home at the scheduled interview time, follow up telephone calls were made and messages were left requesting that the individual call the MDPH at a toll

free telephone number in order to reschedule the interview. A response log and contact sheet was used to record the outcome of each phone contact attempt. If there was no answer after several attempts to contact a study participant by telephone, a follow-up letter was sent by certified mail to the individual requesting that they contact MDPH to reschedule the interview. As a final attempt to increase overall study participation, study participants who had responded positively that they would participate but were non-compliant or difficult to contact at the time of the pre-scheduled interview, were mailed the questionnaire and instructions on how to complete it. Participants who were mailed the questionnaire were provided with the toll free number which they could call and receive telephone assistance from interviewers to complete the questionnaire. Study questionnaires were mailed to 62 individuals. Eight individuals completed the questionnaire by mail. Interviews took place between 2000 and 2002 with the majority conducted in 2000 (N=1,232 or 89%).

G. Case Confirmation/Medical Records Review

Disease status was determined initially based on self-reported responses to questions during the interview. Confirmation of cancer diagnosis reported by study participants at interview was obtained from the Massachusetts Cancer Registry and/or information obtained from medical records reviews. Study participants who self-reported a diagnosis of cancer during interview were instructed to complete the medical records consent form and return the form in a self-addressed stamped envelope to the MDPH. MDPH staff then contacted the appropriate physicians and/or hospitals to obtain and abstract medical records information for confirmation of a cancer diagnosis and any other pertinent information related to disease status. An MDPH consulting physician reviewed medical records and cancer diagnoses.

H. Determination of Exposure Status

The exposure of interest evaluated in this study was contact with the Nyanza Chemical site. As mentioned previously, there was a variety of exposure pathways and exposure points or specific areas of the Nyanza site where children and young adults historically

may have come in contact with contaminants of concern. Both the exposed and non-exposed groups were derived from the same study population. Exposure status for study participants was determined based on responses to the exposure assessment portion of the study questionnaire. The exposed group consisted of study participants who reported contact with and activities at specific areas of the Nyanza site as described and displayed in Exposure Map B (Figure 3). The non-exposed group or comparison population was derived from the same study population of Ashland resident who were between the ages of 10 to 18 during the study period. However, this group was defined as those individuals who did not have contact with the Nyanza site or specific site areas. Because the exposed and non-exposed groups were derived from the same study population, the comparability of the two groups with respect to known confounders is increased.

I. Data Management

Each study participant was assigned a unique numerical identifier to protect the confidentiality of study participants. Both data coding and data entry were accomplished through use of Teleform software. At the completion of data collection through study participant interviews, all questionnaires were scanned with the Teleform software. The Teleform software evaluates completeness of the scanned questionnaire and detects incorrect or questionable markings. Research staff reviewed the scanning results to check the completeness of the questionnaire responses and accuracy of the collected data. If information was missing or apparently erroneous, study participants were contacted by telephone and the correct information noted and initialed on the questionnaire for re-scanning. After completed questionnaires were scanned the data was exported to a SAS dataset for analysis.

J. Data Analysis

Statistical analysis consisted primarily of univariate and multivariate analyses designed to assess the relationship between exposures involving contact with the Nyanza site and other potential risk factors and their potential relationship to a cancer diagnosis. In addition, descriptive statistics for the case group and the cohort as a whole were

calculated and potential confounding factors assessed. All analyses were conducted using SAS Statistical Software. These analyses were mainly performed by using procedures in base SAS and SAS/STAT software (SAS 2001, 2003). Some analyses were conducted with the use the SAS Analyst application or specially written SAS macros.

K. Categorical Analysis

Because the questionnaire was preformatted, the majority of data were collected in a closed ended categorical or dichotomous format. Therefore, categorical analysis was the primary and preferred method of statistical analysis of study data. The majority of statistical analyses were conducted using the 2-way frequency table analysis task in the SAS Analyst application. This procedure utilizes the 2 x 2 contingency table with the rows being the exposure level and the columns being the presence or absence of the outcome of interest. The Pearson chi-square statistic, labeled as "chi-square", was used to test the association between the exposure and outcome and the associated p-value that provides a probability level for statistical significance. This test is based on expected frequencies versus observed frequencies. When sample sizes were small and any of the cell counts of the contingency table were below five, the Fisher's exact test, indicated by "(f)" next to the p-value, was used as a more reliable test of association. The standard 0.05 probability level was used to determine statistical significance for all statistical tests and 95% confidence intervals. Measures of association calculated as either the relative risk (RR) or odds ratio (OR) as well as corresponding 95% confidence intervals were also generated through categorical analysis.

L. Multivariate Logistic Regression

While most study questions were designed for a dichotomous response, making categorical table analysis the preferred method of analysis, many questions involved continuous or ordinal assessments of exposure. For these questions, multiple logistic regression was conducted using SAS Proc Logistic. Regression techniques are used to investigate the relationship between a categorical dependent variable and one or more

independent explanatory variables. Because the outcome of concern in this study is a dichotomous variable (i.e., cancer diagnosis), this analysis employs the technique of standard logistic regression. The Analyst application in SAS was used to generate corresponding test statistics and measures of association. Various statistical chi-square tests were calculated to assess the effects of the explanatory or independent variables, both individually and as a group, on the dependent variable. When multiple explanatory variables were included in the model, the tests describe the effects of each independent variable on the dependent variable while controlling for the other covariates in the logistic regression model, or holding their effects constant. P-values were generated for each test statistic to determine the level of statistical significance of each association. Wald chi-square estimates and corresponding p-values from the Analysis of Maximum Likelihood are reported for logistic regression results in the corresponding tables summarizing results of these analyses. In addition, odds ratio estimates were calculated for each covariate as the primary measures of association with 95% confidence intervals based on a 0.05 significance level.

M. Continuous Versus Categorical Variables

For some exposure variables, the frequency and/or duration of exposure was assessed with open-ended questions resulting in continuous response variables. When indicated, the distributions of these variables, or a combination thereof, were split into quantiles to ultimately create categorical exposure variables. Two methods of determining exposure categories were used. The first method divided the distribution into equal percentiles based on the number of responses. This resulted in exposure categories with equal numbers of responses, however it also allowed for the possibility of each exposure category having a very different range of values. The second method divided the distribution based on the range of response values reported. This method resulted in exposure categories with equal ranges of values, however it allows for the possibility of very different sample sizes within each exposure category. Both methods were used in conjunction with standard logistic regression for the most complete and thorough analysis.

N. Assessment and Control for Confounding

While the primary goal of the study was to assess the relationship between opportunities for exposures involving contact with the Nyanza site and the relationship to a cancer diagnosis, information on other potential risk factors was also collected. If a factor is associated with both the disease outcome and the exposure of interest, it can distort the true relationship between exposure and disease, resulting in an alternative explanation for the observed association. These factors must first be assessed and then held constant, or controlled for, during analysis. Some factors that could be independently associated with disease and exposure are demographic and behavioral characteristics and medical history information. Specific factors which may be particularly associated with a cancer diagnosis and/or contact with the Nyanza site, and for which information was collected include age, gender, family history of cancer, smoking, and alcohol consumption.

The potential confounding effects of age, gender, and family history of cancer were assessed using descriptive analysis and the two-sample tests for means and proportions, which test whether the measurements in the two populations, the case group and the non-case group, are equal or if the observed differences are statistically significant. These tests produce *t* and Z statistics, respectively, and associated p-values to assess the statistical significance of the observed differences. These statistical tests were used to screen for the presence of confounding factors and to assess the magnitude of any confounding so that these variables could be adjusted and controlled for in the statistical analyses.

Confounding factors can be controlled for in analysis through two methods, stratification and multivariate analysis. Stratification occurs when separate analyses are conducted on homogenous categories (or strata) of the confounding variable. The association between the exposure and the outcome for each stratum can then be compared to see if they differ appreciably with each other and with the crude estimate without stratification. If the results for each stratum are similar and these values are similar to the crude estimate, then the factor is not confounding the true association. However, if the results are similar to

each other but differ from the crude estimate, confounding has likely occurred and results from the stratified analysis can be used to estimate the association. While this is the preferred method of controlling for confounding with categorical data, it is difficult to simultaneously control for numerous factors through stratification. Therefore, multiple logistic regression was used to control for several variables at once. Using this multivariate analysis technique, the effect of each variable included in the logistic regression model can be estimated, while controlling for the effects of the other covariates.

Another scenario is that the stratum-specific associations differ appreciably from one another, which indicates that the factor is not a confounding factor but rather acting as an effect modifier or moderator variable. In this case, the factor is not considered as biasing the effect of exposure on disease outcome but rather the factor is interacting with the exposure as a risk factor in the process of producing the disease in the underlying population. Where a factor was identified as a possible effect modifier, all stratum-specific estimates are reported.

IV. Study Results

A. Descriptive Analyses

1. Study Participation

The target study population was all children between the ages of 10 and 18 who resided in Ashland during the years 1965 to 1985. Outreach and research efforts established a possible cohort of 2,751 individuals. Using primarily public school records and graduate lists from Ashland High School as a sampling frame, the study population was estimated at 2,618 individuals. Of these 2,618 individuals, we were unable to contact 564 due to incorrect address or contact information. Therefore the resulting study population was 2,054. Recruitment efforts yielded a sample population of 1,387 individuals who either graduated or attended Ashland schools during the study period of interest resulting in a response rate of 67.5%. One hundred and forty-two individuals refused study

participation (6.9%) and there was no response from 525 individuals (25.6%). The results of study participation are shown in *Figure 4*. Reasons for refusal included not interested in the study (51%), too busy or no time to participate (24%), proxy refusal (14%), not familiar enough with Ashland or area to answer questions (7%), physically unable to participate (4%). The distribution of refusals did not vary by gender or class year.

2. Study Cohort Description

The study cohort of 1,387 individuals was composed of 41% males and 59% females. The mean study age among the cohort was 39.1 years, with a range of 19 to 53 years. The mean age for male study participants was 38.6 years and the mean age for females was 38.1 years. There was no statistically significant difference in mean age between male and female study participants. Nearly all of the study participants were high school graduates (99%) and 53% had college level educations. Only a slightly higher percentage of males than females in the cohort had obtained a college education (53% for males versus 52% for females) and there was no significant difference in education level between male and female study participants. The majority of study participants were alive at the time of the study; eight individuals were deceased.

3. Case Group Description

A total of 73 individuals from the study cohort of 1,387 reported that they had received a diagnosis of some type of cancer at the time of interview. These 73 participants establish the case group. The mean age of the case group was 39.8 years. The age range of the cases at the time of diagnosis was 14 to 46 years. Sixty-six percent of the individuals in the case group were under the age of 35 at the time of their diagnosis. The distribution of cancer types reported by the 73 participants with a self-reported a cancer diagnosis is provided in *Figure 5*.

Of the 73 study participants with a self-reported a cancer diagnosis, the diagnosis was confirmed either by medical record review or through the Massachusetts Cancer Registry

for 55% (n=40) of the case group. For the remaining individuals in the case group the MDPH was either unable to obtain consent for medical records review or unable to obtain appropriate medical records to confirm a cancer diagnosis. The distribution of cancer types reported by the 40 participants with a medically confirmed cancer diagnosis is provided in *Figure 6*.

Thirty-four percent (n=25) of the case group were diagnosed with cancers that could be considered a rare cancer. These cancers were defined as rare because either the cancer type itself is a rarely occurring cancer with a low incidence or because the cancer is a more prevalent cancer type in general but rarely occurs at the ages that were observed in this study cohort. Rare cancers were defined as cancers where the age-specific incidence in the general population is less than 5 cases per 100,000 (SEER 2004). The primary study analyses were conducted with the entire case group of individuals who self-reported a cancer diagnosis (n=73). Additional analyses were also conducted with two separate subsets of the main case group: 1) restricting the case group to only study participants with a medically confirmed cancer diagnosis (n=40) and, 2) restricting the case group to only study participants diagnosed with a rare cancer (n=25).

4. Assessment of Potential Confounders

(a) Age

In order to determine the ages of the participants, all were assigned an interview year of 2001 and the study age was calculated by subtracting the birth year from the assigned interview year (2001). For the eight deceased participants in the cohort, the study age was considered the age at death. The mean study age among the case group and the non-case-group was 39.8 years and 39.0 years, respectively. A comparison of the mean study age between these two groups did not result in a statistically significant difference ($p < 0.325$) with the case group having only a very slightly higher mean age than the non-case group. In addition, logistic regression analysis of study age (as a quantitative variable) comparing the two groups resulted in an estimated odds ratio (OR) of 1.02 (95% CI: 0.98 to 1.06) and a chi-square (x^2) p-value of 0.323. Both results indicate that there is no

significant difference in age between study participants who reported a cancer diagnosis and those that did not. Therefore, age was not considered a likely confounding factor in analyses that compared these two groups and was not adjusted for in the analyses.

(b) Gender

The self-reported case group was composed of 33% males and 67% females, while the remainder of the cohort was composed of 41% males and 59% females. A table analysis resulted in a nearly significant difference between the gender distribution of the case group and the non-case group with a relative risk (RR) of 0.63 (95% CI: 0.39 to 1.02) and a corresponding p-value of 0.058, suggesting that males in the study cohort are at somewhat lower risk of having a cancer diagnosis than females. In comparable analyses, two-sample tests for means and proportions were used to compare the gender composition of each group. The difference in the distribution of males between the two groups resulted in the same p-value of 0.058. This analysis implies that males are at a nearly statistically significant reduced risk of a cancer diagnosis and females are at a slightly higher risk. However, it is possible that the observed association is due to the higher number of gender-specific cancer types comprising the case group. Female-specific cancer types (i.e., breast, cervical and ovarian cancer) accounted for 30% of the diagnoses among the case group. Comparison of the gender distribution of the case group and non-case group removing female-specific cancer types resulted in a relative risk of 1.07 (95% CI: 0.62 to 1.85), indicating no difference in risk between males and females. Because the difference is not statistically significant at the 0.05 probability level, gender was not considered to be a confounding factor included in the statistical analysis.

(c) Family History of Cancer

Family history is a primary factor in the etiology of many cancers and was assessed as a possible confounding factor to the exposure analyses of Nyanza site contact. A positive family history of cancer was considered to be a positive response to the question asking if the participant's biological mother, father, or siblings had received a diagnosis of cancer.

Of the 73 participants in the case group, 42 responded as having a family history of cancer (58%), as opposed to 552 out of the remaining 1,314 in the non-case group (42%). A two sample test of proportions tested the null hypothesis that there was no difference in the proportion of family members with cancer between the case group and the non-case group. This analysis resulted in a highly statistically significant difference between the two groups in the proportion of individuals who reported a family history of cancer ($p < 0.009$). A categorical table analysis confirmed the previous results (x^2 $p < 0.009$). These analyses resulted in a relative risk (RR) of 1.81 (95%CI: 1.15 to 2.84) indicating that study participants who reported a family history of cancer were also more likely to report a cancer diagnosis. Family history of cancer was considered a confounding factor and therefore controlled for in the exposure analyses.

(d) Behavioral Factors

Information collected regarding behavioral factors that may be related to a cancer diagnosis included smoking and alcohol consumption habits. There were no significant differences between self-reported cancer diagnosis and having ever smoked for six months or longer, currently smoking, the number of years smoked the number of cigarettes smoked per day, or smoking on the path in the woods behind Ashland High School. Also, there were no significant associations between having ever consumed alcohol, drinking on the path in the woods behind Ashland High School, or other locations at which drinking occurred.

B. Exposure Analysis

1. Nyanza Site Contact and Activities Analysis

As one method of assessing exposure to the Nyanza site, the study inquired about several activities that could have resulted in contact with site contaminants in specified areas both on and off the site. These areas labeled A through I are delineated on Map B and a description of each area is provided in List 1. Separate questions were also asked about specific locations or features of the Nyanza site within the defined Map B Areas A

through I. These include areas that were reported to have significant contamination such as Chemical Brook and the waste ponds located atop Megunko Hill. The broadest analysis involved having ever played or spent time in each of nine discrete areas labeled A through I located either on or near the Nyanza property. The majority of study participants (98%) reported either playing or having spent time in some area labeled on Map B. This is not unexpected given that the labeled areas contained the Ashland High School (Area C) and Stone Park (Area I), a location where a variety of community recreational events were hosted. The largest percentage of study participants (55%) reported that they spent time in the area north of the Nyanza site property (Area G), which includes a baseball field, the Sudbury River and Mill Pond.

A categorical analysis conducted with each area (A-I) as the exposure variable and a self-reported cancer diagnosis (SR) as the outcome (self report of cancer diagnosis versus no self-report of cancer diagnosis) resulted in a positive relative risk of a cancer diagnosis for all labeled areas depicted on Map B with the exception of Area C (Ashland High School, Ashland High School fields, and the path to Nyanza). The relative risks ranged from 1.17 to 1.82 however, none of the observed results were statistically significant (*Table 2*). As mentioned, the areas delineated on Exposure Map B include locations both on and off the Nyanza site property. Analysis including study participants who reported ever spending time or playing at only the areas within the Nyanza site property boundaries (i.e., Areas A, B, D, and E) showed a small increased relative risk for self-reported cancer diagnosis (RR=1.19) but again the results were not statistically significant (p=0.482).

Because family history of cancer was considered to be a potential confounding factor to the association between Nyanza site contact and a self-reported cancer diagnosis, a stratified analysis of contact with areas of Map B with respect to a family history was performed. These analyses resulted in statistically significant associations for playing or spending time in three areas associated with the Nyanza Chemical site. These areas included Area B (the Eastern Wetlands), Area F (the Sudbury River, near High Street

Bridge), and Area G (Sudbury River, Mill Pond, baseball field) (*Table 3*). Playing or spending time in these areas for those who reported a family history of cancer resulted in an increased risk of cancer that was nearly twice the risk for study participants who reported no family history of cancer (RRs = 1.80 to 1.96). A multiple logistic regression analysis was also performed including each Map B area and a family history of cancer. The overall combined effect of these factors was not significantly associated with a self-reported cancer diagnosis (Wald x^2 p=0.114), nor were the individual effects of any of the Map B Areas A through I. However, the results for family history of cancer did remain statistically significant with estimates of a self-reported cancer risk from the logistic regression analysis comparable to the results of the categorical analysis of family history of cancer and self-reported cancer (OR=1.78, 95% CI=1.08-2.91).

In addition, because one of the principle concerns of this study was contact with contaminants from the Nyanza site, an analysis was conducted to assess the relationship between spending time in specified areas on Map B that were considered on-site versus those where contamination was detected but were located off-site from the actual Nyanza property. Areas A, B, D, and E depicted on Map B were defined as the areas that compose the actual Nyanza site property. Study participants that responded as having spent time in any of these areas were considered as one exposure group (on-site), while those that did not respond to having spent time in any of these areas were considered as a separate exposure group (off-site). Analysis of these two exposure groups stratified by family history of cancer did not result in a significant relationship with a self-report of cancer diagnosis (*Tables 2 and 3*).

In addition to Areas A through I depicted on Map B, a categorical analysis was conducted to measure the associations between contact with specified areas within Areas A-I and an outcome of self-reported cancer diagnosis. These questions independently address specific smaller subsets of the larger general areas on Map B (*List 2*) and areas of the Nyanza site that were known as the most highly contaminated. For example, Chemical Brook is a specific area of known contamination located on the Nyanza Chemical site but is located within Area A. Also, there were two waste ponds located on the top of

Megunko Hill (Area D) and responses to questions about contact with Area D are not necessarily specific to the two ponds alone but the entire Megunko Hill area. These analyses showed that contact with the Megunko Hill ponds resulted in a nearly statistically significant relative risk of a self-reported cancer diagnosis [RR=1.84 (95% CI: 0.99 to 3.43)]. The x^2 p-value for these analyses was 0.054 (*Table 4*). A logistic regression model including all specified areas within the labeled areas depicted on exposure Map B and a self-report of cancer yielded no statistically significant results.

In relation to playing or engaging in any type of activity in these more specific exposure areas, additional questions were asked regarding discoloration of skin or clothing, skin irritations or rashes, and the presence of chemical drums in each area. Categorical table analysis was conducted with each exposure and the outcome of self-report of cancer diagnosis (*Tables 5-7*). The only significant association was a positive relationship between a report of discolored skin or clothing after spending time in the woods behind the Ashland High School playing fields and a self-reported cancer diagnosis. Although the number of exposed individuals who reported a cancer diagnosis was small (n=4), the fisher's exact test yielded a p-value of 0.0146 and a relative risk of 4.21 (95% CI: 1.71 to 10.38). However, these results should be interpreted with caution as the statistics are somewhat unstable given the small number of individuals who reported this type of exposure and the width of the 95% confidence interval. See *Table 5* for a summary of these results. The corresponding analyses for skin irritation or rashes and the presence of chemical drums in the woods behind the Ashland High School playing fields showed no statistically significant associations among individuals who self-reported a cancer diagnosis (*Table 6*). The study questionnaire also inquired whether individuals touched any chemical drums in specified areas of the Nyanza site. However, the frequency of response to these questions was too low to provide any meaningful statistical analysis.

It was reported that disposal of waste from the Nyanza Chemical site occurred at the Howe Street Dump (MDPH 1994). This was an area where teens and young adults were known to frequent and engaged in activities such as motor bike riding. A categorical analysis was also conducted for any contact with the Howe Street dump as the exposure

variable and a self-reported cancer diagnosis as the outcome. These results were not statistically significant (RR=1.24, 95%CI=0.75-2.07). The frequency of activity at the dump was also assessed. The frequency distribution was equally split into quartiles. Because the sample sizes for the upper 75% were small, this portion of the response distribution was designated the more-exposed group, while the lowest 25% was designated the less-exposed group. Analysis of the frequency of play at the Howe Street dump did not result in any statistically significant associations with a self-reported cancer diagnosis.

The next level of assessing exposure to the Nyanza site involved an inquiry regarding particular activities in each area on Map B. The only exposure activity resulting in a statistically significant elevated relative risk for a self-reported cancer diagnosis was swimming or wading in Area D with a chi-square p-value of 0.0236 and a relative risk of 2.48 (95% CI: 1.13 to 5.44) (*Table 8*). Area D represents the Megunko Hill area. A stratified table analysis of swimming was conducted to assess the affect of a family history of cancer on the observed statistically significant association with a self-reported cancer diagnosis. For study participants with a family history of cancer, the association between swimming in Area D and a self-reported cancer diagnosis was significantly strengthened to a relative risk of 3.03 (95% CI: 1.30-7.08). In contrast, for those with no family history of cancer, the association weakened (RR=1.11) and was no longer significant (*Table 8*). This pattern indicates that family history of cancer is acting as an effect modifier to the association between swimming in Area D and a self report of cancer diagnosis.

Further analysis resulted in statistically significantly increased risk of cancer for swimming or wading in very specific locations of the defined Areas A-I on Map B. These areas include the waste ponds located on the top of Megunko Hill, the ponds located near the Nyanza Company buildings, and the wetlands area behind Cherry Street (*Table 9*). The increased risks were markedly greater for swimming exposures in these areas and were further increased for study participants who reported a family history of cancer with relative risks ranging from 5.55 to 14.46.

Activities that did not result in a statistically significant association included playing ice hockey, ice skating, fishing, building forts, hunting, catching turtles or frogs, hanging out with friends, riding a bike, running, hiking or walking, or doing any other activity (*Tables 10-19*).

A more extensive analysis was conducted on swimming and fishing in lakes, rivers, and other natural bodies of water in Ashland. These areas included the Sudbury River, the Mill Pond, the Raceway, High Street Bridge area, the Ashland Reservoir, and the Framingham Reservoir. No statistically significant associations were seen with either swimming or fishing in these areas and a self-reported cancer diagnosis (*Tables 20, 21*).

Fish consumption habits were also assessed with respect to the natural bodies of water mentioned above. The frequencies of fish consumption from these areas were analyzed both categorically and through logistic regression. The analysis was conducted with two different exposure variables consisting of, ever/never consumed fish from these areas and high/low consumption of fish from these areas. Both exposure variables were created by grouping differently the frequency category response options of: 3 or more times a week; 1-2 times a week; 1-3 times a month; less than once a month; and never consumed fish. In addition, a separate logistic regression model was used to analyze frequency of fish consumption for each area. None of the analyses resulted in statistically significant associations, the majority of relative risks were less than 1.0 (*Tables 22, 23*).

Categorical analyses of participation in any sports, including intramurals, at Ashland High School and a self-report of cancer diagnosis did not result in statistically significant associations (*Table 24*). Questions were asked about particular sports played, the duration of participation, and locations of activities corresponding to the areas on Map B. As expected, Area C (Ashland High School and associated fields) and Area I (Stone Park) were the most frequented areas, as they are associated with playing sports, however none of the additional analyses on sports activities in these locations resulted in any statistically significant associations. Because Areas A, B, D, and E are contained within

the Nyanza site property, analyses were conducted for study participants that indicated practicing any sports in these areas, at any time. Individuals in this group were considered as the on-site exposure group, and compared to study participants that did not practice any sports in these areas. Categorical analysis of this exposure did not result in a significant relationship, with only 23 study participants who reported sport activities or practice at areas within the Nyanza property.

In addition, use of the path behind Ashland High School was assessed for study participants that participated in sports, as well as for the entire cohort (*Table 24*). Categorical analysis of this exposure did not result in any statistically significant associations with an outcome of self-reported cancer diagnosis. Additional frequency and duration information was also collected in relation to the use of the path; however none of the additional analyses on these variables resulted in statistically significant associations.

A categorical table analysis of participation in community organizations in Ashland (e.g., Girl or Boy Scouts, Little League, community sports, etc.) and a self-report of cancer diagnosis also did not result in a statistically significant association (*Table 24*). Additional questions assessed the locations labeled on Map B at which these organizations met; however none of the additional analyses on these variables resulted in any statistically significant associations. Again, because the main concern of the study was contact with the Nyanza site, a binary exposure variable was created with respect to the areas surrounding the Nyanza site at which these organizations regularly met. Because areas A, B, D, and E comprise the Nyanza site, those that indicated having met in any of these areas, for any organization listed, were considered in the on-site exposure group, while those that did not were considered in the off-site exposure group. These analyses did not result in a significant relationship with a self-reported cancer diagnosis.

The above analyses examining exposures related to contact with Nyanza site areas depicted on Map B and activities in these areas were repeated restricting the case group to only individuals for which a cancer diagnosis was confirmed through medical record

review or the Massachusetts Cancer Registry (n=40). By restricting the case group to only those with a confirmed cancer diagnosis rather than including all study participants with a self-reported cancer diagnosis (n=73), much of the cancer types considered to be not associated with an environmental exposure were excluded from the analysis (e.g., skin cancer and cervical cancer. The distribution of cancer types among study participants with a confirmed cancer diagnosis is provided in Figure 6. Therefore, the analyses restricting the case group to only those with a confirmed cancer diagnosis provide a more targeted assessment of Nyanza site exposure and its effect on the incidence of cancer in this study population.

Analyses with the case group restricted to only study participants with a confirmed cancer diagnosis showed positive associations among study participants who reported contact with Areas A, D, E and I. Similar to results of analyses that included all self-reported cancer diagnoses as the case group, the relative risks were generally small and were not statistically significant (*Table 25*). Further, the relative risks for study participants who reported contact with any of the remaining five areas depicted on Map B did not indicate an increase risk in cancer among study participants with a medically confirmed cancer diagnosis (*Table 25*). Contrary to the results observed for all individuals with a self-reported cancer diagnosis, when the analyses considered a family history of cancer, the association between Nyanza site contact for Areas A-I and confirmed cancer diagnosis were increased but were none were statistically significant. Further, the relative risks of a confirmed cancer among study participants who reported a family history of cancer compared to study participants with no family history of cancer were not measurably different (*Tables 26*).

2. Water Contact Exposure Analysis

The statistically significant association observed between swimming or wading and the self-reported cancer diagnosis group in the initial activities analysis prompted further exploration into exposure involving external contact with water. Because both swimming or wading and fishing involve water contact and none of the land-based activities nor

consumption of fish showed significant positive results, the responses to swimming or wading and fishing were combined to form a cumulative overall water-contact exposure variable. This binary variable assesses contact with water through swimming or wading and/or fishing. It does not distinguish between those that have answered positively to both swimming and fishing and those that have answered positively to one or the other activity per area. Further analysis was conducted in order to measure the association between overall water exposure and a self-reported cancer diagnosis for each Areas A-I on Map B.

The results for overall water-contact exposure to Areas A-I on Map B and self-reported cancer diagnosis confirm those previously observed for the individual analyses on swimming in Areas A-I. Where previous results for swimming or wading on Megunko Hill (Area D) showed a statistically significant relative risk of 2.48 (95% CI: 1.13 to 5.44), the association between overall water-contact in the Megunko Hill area and a self-reported cancer diagnosis remained the only statistically significant association with a similar relative risk of 2.43 (95% CI: 1.11 to 5.33) (*Table 39*). When stratified by a family history of cancer, this association strengthened to a relative risk of 3.03 (95%CI: 1.30 to 7.08). The non-family history of cancer stratum showed no significant association (*Table 40*). This supports the previous suggestion that family history of cancer is acting as an effect modifier to the association between water-contact exposure and a self-report of cancer diagnosis. In addition, these results further support the dominance of swimming or wading on the association between overall water-contact exposure and a self-reported cancer diagnosis and that fishing is likely only a negligible contribution, if any at all.

Analyses of overall water contact in any areas depicted on Map B were repeated restricting the case group to only those individuals with a confirmed cancer diagnosis. The results produced similar results to those previously described above. A statistically significant association was observed between study participants who reported any water contact with Area H (an area of the Sudbury River near Myrtle Street) and individuals with a confirmed cancer diagnosis (*Table 41*). These results showed that the relative risk

of confirmed cancer among individuals who had any type of water contact with Area H was slightly more than twice the risk than those who reported no type of water contact in this area (RR=2.07, 95% CI: 1.08-3.99). Previous analyses of Area H that included all individuals who self-reported a cancer diagnosis showed increased relative risks that approached statistical significance. As was also seen in the analyses including all self-reported cancer diagnoses as a group, the relative risk increased and remained statistically significant for those who had any water contact with Area H and reported a family history of cancer (RR=3.63, 95% CI 1.47-8.96).

Although virtually no statistically significant associations were observed for water contact with general areas of the Nyanza Site (Areas A through I) as defined on exposure Map B and a confirmed cancer diagnoses, significant associations occurred between overall water-contact exposures at Area H, the Sudbury River area near Myrtle Street (*Table 41*). These results were statistically significant among individuals who had a confirmed cancer diagnosis and remained statistically significant when a family history of cancer was considered in the analyses.

Further examination of the associations between overall water-contact exposure and Nyanza site areas was conducted restricting the case group to those individuals in the study population whose cancer diagnosis was considered a rare cancer. As stated earlier, 33% (n=25) of the individuals who reported a cancer diagnosis were defined as having a rare cancer. These analyses were consistent with the results obtained for overall water-contact exposures at the Nyanza site and a confirmed cancer diagnosis. Again, statistically significant associations occurred for exposures at the Nyanza site involving water contact at Chemical Brook and the two waste lagoons located on Megunko Hill. These results remained statistically significant when a family history of cancer was considered in the analyses.

3. Occupational History Analysis

The questionnaire assessed occupational histories with specified sites and tasks including: working at the Nyanza Chemical Company; working at the Derby Chemical Company;

jobs involving agriculture or woodworking; and any job that brought respondent in contact with areas labeled on Map B. Categorical analysis for each of these employment variables did not result in any significant associations with a self-reported cancer diagnosis (*Table 43*). However, the number of study participants who responded positively to questions about working at Nyanza Chemical Company or Derby Chemical Company was relatively small (n=12).

Information was also collected regarding specific materials used at least once a week at a job, inside, or outside the home. (See *List 3* for a listing of occupational materials). Categorical table analyses for exposure to each material and an outcome of a self-report of cancer diagnosis were conducted (*Table 44*). Analysis for paints, thinners, removers (paint products) and self-reported cancer diagnosis resulted in a significant negative association with a p-value of 0.0076 and relative risk of 0.32 (95% CI: 0.13 to 0.79). Although this negative association implies a protective effect of paint products and a self-reported cancer diagnosis, it is likely a reflection of this sample of participants, rather than a protective effect of the materials.

A multiple logistic regression analysis was conducted for all materials and a self-reported cancer diagnosis to explore the effects of each occupational material while controlling for those of the other materials in the model. Again, paint products remained statistically significant with a p-value of 0.0197 and a relative risk of 0.26 (95% CI: 0.09 to 0.81). The association with chlorinated chemicals became statistically significant in the regression analysis with a p-value of 0.0032 and relative risk of 2.99 (95% CI: 1.44 to 6.18).

In order to control for possible confounding, table analyses for paint products, chlorinated chemicals, chemical dyes and a self-reported cancer diagnosis were individually stratified by family history of cancer (*Table 45*). The relationships between paint products, chemical dyes and a self-reported cancer diagnosis were not influenced by family history of cancer. Alternatively, family history of cancer is likely an effect modifier to the relationship between chlorinated chemicals and self-reported cancer diagnosis, as the

estimated relative risk for those with a family history of cancer exceeded the value of the crude relative risk, while the relative risk for those with no family history of cancer negative stratum approximated the null value of 1.00. These results were confirmed through logistic regression analysis. In addition, information was also collected on the duration and frequency of use of these materials in an occupational setting. Because these variables are continuous, they were individually analyzed through logistic regression analysis. This analysis showed a small yet statistically significant increase in risk of self-reported cancer diagnosis with increased duration of use (years) of chlorinated chemicals only with a p-value of 0.0098 and an estimated odds ratio of 1.09 (95% CI: 1.02 to 1.16) while controlling for family history of cancer.

Because table analyses of chlorinated chemicals and chemical dyes resulted in nearly statistically significant positive associations and both were prominent at the Nyanza site, a binary cumulative chemical exposure variable was created by combining these responses. This variable assesses cumulative chemical exposure through chlorinated chemicals *and/or* chemical dyes; however it does not distinguish between exposures to only one or to both. Categorical table analysis of this variable and a self-reported cancer diagnosis resulted in a statistically significant association with a p-value of 0.0083 and a relative risk of 2.05 (95% CI: 1.20 to 3.49) (*Table 44*). Stratification by family history of cancer again indicated that family history is acting as an effect modifier to the relationship between cumulative chemical exposure and self-reported cancer diagnosis, with results similar to those for chemical dyes alone (*Table 45*).

4. Occupational and Water-Contact Exposure Analysis

Chlorinated chemicals and chemical dyes are highly utilized materials in the chemical dye manufacturing industry and have been linked to cancer development, therefore the significant positive associations observed in the previous analyses warranted further investigation. In particular, because the results for chlorinated chemicals and cumulative chemical exposure with a self-reported cancer diagnosis indicate effect modification by a family history of cancer, it may also be important to control for exposure to chemicals when conducting additional water-contact exposure analyses.

A logistic regression analysis was conducted with a model that included cumulative chemical exposure, water-contact exposure in the Megunko Hill area (Area D), and family history of cancer as covariates and a dependent variable of self-reported cancer diagnosis. This is an important analysis, as each exposure variable has been shown to be significantly associated with a self-reported cancer diagnosis, as well as interacting with or being modified by a family history of cancer. Therefore, this analysis described the effects of each independent exposure variable on the dependent variable, while controlling for the other covariates in the model. The overall model was significant with a chi-square p-value of 0.007 and the individual effects of the covariates remained significant with odds ratio estimates comparable to the results from their respective unstratified categorical analyses (*Table 46*). The chi-square p-value for overall water-contact exposure in Area D was 0.053.

These analyses confirm that there is a statistically significant positive association between overall water-contact exposure in Area D and a self-report of cancer diagnosis. Exposure includes swimming or wading and fishing. However, the exposure is dominated by swimming or wading, while fishing appears to have little influence on the observed associations. More specifically, those that reported water-contact in the Megunko Hill area (Area D) and the ponds on Megunko Hill had approximately 2.5 times greater risk of self-reporting a cancer diagnosis than those that did not report these exposures. Also confirmed is the statistically significant positive association between occupational exposure to chemicals, which includes chlorinated chemicals and chemical dyes, and a self-report of cancer diagnosis. More specifically, study participants that reported using chlorinated chemicals and/or chemical dyes at least once a week on a job had approximately twice the risk of a self-reported a cancer diagnosis than those that did not report this exposure. These associations remain on despite the influence of a family history of cancer, which is also significantly associated with a self-reported cancer diagnosis with a relative risk of approximately 1.8 (95% CI: 1.1 to 2.9).

5. Agricultural and Pest Management Analysis

Exposure to various forms of agricultural treatment and pest management were assessed as possible risk factors for self-reported cancer diagnosis. Each set of questions included a dichotomous ever/never response, the duration in years, and frequency in specified times per year. Exposures consisted of pesticide or insecticide spraying inside the home, insecticide or herbicide treatment on the yard, tree or agricultural spraying at or near the home, and mosquito spraying or fogging at or near the home. Categorical table analyses of the dichotomous ever/never exposure variables with a self-report of cancer diagnosis, unstratified and stratified by family history of cancer, did not result in any significant associations (*Tables 47, 48*).

For those ever/never exposure variables to which participants responded positively, the number of years of exposure was analyzed using multiple logistic regression. A separate analysis was conducted for each variable with only itself and family history of cancer included as covariates in the model (i.e., four separate models). Only exposure to tree or agricultural spraying resulted in a significant association with a chi-square p-value of 0.012 and an estimated odds ratio of 1.05 (95% CI: 1.01 to 1.09). While statistically significant, the estimated odds ratio, is very small and does not indicate an appreciable increase in risk of self-reported cancer diagnosis with each year that tree or agricultural spraying took place at or near the home *(Table 49*).

Cumulative chemical exposure and water-contact exposure in Area D were then included in the logistic regression model with each exposure variables for tree and agricultural spraying as well as a family history of cancer. Similar to previous results, exposure to tree or agricultural spraying remained significant with a p-value of 0.0089 and an estimated odds ratio of 1.06 (95% CI: 1.01 to 1.10). In addition, water-contact exposure in Area D was highly significant with a p-value of 0.0006 and an estimated odds ratio of 12.16 (95% CI: 2.90 to 51.01) (*Table 50*). The odds ratio is elevated and the confidence intervals wider than in previous analyses, as this analysis included only a subset f the data and only those participants who responded positively to the ever/never occurrence of tree

or agricultural spraying at or near the home were included. Nonetheless, these results indicate that although exposure to tree or agricultural spraying is statistically significant, the risk of a self-reported cancer diagnosis is low, and water-contact exposure to Area D (Megunko Hill) remained the most influential factor in the model in determining risk for a self-report of cancer diagnosis.

Individual regression analyses of the frequency variables for each exposure associated with tree and agricultural spraying with a self-report of cancer diagnosis, while controlling for family history of cancer, did not result in any significant associations (*Table 51*). A cumulative exposure variable was created by multiplying the duration and the frequency variables for each agricultural exposure. Logistic regression analysis of the cumulative variable for tree and agricultural spraying with family history of cancer, cumulative chemical exposure, and water-contact exposure in Area D resulted in exposure to tree or agricultural spraying only slightly elevating the risk of a self-report of cancer diagnosis (OR=1.04, 95%CI: 1.01 to 1.07), while controlling for the other variables in the model (*Table 52*). This analysis also re-emphasizes that water-contact exposure in Area D, including swimming or wading and fishing, remains the most significant risk factor for a self-report of cancer diagnosis (OR=12.7, 95%CI: 3.08 to 52.18).

6. Descriptive Residential Analysis

The residential history analysis is based on an ever/never exposure to a defined exposure area surrounding the Nyanza site. Exposure to this area was determined by spatially mapping each of the addresses at which study participants lived for at least 2 years. A participant was considered ‘ever exposed’ if any of their reported addresses fell within the designated exposure region.

There were 1384 participants out of the 1387 in the cohort for whom at least some part of residential history could be spatially mapped. Out of this 1384, there were 15 participants who provided a partially incomplete residential history with one or more addresses that could not be mapped and exposure not determined. Therefore the total

number of participants for whom a residential history exposure analysis could be conducted was 1369. The total number of participants having ever resided in the designated exposed region was 125/1369; therefore analysis on the duration of exposure is only possible for this subset of 125 participants. Exposure ranged between 2 and 28 years with the mean, median, and mode at 9.06, 7.00, and 2.00 years, respectively.

7. Residential Exposure Analysis

Ever/never residential exposure to the designated exposure region was analyzed categorically with the outcome variable of a self-reported cancer diagnosis, unstratified and stratified by both family history of cancer and gender (*Tables 53 & 54*). In addition, logistic regression analysis was used to assess the potentially confounding affects of family history of cancer, gender, and study age (*Table 55*). There were no statistically significant associations between having ever lived in the designated "exposed" region surrounding the Nyanza site and a self-reported cancer diagnosis. Neither stratification nor the inclusion of various combinations of covariates in the logistic regression models influenced this result; however the latter did reiterate the significant association between family history of cancer and a self-reported cancer diagnosis.

A continuous variable representing cumulative duration of years lived in and out of the exposed area was created by summing the number of years spent at each address. The number of years of residential exposure was analyzed using logistic regression with a self-reported cancer diagnosis as the outcome of interest. Various models were used to assess the potentially confounding affects of family history of cancer, gender, and study age (*Table 37*). There were no statistically significant associations between the duration of time lived in the designated exposure area and a self-reported cancer diagnosis.

V. Discussion

The results of the analyses indicate that although most study participants reported either playing or spending time at some area either on or off the Nyanza site property as depicted in Exposure Map B, that only a small percentage of individuals engaged in

activities associated with an increased risk of developing cancer (i.e., exposures involving swimming/wading or fishing). Even though a positive relative risk of a self-reported cancer diagnosis was observed for study participants who reported spending time in most exposures Areas labeled A through I and depicted on Map B, the risks were not statistically significant, indicating that there was no difference in cancer risk among study participants who reported spending time in these areas as compared to those who did not spend time in these areas.

However, when a family history of cancer was considered in the analyses, statistically significant increased risks were observed for three areas delineated on the exposure map (Areas B, F and G). Individuals who reported both a family history of cancer and contact with these areas of the Nyanza site had nearly twice the risk of cancer than study participants who did not report contact with these areas. Even though only one area (Area B) is located within the actual Nyanza site property bounds, all of the areas (Areas B, F, and G) contain wetlands or are associated with the Sudbury River. Therefore, exposure in these areas would have primarily involved some type of water contact.

Among study participants who reported no family history of cancer, there was no association with exposure to areas on the Nyanza site. When specific exposure points located within the delineated exposure areas on Map B were considered, contact with Megunko Hill and specifically the waste ponds located on Megunko Hill, resulted in a nearly statistically significant association with having a self-reported cancer diagnosis. This area was one of the most highly contaminated areas of the Nyanza Chemical site. Of note is that study participants who reported skin irritation or rash after contact with this area of the site, experienced a four-fold relative risk of having a cancer diagnosis (RR=4.02, p=0.2280). However, the number of individuals who reported contact with this area was relatively small and the association was not statistically significant. A similar result was observed for study participants who reported contact with the woods behind Ashland High School. Although no increased risk of a self-reported cancer diagnosis was observed among study participants who reported general contact with the woods behind Ashland High School, individuals who reported discolored skin or clothing

after contact in this area had a statistically significant increased risk of cancer (RR=4.21, p=.0146).

Repeating the analyses restricting the case group to only study participants with a confirmed cancer diagnosis showed positive associations with the majority of exposure areas described on Map B. However, unlike the results observed with self-reported cancer diagnosis as the case group, the results for the individual exposure Areas A-I and a confirmed cancer diagnosis were not statistically significant when the analyses considered a family history of cancer. Analysis of overall water-contact exposures in Areas A-I of the Nyanza site again did not result in statistically significant associations with a confirmed cancer diagnosis. But these analyses did show a statistically significant association between Area H (an area of the Sudbury River near Myrtle Street) and individuals with a confirmed cancer diagnosis. The relative risk among individuals who reported any type of water-contact (i.e., swim/wade or fish) with Area H was slightly more than twice the risk compared to individuals with no contact with this area of the Nyanza site (RR=2.07, 95% CI 1.08-3.99). Previous analyses of this area that included all individuals who self-reported a cancer diagnosis showed results that approached statistical significance (e.g., lower bound 95% CI of 0.92, 0.94 etc.). As was also seen in the analyses including all self-reported cancer diagnoses as a group, the relative risk increased and remained statistically significant for those who had any water-contact exposure with Area H and reported a family history of cancer (RR=3.63, 95% CI 1.47-8.96).

Although no statistically significant associations were observed for contact with general areas of the Nyanza site (i.e, Areas A through I) defined on exposure Map B and confirmed cancer diagnoses, when specific exposure locations within Areas A-I were evaluated, significant associations occurred between overall water-contact exposures at Chemical Brook and the two waste lagoons located on Megunko Hill. These results were statistically significant among individuals who had a confirmed cancer diagnosis as well as those individuals whose diagnosis was considered a rare cancer. Again, the results were confirmed and remained statistically significant when a family history of cancer was

considered in the analyses. Therefore, even though the analyses restricting the case group to study participants who had a confirmed cancer diagnosis did not consistently confirm the associations initially observed between areas defined on Map B and self-reported cancer diagnoses, the results for overall water-contact exposures at the site were consistent for specific site locations (Chemical Brook and Megunko Hill ponds).

Although five individuals in the case group were diagnosed with various types of sarcoma, contrary to information initially reported by the Ashland community, the MDPH was unable to confirm a cluster of five young men of relatively similar age diagnosed with this type of cancer among the study group. What is interesting to note is that the case group as a whole demonstrated an atypical pattern of cancer diagnoses in that 15% of the individuals were diagnosed with cancers that could be considered either a rare cancer type or rare for the age at which the diagnosis occurred. The age range of the cases at diagnosis was 14 to 46 years and 66% of the individuals in the case group were under the age of 35 at the time of their diagnosis.

Further examination of the associations between overall water-contact and Nyanza site areas restricting the case group to those individuals in the study population whose cancer diagnosis was considered a rare cancer showed results consistent with those obtained for overall water-contact exposures at the Nyanza Site and a confirmed cancer diagnosis. Again, statistically significant associations occurred for exposures at the Nyanza site involving any water contact at Chemical Brook and the two waste lagoons located on Megunko Hill. Again, these results remained statistically significant when a family history of cancer was considered in the analyses.

By restricting the case group to only those with a confirmed cancer diagnosis or a rare cancer diagnosis rather than including all study participants with a self-reported cancer diagnosis is important because much of the cancer types considered as not associated with an environmental exposure (e.g., skin cancer and cervical cancer) were excluded from these analyses. Therefore, the subset analyses restricting the case group provide a more targeted assessment of Nyanza site exposure and its effect on the incidence of

cancer in this study population. Taken as whole, the findings suggest that a possible gene-environment interaction could be occurring for individuals in the study population who reported exposures involving any type of water contact and have a family history of cancer.

The study had a 67% response rate but approximately 40% of the overall estimated target cohort was lost to follow up. That is, these individuals either did not respond to recruitment efforts or we were unable to locate and contact them to request study participation. Although lost to follow-up can be a concern to the validity of a study such as this, it is unlikely that selection bias had a strong affect on study validity. Selection bias exists if the observed association among the actual study participants is systematically different from the association obtainable from those who are eligible but were either excluded or did not participate/withdrew from the study. Generally, loss to follow-up will only introduce selection bias if it depends on the disease status among the exposed or non-exposed population. If there is equal probability of loss to follow-up or it only varies by exposure status or only varies by disease status then no selection bias will occur.

Due to the historical nature of the Nyanza Chemical Waste site, the exposure assessment for the Ashland Nyanza Health Study was limited. That is, prior to 1982, when the Nyanza Chemical Waste site was closed and investigations into to the extent and nature of contaminants present at the site took place, little or no data existed that could be used to empirically measure exposure from site contaminants. Therefore, it was necessary to estimate exposure based on reported information from study participants about the nature and types activities they engaged in at the site and pair the collected data with information about locations at the site where contaminants were known to have been present.

In any observational study, some degree of inaccuracy in assessing exposure is inevitable. Although using exposure estimates, as was the case in the Ashland Nyanza Health Study, likely introduced some misclassification to the study, the MDPH believes that this did not

introduce a high potential for error. Given that the study design was a retrospective cohort and the exposure assessment relied on self-report from study participants of activities at the Nyanza site, the most likely source of error was introduced in the form of recall bias. Recall bias is a systematic error that can occur in studies due to differences in the accuracy or completeness of recall to memory of prior events or experiences (Last 1988). Because of the amount of time that has passed, it is difficult for the study participants to provide information regarding the exact location and frequency of activities they may have participated in the past.

This type of information bias is more of a concern to the validity of the study if it results in differential misclassification of exposure. That is, if exposure is classified differently as a result of memory recall for individuals in the study who reported a cancer diagnosis as opposed to those who did not, the association between exposure and disease may be exaggerated or overestimated. This is because those who have a diagnosis of cancer may tend to remember or recall events in their past more specifically and attribute these events to their diagnosis. Therefore, these study participants are disproportionately classified as "exposed" because they tend to recall exposure events more specifically than those with no cancer diagnosis. However, if the exposure is classified incorrectly for equal proportions of study participants in the two groups being compared, then this type bias results in non-differential misclassification. That is, if the exposure actually increases the risk of disease then non-differential misclassification would influence the observed risk estimate towards the null value or no association between exposure and disease. Non-differential misclassification bias therefore is a lesser concern to the study validity because the bias is always in the direction of an underestimate of the association between exposure and disease. This is because the estimate of the effect without the misclassification would always be greater than that observed. Some may argue that poor exposure data (or poor disease classification) invalidates study results, but this argument is incorrect if the results indicate a nonzero effect (Rothman 1986). In the Ashland Nyanza Health Study differential misclassification bias is probably not an issue because the study results overall did not show a strong effect of exposure on the incidence of

cancer and the impact of differential misclassification would only tend to overestimate the true effect of exposure on disease.

The MDPH attempted to minimize any effect this type of bias may have had by using consistent and replicable methods to assess exposure. Standard and replicable methods were used in the exposure assessment for the study. Two maps were provided to all study participants. The maps provide a standard source of information from which study participants could respond to questions about locations and activities they engaged in at the Nyanza site. A standard script was used during the interview process to allow study participants to familiarize themselves with the map locations and orient themselves to specific exposure areas defined on the map. Therefore, although study participants were asked to recall events that occurred possibly 20 to 40 years in the past, because standard methods and prompts were used, memory recall should be similar among both the exposed and unexposed groups. Also, during the study period of interest (1960s to 1980s), the Nyanza site was a large uncontrolled hazardous waste site. If exposures occurred, conditions at the site during this period of time were such that study participants would have specific recollections of exposure events. For example, Chemical Brook was an intermittent stream where the Nyanza Company discharged waste from its dye manufacturing processes. Chemical Brook was also called Purple Brook for the fact that that the brook would change colors according to the chemical dye processes that took place at the Nyanza facility. Therefore, if a study participant had fallen into Chemical Brook as a young child or teen and emerged with discolored skin or clothes, they were likely to have remembered this type of event with certainty even though it may have occurred 20 years prior. These were the types of exposure events that were reported frequently at the Nyanza site.

Map B (the exposure map used in the study) was developed using a geographic information system based on historical aerial photographs of the site taken during the 1960s and 1970s. The Map was created to accurately depict historical features and landmarks of the Nyanza site during the time period of interest so as to minimize any potential misclassification of exposure. The specific site areas labeled as Areas A

through I on Map B were shown as discrete areas so that reports of activities or contact at theses sites would be with some degree of accuracy in terms of actual location. Thus any recall bias introduced to the study for the most part would then be non-differential because it likely occurred equally between the two groups (exposed and non-exposed). Again, this type of bias tends to underestimate the observed association between exposure and disease.

For the Ashland Nyanza Health Study any selection bias would be non-differential since both exposed and unexposed groups were applied the same criteria in recruitment and selection. In addition, most of the refusals were for reasons of time or interest and very few potential study participants indicated that their refusal was related to possible exposure status. Although most exposure areas of the Nyanza site showed no statistically significant association with an increased risk of cancer, the study had some limitations that could have affected its ability to detect stronger associations between site exposures and disease. The nature of contamination at the Nyanza Chemical Waste Dump is that of a complex mixture of contaminants in both on-site and off-site locations. Because little to no environmental sampling data exists that could be used to estimate exposures for study participants during the time when the site was operational, it was not possible to recreate more precise exposure estimates. Therefore, reliance of self-report of contact with general exposure areas on Map B as the primary assessment of exposure could yield imprecise exposure estimates and potential misclassification of exposure. However, any misclassification of exposure introduced to the study is likely to be non-differential between the exposed and non-exposed groups. Further, the study inquired about specific exposure activities and locations within the defined exposed areas A through K in an effort to more specifically define exposure among study participants and reduce any potential exposure misclassification.

The MDPH acknowledges that like all retrospective studies that given the historical nature of the Nyanza Chemical Waste Site, recall bias is an issue of some concern to the study. The area in which recall bias has the most impact on the study results is the potential for misclassification of exposure. This means that study participants who may

actually have been exposed as a result of contact with the Nyanza Chemical Site or participation in certain activities at the site do not recall these events accurately because of the length of time that has passed since the events occurred. So those who have actual exposure are then misclassified as unexposed due to poor memory of events in the past. It can also be argued that individuals with a cancer diagnosis are more likely to recall certain events in their life as related to their cancer diagnosis and therefore respond differently to questions about activities and contact with the Nyanza Chemical Site. Therefore, the bias due to recall leads to misclassification of the exposed and unexposed groups by incorrectly classifying exposure status.

While it is likely that recall bias exists and had some impact on the study results, the real issue is whether recall of events differed significantly between those with and without a cancer diagnosis. If the recall occurred evenly or in a comparable way among both groups then the bias is non-differential and would not have a large impact on the observed association. If the cases were more likely to report exposure than non-cases then the result would be a bias in the direction of a positive association (differential bias). The study used a standard method of assessing exposure for all study participants regardless of whether they had a diagnosis of cancer. A standard map was used as a memory prompt for participants to locate areas of the Nyanza Chemical site. These steps were taken to establish a standard method for assessing exposure for all study participants and therefore minimize any effect that recall bias may introduce to the study results. However, it is difficult to say with certainty whether recall bias resulted in differential misclassification of exposure. Viewing the study results in their entirety suggests that any impact of recall bias was not in biasing the observed associations towards an increased risk of cancer since most analyses did not result in a statistically significant association between Nyanza site exposure and an increased risk of cancer diagnosis.

VI. Conclusions and Recommendations

The findings of this study suggest that the relative risk of developing cancer was greater for study participants with some types of reported exposures in areas of the Nyanza

Chemical Waste Dump. Although no specific activity appeared to increase the risk, statistically significant associations were observed among study participants who reported exposures involving any type of water contact in the Megunko Hill area, specifically the ponds located at Megunko Hill as well as Chemical Brook and certain areas of the Sudbury River (Area H on Map B). Further, the relative risks were markedly increased among study participants who reported a family history of cancer. The results for reported exposures with any type of water contact were observed in areas both on the Nyanza property itself as well as locations where contamination associated with the Nyanza site was detected not located on the Nyanza property. These findings showed a consistent pattern when the analyses included all study participants who self-reported a cancer diagnosis as the case group, only those individuals with a confirmed cancer diagnosis as the case group, and individuals with a rare cancer type. These findings considered as a whole are suggestive of a possible gene-environment interaction and between Nyanza site exposures involving water contact and study participants with a family history of cancer.

The MDPH acknowledges that there are limitations to the exposure assessment portion of the study given the retrospective nature of the study and therefore the necessary reliance on proxy data to estimate individual exposures. As a result, it is entirely possible that some misclassification bias was introduced to the study due to poor recall of exposure activities. However, the MDPH attempted to minimize any potential error in exposure measurement by quality control in the design of the study questionnaire, interview procedures, and standardization of response. Despite this and because of the historical nature of exposures at the Nyanza site and the combined mixtures of contaminants that were present at the site, the study is limited in its ability to draw conclusions about cause and effect relationships between specific contaminants and risks of cancer among this population.

As part of the remedial plan for the Nyanza site, the USEPA excavated and landfilled contaminated sediments from the Eastern Wetland, the Trolley Brook, Outfall Creek and the Lower Raceway. These areas correspond to the study exposure Areas B (Eastern

Wetlands and Trolley Brook Wetlands) and portions of Area A (Outfall Creek) and Area H (Lower Raceway). Therefore current exposures to these areas are not likely. Based on the study findings related to specific exposures and exposure areas (i.e., water contact exposures), the MDPH recommends that individuals consult with their medical provider about their individual cancer risk. This is particularly important for those with a family history of cancer, who engaged in activities that may have brought them in contact with the Eastern Wetlands of the Nyanza site and portions of the Sudbury River near Mill Pond and High Street during the 1960s through the 1980s when the Nyanza site was operational and unrestricted.

The USEPA is currently considering a more permanent remedial solution to reduce or eliminate risks posed by mercury contamination in the fish and sediments of the Sudbury River. The MDPH recommends that the USEPA considers the findings of this study in their human health risk assessment for the site.

VII. References

Ashland, Town of. 2004. Department of Housing and Community Development, Ashland Middlesex County. Available at: http:\www.ashlandmass.com\Town Hall\planning\profile.html

Costani G, Rabitti P, Mambrini A, Bai E, Berrino F. 2000. Soft tissue sarcomas in the general population living near a chemical plant in Northern Italy. Tumori. Se p-Oct 86(5): 381-3.

Eriksson M, Hardell L, Adami HO. 1990. Exposure to dioxins as a risk factor for soft tissue sarcoma: a population based case-control study. Journal of National Cancer Institute 82:486-490.

Ericksson M, Hardell L, Adami HO. 1988. Exposure to dioxins as a risk factor for soft tissue sarcoma: a population-based case-control study. Journal of the National Cancer Institute. 82(6): 486-90.

Fauss LM. 1989. Photogeologic analysis Nyanza Chemical Company Ashland, Massachusetts by L. Mike Fauss, Imagery Analyst, The Bionetics Corporation, Warrenton, VA. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Office of Research and Development, Las Vegas, NV. July 1989.

Hanson H, DeRosa CT, Pohl H, et al. 1998. Public health challenges posed by chemical mixtures. Environmental Health Perspectives 106 (Supp. 6): 1271-1280.

Hardell L, Eriksson M. 1988. The association between soft tissue sarcoma and exposure to phenoxyacetic acids or chlorophenols. British Journal of Cancer 54:819-823.

Hayes S. 2000. Personal communication (meeting with Sharon Hayes, Project Manager, Nyanza Chemical Dump, USEPA Region I, Boston, MA on January 4, 2000, regarding Ashland Nyanza Health Study exposure map and areas of Nyanza site contamination). U.S. Environmental Protection Agency, Boston, MA.

Hoppin JA, et al. 1999. Occupational risk factors for sarcoma subtypes. Epidemiology. 10(1): 300-306.

Jaratowicz G. 1952. Aerial photograph of Nyanza Chemical Dump with USEPA overlay from Ashland Public Library. supplied by Gary Jarotowicz, Tricil.

Kogan MD, Clapp RW. 1988. Soft tissue sarcoma mortality among Vietnam veterans in Massachusetts, 1972-1983. International Journal of Epidemiology 17:39-43.

Last JM (ed). 1988. A dictionary of epidemiology. 2nd ed. Edited for the International Epidemiological Association. New York: Oxford University Press.

MDEQE. 1980. Nyanza, Chemical waste dump, Megunco Road Ashland, Massachusetts preliminary site assessment report. October 23, 1980. Massachusetts Department of Environmental Quality Engineering, Boston, MA.

MDPH. 1988. Bladder and kidney cancer incidence and mortality, infant and neonate mortality in the town Ashland, Massachusetts. December 16, 1988. Massachusetts Department of Public Health, Community Assessment Unit, Boston, MA.

MDPH. 1990. Follow-up report on the occurrence of bladder and kidney cancer in Ashland, Massachusetts. December 5, 1990. Massachusetts Department of Public Health, Division of Environmental Health Assessment, Boston, MA.

MDPH. 1994. Public health assessment for the Nyanza Chemical Waste Dump, Ashland Massachusetts, CERCLIS NO. MAD990685422. June 29, 1994. Prepared by the Massachusetts Department of Public Health under a cooperative agreement with the Agency for Toxic Substances and Disease Registry, Boston, MA. .

NUS. 1992. Final remediation investigation report oversized figures volume iv of iv – Nyanza operable unit III – Sudbury River Study Middlesex County, Massachusetts. NUS Corporation. EPA work assignment no. 02-1L15. May 1992.

Rothman KJ. 1986. Modern epidemiology. Boston: Little, Brown and Company. p.83-9.

SAS. 2001. SAS Software, version 8.1. 1999-2001. SAS Institute Inc. Cary, NC.

SAS. 2003. SAS Software, version 9.1. 2002-2003. SAS Institute Inc. Cary, NC.

Schmidt, CW. 1999. A closer look at chemical exposures in children. Environmental Science and Technology 72A-75B.

SEER. 2004. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 9 Regs Public-Use, Nov 2003 Sub (1973-2002), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2005, based on the November 2004 submission.

USEPA. 2004. Nyanza Chemical Waste Dump, Ashland, Massachusetts. U.S. Environmental Protection Agency. Available at: http:\www.yosemite.epa.gov [cited Nov 7, 2004].

Stokely PM. 1982. Nyanza hazardous waste site Ashland, Massachusetts by Peter M. Stokely, Imagery Analyst, The Bionetics Corporation, Warrenton, VA. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Office of Research and Development, Las Vegas, NV. July 1989

Verity. 1999. Teleform Scanning Software, 1999. Verity Inc. 894 Ross Drive, Sunnyvale CA.

Zahm SH, Fraumeni JF. 1997. The epidemiology of soft tissue sarcoma. Seminars in Oncology. 24(5): 504-514.

Zahm SH, Blair A, Holmes FF, Boysen CD, Robel RJ. 1988. A case-referent study of soft tissue sarcoma and Hodgkin's disease, farming and insecticide use. Scan J Work Environ Health. 14(1988): 224-230.

Zahm SH, Blair A, Holmes FF, et al. 1989. A case-control study of soft-tissue sarcoma. American Journal of Epidemiology 130: 665-75.

Zahm SH, Tucker MA, Fraumeni JF Jr. 1996. Soft tissue sarcoma. In: Shottenfeld D, Fraumeni JF Jr. Cancer epidemiology and Prevention. New York: Saunders.

Zeliger, HI. 2003. Toxic effects of chemical mixtures. Arch Environ Health. 58(1): 23-9.

Zeliger, HI. 2004. Unexplained cancer clusters: common threads. Arch Environ Health. 59(4): 172-6.

Figures

Tables

Appendix A

External Peer Review Committee Comments and MDPH Response

Appendix B

Ashland Nyanza Health Study Participant Contact Letter