Epidemiological Approaches to Testing the Population Mixing Hypothesis

Summary Minutes

March 27, 2002
# TABLE OF CONTENTS

*March 27, 2002*

I. **Introduction and Overview of Meeting Objectives** .................................................................1

II. **Overview of Existing Data Relevant to the Population Mixing Hypothesis** ........2

   - Overview of existing data relevant to the Population Mixing Hypothesis — Malcolm Smith (for Randy Todd) .................................................................2
   - Studies testing the Population Mixing Hypothesis — Leo Kinlen .................................................3
   - Testing the Population Mixing Hypothesis for non-cancer chronic diseases — Graham Law ........4
   - Historical overview of leukemia cluster evaluations in the U.S.: relevance to Population-Mixing Hypothesis — Clark Heath ........................................5
   - Recent childhood leukemia cluster evaluations in California — Peggy Reynolds ..........................6
   - Population mixing and childhood leukemia in Canada — Will King ...........................................7

III. **Key Concepts and Open Discussion** .........................................................................................8

IV. **Infectious Disease Correlates of Population Mixing Hypothesis** ........................................9

   - Previous evaluations of specific infectious agents for childhood ALL — Malcolm Smith ...........9
   - Methods applicable to identifying infectious etiologies of cancers — Siobhan O’Connor ......10
   - Molecular epidemiology of population mixing — Martyn Smith ............................................10
   - Potential immunologic correlates of the Population Mixing Hypothesis — Steve Chanock .......11
   - Potential applications of GIS to evaluating the Population Mixing Hypothesis — Daniel Wartenberg ...

V. **Issues in Study Design for Testing the Population Mixing Hypothesis** ...............................14

VI. **Summary and Discussion of Action Items** .............................................................................16

VII. **Background References Related to Population Mixing** ......................................................16

   - Childhood Cancer Clusters and Clustering: .................................................................................16
   - In utero Initiation of Childhood ALL: .........................................................................................16
   - Population Mixing References: ......................................................................................................17
   - Population Mixing and Childhood Diabetes: ..............................................................................17
   - Possible Role of Infectious Agents in Childhood ALL Etiology: .................................................18

VIII. **Meeting Participants** ..............................................................................................................19
I. Introduction and Overview of Meeting Objectives

The objectives of the meeting were introduced by Drs. Tom Sinks and Malcolm Smith. This meeting directly followed a larger workshop held the two previous days, entitled “Gene-Environment Interactions in the Etiology of Childhood Cancer.” Dr. Sinks alluded to his presentation on cancer cluster investigations at this earlier workshop, then noted that a particular cluster of childhood acute lymphoblastic leukemia (ALL) in Fallon, Nevada, was one of the stimuli for scheduling the present meeting. Dr. Sinks observed that the town of Fallon and its associated naval air station possibly fulfilled the criteria for population mixing, and that an Expert Review Committee advising the Nevada State Health Department had listed “population mixing” as a possible contributor to the cluster.

Several questions were posed to the meeting participants by Drs. Sinks and Smith in relation to the population mixing hypothesis in general and more specifically to its possible applicability to the cluster of childhood leukemia cases in Fallon, Nevada. These questions included:

1) How can key concepts of the population mixing hypothesis be defined in a way that allows their evaluation in analytic studies?
   a) Defining “isolated population”
   b) Defining “susceptible persons”
   c) Defining “population mixing”

2) What biological studies relating to infectious agents might be used in testing the population mixing hypothesis or in understanding the underlying biological basis of the population mixing hypothesis?

3) What biological studies relating to immunologic development and function might be used in testing the population mixing hypothesis or in understanding the underlying biological basis of the population mixing hypothesis?

4) Are there “natural experiments” that could be used to study the population mixing hypothesis? If so, what are these situations and how could they be evaluated?

5) What existing databases could be used to study the population mixing hypothesis? How could these databases be utilized for this purpose?

6) Can observational study designs (e.g., case-control, cohort, or other approaches) be used to sufficiently test the population mixing hypothesis?

7) Concerning the Fallon leukemia cluster:
   a) Is “population mixing” a plausible contributor to the cluster of childhood ALL in Fallon, Nevada?
b) Are there factors relevant to the “population mixing” hypothesis that should
   be determined for the Fallon cases/controls?

c) Should “population mixing” be studied in rural communities with large government
   facilities? If so, how might this be done?

After being charged by Dr. Smith with these meeting objectives, participants listened to talks on
existing data relevant to the population mixing hypothesis, infectious disease and immunologic
correlates of this hypothesis, and contributions by molecular epidemiology, as well as potential
applications of geographic information systems (GIS) methods, to evaluating the population
mixing hypothesis.

(Note: References to published literature relevant to the population mixing hypothesis are
provided at the end of the meeting summary, as is a list of meeting attendees.)

II. Overview of Existing Data Relevant to the Population Mixing Hypothesis

Overview of existing data relevant to the Population Mixing Hypothesis—Malcolm Smith (for
Randy Todd)

Dr. Smith presented slides detailing the cancer cluster in Nevada that were prepared by Dr.
Randy Todd, Nevada State Epidemiologist, who could not be present. Fallon, Nevada, located in
Churchill Country, is a community of 8,000 people associated with a large naval air station.
Since 1997, there have been 15 cases of leukemia among children who were either current or
past residents of Churchill Country, with most of the cases occurring in 2000 (9 cases) or 2001 (3
cases). Of the 15 cases under investigation, 11 were residents of Churchill County at the time of
diagnosis. Four additional cases were previous residents of Churchill County but were living
elsewhere at the time of diagnosis.

The 15 leukemia cases comprised 8 females and 7 males, with 8 of the cases occurring in 3- to 5-
year-olds. Four of the cases occurred in 15- to 19-year-olds. In a majority of the cases, the
mother lived in Churchill County when she was pregnant, and the children lived in the county at
birth, at least 2 years before diagnosis, and at the time of diagnosis. Eleven children had B-
precursor acute lymphoblastic leukemia (ALL), three had T-cell ALL, and one had acute
promyelocytic leukemia (APL). Two of three T-cell ALL cases were in the 15-19 age group. A
third T-cell ALL case was in the 9-11 age group. The age distribution of the B-precursor ALL
and the T-cell ALL cases is similar to that observed in the general population.

Tests to date have revealed no unusual environmental exposures in the county except for a high
level of arsenic in both the municipal water supply (approximately 100 ppb) and in family wells
(variable levels, ranging up to 700 ppb).

An exceptional feature of Fallon is its naval air station that has an annual total personnel of
approximately 50,000. The majority of the base personnel are temporary (< 6 months). An
increase in the total personnel associated with the naval air station occurred between 1998 and
1999 (20-30,000 prior to 1999 and approximately 50,000 afterwards). Some Fallon citizens have
expressed concern about the amount of jet fuel to which they might be exposed from this air
station. Substantial interest has focused on the presence of a jet fuel pipeline that runs about 60 miles from just outside of Reno to the NAS-Fallon. A portion of this line runs directly under the town of Fallon. The Nevada Division of Environmental Protection has walked the portion of the line that goes through town looking for gross evidence of leakage. They have also flown the entire length of the line using infrared videography to look for gross leakage. No evidence of leakage was noted. The owner of the pipeline has monitored input and output for evidence of leakage. They have also hired a firm to look for small leaks by injection of an inert substance into the line and then dragging a detection sled over the top. These have also resulted in no evidence of leakage. Navy monitoring wells have indicated no movement of contaminated water off of base property.

The Expert Review Committee advising the Nevada State Health Department recommended that a cross sectional exposure assessment be conducted. The CDC has been the lead agency on this project. Cases and frequency-matched controls (1:4) have participated in biological and environmental sampling (e.g., blood, urine, dust, drinking water, and indoor air samples). Samples have been collected and analysis is presently underway. Biological analytes were selected to correspond to the NHANES analyte list. This will provide a national reference sample. The committee also recommended an environmental pathways assessment. The Agency for Toxic Substances and Disease Registry has been the lead agency for the pathways assessment project which is also currently in progress.

Additional information about the Fallon leukemia cases is available at the Nevada State Health Department website (http://health2k.state.nv.us/healthofficer/Leukemia/Fallon.htm).

**Studies testing the Population Mixing Hypothesis—Leo Kinlen**

Dr Kinlen first illustrated one aspect of the population mixing hypothesis with an example from classical infectious disease epidemiology. When the US first mobilized an army in World War I, striking differences in severe infective outbreaks were noted between camps (sometimes close together) according to the place of origin of their intakes. Epidemics were much more frequent in camps that drew their recruits from the sparsely settled states than in camps with city bred recruits. As the epidemiologists of the US Army Medical Corps recognized, the higher prevalence of susceptible individuals from areas of low population density was potent in fuelling epidemics caused by the mixing.

The population mixing hypothesis holds that childhood leukemia is a rare response to a common but unidentified infection, the transmission of which is promoted by the mixing of rural and urban populations. Thus, among the New Towns created in Britain in the 1950s, significant increases of childhood leukemia were found in the rural towns, but not in the overspill towns set up for Londoners. In the former, new residents were exposed to an increase in population density whereas the reverse was the case for the ex-city dwellers. This has been further tested (and supported) in other situations involving marked rural population mixing, including rural increases in servicemen numbers and large rural construction projects with work camps, where the succession of workers doing different tasks bears comparison with a military transit camp. Recently, a cohort study has demonstrated an almost 4-fold increase of childhood leukemia in
local children in the Orkney and Shetland islands exposed in World War II to the servicemen stationed there, outnumbering local people.

In all these studies, the relative ‘resistance’ of urban areas to the effects was striking, presumably reflecting their lower prevalence of susceptible individuals. Also evident was the fact that the mixing need not directly involve children for adults can mediate the effect. A predilection was noted for the leukemia to affect the children of parents in ‘high contact’ occupations such as teachers or workers in the construction industry. The effects were more marked at ages 0-4 years though the increases were also significant at ages 5-14 years. The excesses found in these studies were short-lived, ending presumably when the prevalence of susceptible individuals had declined sufficiently. A corollary of the hypothesis is that parental exposure in rural areas to large numbers of people through their work may increase the leukemia risk among their children and this has been supported by two recent case-control studies in rural Scotland and rural Sweden.

Dr. Kinlen felt that the relevance of these studies to the Fallon cluster is clear. Situated in the Nevada desert, Fallon is certainly in a rural area that has seen recent population mixing. Indeed, the increase at the naval air station in the number of temporary and other personnel (from 20-30,000 prior to 1999 to approximately 50,000 afterwards) just before the cluster began exceeds the changes in any of the situations so far studied - and so also does the magnitude of the leukemia excess recorded in Fallon. The town is well linked to the base by shared schools, residents working on the base and visits by base personnel.

In answer to questions following his talk, Dr Kinlen stated that the infection causing the leukemia risk cannot be limited only to exposure during pregnancy since in the studies summarized earlier excesses were also evident among children who were born well before the population mixing.

**Testing the Population Mixing Hypothesis for non-cancer chronic diseases—Graham Law**

Dr. Law’s presentation served to expand the scope of the population mixing hypothesis at this meeting from childhood cancer to another disease: childhood diabetes. Population mixing becomes a relevant component to the “hygiene hypothesis” of childhood insulin-dependent diabetes. This hypothesis states that the developing immune system in children matures upon exposure to infectious organisms, and without such exposure—as may occur in our modern, sanitary conditions—the immune system is more susceptible to autoimmune diseases such as diabetes.

Dr. Law presented data using a population-based registry from Yorkshire, UK. From the years 1986 to 1994, 994 cases of childhood diabetes were registered, and the exposure of these children to population mixing was assessed. Dr. Law noted the complex nature of the population mixing concept, in that there are different types of mixing (e.g., commuting versus residential migration) and different scales of mixing (e.g., individual and family level versus regional level). Dr. Law explained that the measure of population mixing requires assessment of both the ‘Volume’ of mixing (i.e., the total numbers of persons migrating into or out of an area) and the ‘Diversity’ of mixing (i.e., the number of different originating regions for the population migrating into a particular area), which can be measured by the “Shannon Index of diversity”.

12 September 2002
[incorporates both of these aspects of population mixing and was used as a measure of population mixing]. Looking only at the diversity in population mixing [this measure] for child migrants, Dr. Law found that areas of low childhood population mixing had higher incidence of childhood diabetes than areas of higher childhood population mixing.

Dr. Law summarized the hypotheses for diseases as postulated by Drs. Greaves and Kinlen and reflected on the role of population mixing as a source of infection in Dr. Kinlen’s population mixing hypothesis. Dr. Greaves’ hypothesis incorporates the idea that an unusual response to a common infection causes the second hit necessary for the advent of diseases such as common ALL. Each hypothesis features a role for common infections leading to disease.

In response to a question, Dr. Law noted that he had not looked at rural versus urban areas, but he noted that the incidence of childhood diabetes is increased in rural areas. A discussion ensued over the definitions of rural and urban areas. A division could be made on the basis of population density, use of public utilities, social contacts, or area available for residences. It became clear that further consideration needs to be given to the factors to consider in defining “rural” and “urban” populations.

Historical overview of leukemia cluster evaluations in the U.S.: relevance to Population-Mixing Hypothesis—Clark Heath

Dr. Heath gave a historic overview of U.S. studies of leukemia clusters to draw attention to the “New Town”, or population mixing, idea. Although now retired, Dr. Heath worked for the CDC in 1961 when the Niles, Illinois, cancer cluster was recognized and investigated. Eight cases of leukemia in 0- to 14-year-olds occurred between 1956 and 1960 in a town that had 7,000 children under 15 years old. The eight cases occurred in children in St. John Brebeuf Parish, and a perceptive nun noticed and reported the cluster to the CDC. This nun also recognized that prior to the appearance of the leukemia cases, a rheumatic fever-like illness had affected children in the parish. Although ultimately a “cause” for the leukemia cases was not uncovered, investigators noted that this area of Niles, Illinois, had recently undergone significant population growth.

Dr. Heath also spoke about four other cancer clusters that were subsequently investigated: Kendall Park, New Jersey; Middleton, Connecticut; Dubois, Pennsylvania; and Winchester, Virginia. All but one location showed rapid population growth prior to the time of the cancer occurrences, showed clustering in a narrow region (in three of the four cases, the clustering was within Catholic parishes), and showed prior incidence of other illnesses. These prior illnesses suggest an infection-related origin for some types of childhood cancer.

After the presentation, it was noted that a 1991 American Journal of Epidemiology supplement provides a very relevant, if somewhat older, source of information about cancer clusters, and that this supplement can be used to educate students about different approaches to studying these situations.
**Recent childhood leukemia cluster evaluations in California—Peggy Reynolds**

Dr. Reynolds emphasized in her talk that a suspected cancer cluster in the state of California may undergo detailed investigation if a number of criteria are met, including: high observed (O) versus expected (E) ratio; statistical significance of occurrence; homogeneity of tumor types; a suspected agent; biologic plausibility; and community concern.

In a 12-month period (July, 2000 through June, 2001), the California Cancer Registry received 8,000 public inquiries, of which 120 involved suspected cancer clusters. Only 8 of the suspected cancer clusters were found to have a statistically significant excess of cancer cases.

Dr. Reynolds described seven suspected childhood leukemia clusters in California from the 1970s through the 1990s. She also described several other childhood cancer clusters involving other types of cancers. Although 85 percent of California’s population lives in urban areas, most of the described cancer clusters occurred in rural areas.

Dr. Reynolds highlighted in particular the McFarland childhood cancer cluster involving 10 cases of cancer in children between 1975 and 1985. These cases became the subject of intense study by the California Department of Health Services. The cluster did not involve a specific cancer type, but rather reflected one or two cases each of multiple cancer types, including leukemias, non-Hodgkin lymphoma, neuroblastoma, osteosarcoma, rhabdomyosarcoma and others. The investigation included a case-control study with environmental testing and child health screening and monitoring for new cases. Despite tremendous effort on the part of investigators and the crucial support of the community, no conclusion could be drawn about the etiology of these cancers. Reports in the media reflected the frustration of the affected populations and did not reflect favorably on the effectiveness of the research teams.

The California State Health Department then decided to take a broader view and evaluated childhood cancer rates in four counties in the rural area that surrounds McFarland. The study revealed that three communities out of 101 had much higher-than-expected rates of cancer, and three communities had much lower-than-expected rates. McFarland was one of the former communities. The distribution of the number of cancer cases among the communities in the four county study was compatible with a Poisson distribution of cases, suggesting that the variation in the number of cancer cases between the communities in the region reflected chance variation.

The relationship between population growth in isolated rural areas and the occurrence of childhood cancer in the four-county area was also evaluated. The “new town” concept is related to the population mixing hypothesis, in that “new towns” are defined as geographically isolated regions that are rapidly increasing in population (2-10 fold increase in 10 years) and that are experiencing an influx of people from diverse areas. In 5 of the 101 communities of the four-county region there was a twofold increase in population growth, but none of these communities experienced an excess of childhood cancer.

Dr. Reynolds also presented studies evaluating other indirect measures of early exposures to infection. A California Department of Health Services study found no association between childhood cancer risk and birth order (1st born), time since last sibling birth (≥ 7yrs), or parity.
(3+ prior live births). However, the Northern California Childhood Leukemia study led by Dr. Pat Buffler of the University of California at Berkeley did observe a significant reduction in risk of ALL for extended daycare attendance, which may involve a substantial exposure to infectious agents.

Dr. Reynolds concluded that investigating individual cancer clusters has generally been an exercise in futility. Moreover, she stressed care in publicizing the investigation of a potential cluster because this may inadvertently increase public expectations, despite warnings that investigations seldom identify an etiologic agent associated with clusters.

**Population mixing and childhood leukemia in Canada—Will King**

Dr. King introduced the study results he presented by noting that Ontario, Canada, has a population of 11.8 million. During the period of investigation—1978 to 1992—the Ontario population increased by one-third. Data on leukemia diagnoses in children aged 0 to 14 years were taken from the Ontario Cancer Registry, and 98 percent of the cases had pathologic confirmation. The ecologic unit for the study was census subdivisions (CSD) in 5-year periods (1978-82, 1983-87, 1988-92). The Ontario province has 850 CDSs, with the median population of each CSD being 2,000 and the median area being 150 km$^2$. Rural CSDs were defined by a population of < 1,000 or by a population density of < 400 per km$^2$. Seventy-eight percent of the CSDs are rural, with 35% of the population at risk living rural in a rural CSD.

Dr. King defined population mixing in this study as the positive population change in a CSD. High exposure was defined as 20% population increase in a CSD. During the 15 year period, 1,394 diagnoses of leukemia were reported, with 1,158 of these being for ALL. When Dr. King focussed on ALL in children 0 to 4 years old, he noted that the rate ratio nearly doubled in rural areas with population mixing as compared with rural areas without population mixing (see table below).

<table>
<thead>
<tr>
<th>% Change in Population</th>
<th>Rate Ratio for ALL in 0-4 Year Olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Urban</td>
</tr>
<tr>
<td>&lt; 0</td>
<td>1.0</td>
</tr>
<tr>
<td>1-20%</td>
<td>1.8 *</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>1.8 *</td>
</tr>
</tbody>
</table>

* p-value < 0.05
+ p-value interaction < 0.01

In contrast to the association between population mixing and childhood ALL in rural areas, no such correlation was observed for 0-4 year olds living in urban areas. Similarly, no statistically significant increased risk of childhood ALL occurred in the 5- to 9-year-old or 10- to 14- year-old age groups among children residing in rural areas in which population mixing occurred. Dr. King and his colleagues considered possible confounding by socioeconomic status and found that the increased risk associated with population mixing diminished after control of this confounder. Dr. King noted that other measures of population mixing could be considered (e.g., determining the percent of persons immigrating into a given geographic region). In addition, other definitions of isolated populations could be considered (e.g., distance from an urban center).
III. Key Concepts and Open Discussion

Discussion followed naturally from the previous talks. The census data in the UK were noted as very helpful and more useful to epidemiologists than the U.S. census data. Because the 10-year U.S. census collection does not have information on where people lived previously or where they were born, it has limited utility in measuring population mixing within an area.

A distinction was noted between population growth and population mixing. Several comments followed: “marked” population growth does not usually mean an increase in birth rates; rural areas achieve larger absolute growth far more easily than urban areas; a crucial measure is that of population diversity; and large-scale emigration from other countries is focused on urban areas because of employment opportunities.

The discussion then led to thinking about testing the population mixing hypothesis by studying people moving from rural areas to urban settings. Areas of Eastern Europe might be appropriate for migration studies of this sort. Another suggestion was to use SEER data to compare incidence of childhood cancer in rural areas versus urban areas in the parts of the U.S. covered by SEER registries and to take advantage of registries included in the IARC Cancer Incidence in Five Continents to carry out similar comparisons.

The topic was raised of the response to—not just the risk of—infection. The genetics and, in particularly, the immunologic diversity of susceptibles needs to be studied. It was noted that there is a paucity of ALL cases in black children versus white children in the United States, and that no hypothesis has been forwarded to explain this difference. Past epidemiological studies identifying risk factors for childhood ALL have not explained the findings of these studies in relation to immunologic and genetic characteristics.

A thought was advanced that a very specific set of hypotheses should be available for use when confronted with potential cancer clusters. These clusters need to be evaluated to address public concerns, but, as Dr. Reynolds noted, extracting scientifically useful insights from cancer clusters seldom occurs. The California experience has led the state Public Health Office to take a more encompassing view of the issues through use of larger population studies. An advantage of this approach is the greater numbers with which to work. Implementation of detailed, large population-based case-control studies is a more scientifically sound means of addressing public concerns, and affected populations may be encouraged to learn that a statewide study is being conducted.

Dr. Malcolm Smith turned the conversation to the definition of isolated population. Definitions of a “rural community” involve distance from urban centers and low population density. Dr. Kinlen argued for the conventional definition to avoid accusations that inappropriate changes were made for the particular study at hand. A respondent suggested that rural towns by this definition should be evaluated by other definitions. If the rural nature of a town is not obvious, it may be wrong to proceed.

A question about the existence of new isolated areas raised the idea of immunologic isolation. An individual child could be immunologically isolated in the middle of an urban area if the child
is breastfed, kept at home, and otherwise not exposed to infections. A study of home-schooled children was suggested in response to this concept.

Finally, the idea of population movement was broached. China’s large, formerly isolated population is now beginning to be more fluid, making this country a potentially valuable geographic area for studying the effect of population mixing on rates of childhood leukemia.

IV. Infectious Disease Correlates of Population Mixing Hypothesis

*Previous evaluations of specific infectious agents for childhood ALL—Malcolm Smith*

Dr. Smith summarized the two-hit theory of ALL etiology, defining the first hit as that occurring *in utero* and the second hit as that occurring postnatally. An infectious agent may either directly or indirectly affect the leukemia precursor cells at either stage.

Dr. Smith described a model of leukemogenesis that he published with Drs. Chen and Simon of the NCI. This model supports an *in utero* initiation of childhood B-precursor ALL. Three variables were described, representing: 1) the *in utero* event (parameter K); 2) the rate of occurrence of the second hit (parameter $\rho$); and 3) the rate of extinction of clones with only the first hit (parameter $\eta$). The model closely fits the age-incidence profile for childhood B-precursor ALL.

An increase in the occurrence of ALL can involve any of the 3 variables, and each of these variables could conceivably be affected by infectious processes and each could be affected in cluster situations. Implications of the model include an argument against an effect on parameter K (the parameter involving the *in utero* event), because children in cancer clusters often present at different ages.

The infectious process may result in a direct and permanent effect on the cells (e.g., human papillomavirus for cervical cancer), a direct but transient “hit-and-run” effect, or an indirect effect. Dr. Smith gave as an example of the latter situation the indirect effects of cytokines produced by the immune system, in that these could potentially promote the growth of (or alternatively extinguish) preleukemic clones. To date no studies have documented the association of a specific virus with childhood ALL. Such a relationship could more easily be uncovered if the virus has a direct effect and remains in the genome of leukemia cells and if such a virus acts alone without the requirement for other factors.

In discussion following Dr. Smith’s talk, TEL-AML1 ALL was used as an example of a leukemia in which the first hit (the TEL-AML1 translocation) is known to occur *in utero*. The second event for cases of ALL with the TEL-AML1 translocation was envisioned as deletion of the normal TEL gene on the chromosome unaffected by the TEL-AML1 translocation. More generally, it was suggested that the second event stimulating leukemia development could be a proliferative event; infection may promote proliferation of cells of the lymphoid system. It was also suggested that the first hit (e.g., the TEL-AML1 translocation for those leukemias with this molecular abnormality) confers a survival advantage to the preleukemic cells.
Methods applicable to identifying infectious etiologies of cancers—Siobhan O'Connor

Dr. O’Connor spoke on the topic of “Field Investigations: Assessing Infectious Etiology in Clusters and Population Mixing.” She emphasized that because investigations for an infectious agent can be labor- and resource-intensive, one needs to prioritize which cancer clusters warrant detailed study. Moreover, in cases of childhood leukemia, prioritization needs to occur in the use of the collected samples because only small amounts are obtained. In general, investigations of an infectious etiology require measurement of antibody, nucleic acid, and/or antigen levels.

The limitations and characteristics of each measure must be understood. For instance, antibody can be assayed using plasma, serum, or blood spots (e.g., Guthrie cards). The presence of a particular antibody type—IgM versus IgG—can be ascertained if present in high enough concentrations and/or the actual level of antibody (i.e. the titer) can be assessed, depending on the information needed. For nucleic acid assays, sensitivity and specificity issues are crucial. Another consideration here is that the infectious agent must be present in the particular specimen tested by PCR or RT-PCR if recognition of its presence is to occur. Moreover, there are challenges associated with field collection of nucleic acids. Similar issues apply to the detection of antigen.

These technologies may associate an infectious agent with disease, but such an outcome also depends on the timing of infection and diagnosis and, again, on the sensitivity and specificity of the assays. Moreover, the mode of transmission imparts its own restrictions. Study of a cancer cluster occurs after the children have been diagnosed and, therefore, potentially after they have cleared a virus. Testing of family members may be important.

Finally, challenges arise in the application of these technologies. Appropriate specimens must be collected under appropriate conditions; the assay(s) should be reproducible, standardized, and validated; and careful interpretation of the data must be made remembering that molecular detection does not equal causality.

Molecular epidemiology of population mixing—Martyn Smith

Dr. Smith noted the lack of molecular epidemiology studies of population mixing and emphasized the multiple problems that could affect future studies evaluating this hypothesis. The problems could include political and legal issues (as the data may be highly scrutinized by the public), and could also include scientific issues such as whom to study and what to measure. Moreover, population mixing studies of childhood leukemia involve young children, and there are inherent problems in studying this population, especially healthy young control children.

In addressing the question of potential study populations, Dr. Smith noted that studying the cases in a cluster is problematic because of the limited information discernible from the small number of cases in a cluster, the question of whether all identified cases are actually part of the same cluster, and the difficulty of defining an appropriate control population.
An alternative population study design involves comparisons of children/families from rural versus urban populations. Such studies would need to account for differences in socioeconomic status and lifestyle, which are not insignificant confounders.

A third possibility for investigations employing molecular epidemiology methods is to study settings in which populations are undergoing mixing. These include areas in which new towns are being built, remote areas in which military/industrial facilities are placed, and areas with increased mobility associated with major political changes (e.g., as in China).

In addressing the question of potential biomarkers, Dr. Smith first considered biomarkers that could be measures of exposure, including exposure to infectious agents (e.g., antibodies and viral DNA) and exposure to other people (biomarkers of contact). Another measure of exposure would be biomarkers indicative of personal hygiene. Dr. Smith then considered biomarkers of early effect, but unfortunately there are no validated markers of early effect. Possible candidates are genetic components (e.g., measure chromosome translocations of relevance to leukemia) and immunological components (e.g., level of T-helper cell subsets and cytokine profiles).

From his own research of chromosomal translocations, Dr. Smith suggested that translocations are frequently found in people who do not go on to develop disease. Translocations may therefore be challenging to use as biomarkers of early effect. However, immunologic features may provide measures relevant to the disease situation. In particular, assessment of the Th1/Th2 phenotype could conceivably be correlated to incidence of leukemia. Intracellular cytokine profiling to characterize the Th1/Th2 phenotype can now be performed by FACS analysis of blood samples small enough to be compatible with amounts obtainable from children. A third general type of biomarker discussed by Dr. Smith are those related to susceptibility. For example, genes controlling susceptibility to infection and immune-response genes are candidate biomarkers of susceptibility.

After Dr. Smith’s talk, he was questioned about gene profiling of the cases in a cancer cluster as a way to discover a feasible biomarker. Dr. Smith remarked that most cases within a cluster have already occurred by the time a cluster is recognized. RNA samples are not usually obtained at the time of diagnosis for gene expression studies, and thus are not available for use in cluster investigations.

**Potential immunologic correlates of the Population Mixing Hypothesis—Steve Chanock**

Dr. Chanock speculated about whether an immunologic fingerprint could be defined that might point to certain types of cancer as the body’s response to infection or as the direct result of an infectious agent. It is difficult to determine the host response in the case of leukemia because, by definition, leukemia is a malignancy of those cells that fight infection. In the 1970s and 1980s, a number of studies tried to relate T-cell function and measurement of cytokines to certain outcomes or to degree of potential susceptibility to disease. Dr. Chanock noted these studies lacked knowledge of appropriate and specific immunologic markers. Similarly, there was great interest at this earlier time in passive immunization as a way to protect children treated for leukemia. By using the host response, researchers felt they could identify important parameters of disease that would lead to preventive measures.
With the advent of the genomics era, tremendous potential exists for characterization of variation, either in susceptibility factors or as a consequence of disease. Up to 200,000 SNPs are expected to be identified over the next 2 to 3 years during the annotation of the human genome. These SNPs can be grouped into families to be used in hypothesis-driven approaches. However, genomics information also provides a unique opportunity for stimulating hypothesis-generating approaches. cDNA profiles may not be understood at present, but the accumulation of these data may provide fertile ground for future analysis. Dr. Chanock urged meeting participants, as a group, to think on “as large a scale as humanly possible” in terms of collecting biological material that will be useful for future analyses.

The critical question voiced by Dr. Chanock is whether the population mixing hypothesis is a way to address the issue of infectious predisposition. Is exposure really a trigger for propagating infection or does it incite a response to infection? Specific genetic parameters identified in isolated groups and measured at the population level may lead to productive hypotheses. It may be the permissive nature of the host response that allows the required “second” and maybe even “third” hits that convert the susceptible cell to a leukemic cell and that achieve maintenance of the leukemic clone.

To assess serologic changes, timing becomes a difficult issue. One or two timepoints will not be enough; serologic evaluation must occur over an extended period of time spent looking at multiple measures of a particular endpoint in any given individual. Dr. Chanock noted that this is an extremely arduous and difficult proposition, especially if the cancer being investigated is rare and occurs in disparate places. Moreover, serologic evaluation must be coupled with an assessment of genetic components. Dr. Chanock encouraged collaboration with developmental researchers who have experience studying multidimensional processes.

Next, Dr. Chanock introduced into discussion a role for proteomics in detailing immunologic changes. He cited the work of Dr. Lance Liotta on ovarian and colon cancers, asserting that such work detailing profiles of proteins can be predictive of particular disease states. Dr. Chanock encouraged researchers investigating the population mixing hypothesis to engage proteomics researchers to look at multiple cytokines and other gene products related to an immunologic response to an infection (or a series of infections).

Using cDNA arrays alone lacks the specificity achieved with the more arduous pursuit of proteomic measures, but genetic measures such as SNPs can be meshed with proteomics to ask questions on the population level.

Dr. Chanock stated that progress will require studying multiple cancer clusters in multiple rural settings. Moreover, information should be acquired about the extended family and the migration of family members. Dr. Chanock suggested tracing microsatellites and unique markers such as SNPs within large populations. Instead of being criticized for “mixing too many groups,” epidemiologists could deliberately do so to use the population’s stratification to find disease determinants not previously recognized.
The choice of a person’s haplotype as an important parameter to measure stems from its use in study over the past 25 to 30 years. Dr. Chanock felt that there was no question that HLA class I and II genes and the associated genes in the HLA complex need to be carefully characterized. Dr. Chanock next addressed cytokines and argued for looking at those that are clustered together in the genome or that are related at the sequence level. He mentioned that activation of the IL-4 pathways involved 18 different gene products, any of which could be critical in modulating the pathway.

In conclusion, Dr. Chanock endorsed using large population studies to investigate the population mixing hypothesis. Only through studies involving multiple study sites will there be sufficient statistical power to identify important etiologic associations.

**Potential applications of GIS to evaluating the Population Mixing Hypothesis—Daniel Wartenberg.**

Dr. Wartenberg’s presentation began with a review of risk factors for childhood leukemia and the hypotheses proposed by Drs. Kinlen and Greaves. The Kinlen population mixing hypothesis proposes that when children from isolated populations with presumed decreased exposure to infections are exposed to others from regions of greater population density, they are at increased risk of developing leukemia, likely due to an infectious etiology. The Greaves hypothesis proposes that a delay of exposure to infections from infancy through childhood, such as due to improvements in socioeconomic status and hygiene, puts children at increased risk of leukemia and lymphomas.

Dr. Wartenberg focused on the potential contributions of geographic information system (GIS) methods to investigation of the population mixing hypothesis. GIS is a relational database that is organized by location. It can display data to produce “hot spot” maps, trends, and space-time animation. It can retrieve data by location, including demographics, socioeconomic characteristics, and environmental exposures. It can facilitate spatial/space-time analysis to identify homogeneous regions and to guide spatially weighted regression analysis for studies investigating the relevance of proximity to specific geographic factors. The primary limitation of using GIS is data availability. Dr. Wartenberg stressed the importance of obtaining specific individual data to facilitate the application of GIS methods to epidemiological studies of childhood cancer.

To provide an example, Dr. Wartenberg used GIS to analyze SEER data for acute lymphoblastic leukemia (ALL) incidence in children 0 to 4 years old in four states. The odds ratios for ALL development were increased in rural counties that experienced population growth compared to rural counties with stable or decreasing population. There was an apparent dose response for increasing risk associated with population growth for the categories 0-10%, 10-20% and > 20% increase in population size. A similar relationship between ALL risk and population growth was obtained when looking at incidence data for children 0 to 19 years of age, though the dose response was less clear. Inclusion of more urban areas weakened the association between population growth and ALL risk, though a similar pattern of risk remained.
Dr. Wartenberg noted the limitations of the data and analytic methods used for this preliminary study of the relationship between population growth and risk of ALL. One limitation is the poor quality of the geographic data: the SEER data provide only state of birth and county of diagnosis. Forty-five percent of cases are not born in the same state in which they are diagnosed with childhood cancer, which could lead to possible selection bias. Other limitations include the arbitrary breakpoints used for categorizing change in population size, the limited number of cases, and the lack of consideration of proximity to urban areas.

Dr. Wartenberg concluded his talk by summarizing the utility of GIS. He felt that the GIS could facilitate direct evaluation of the population mixing hypothesis and could also be used to evaluate the reverse hypothesis—that is, to compare those born in rural states and diagnosed in urban areas, with those both born and diagnosed in urban areas.

V. Issues in Study Design for Testing the Population Mixing Hypothesis

Drs. Martha Linet and Eve Roman moderated the discussion of study design issues related to testing the population mixing hypothesis. Discussion was initiated informally with a request for a summarization of associations of ALL risk with: 1) a history of vaccination, or 2) the occurrence of illness in siblings. It was felt that a better understanding of such data would direct collection of appropriate specimens (urine, saliva, etc.) in future studies. The consensus of those present at the meeting was that case-control studies had not shown such an association. Significantly, there are few unvaccinated children to use as a control group.

Dr. Ross raised the idea of studying individuals with Down Syndrome as a uniquely susceptible population. These individuals have a 1 in 100 chance of developing acute leukemia (either ALL or AML). Serial blood samples could be taken among cohort members with Down syndrome to help assess susceptibility factors. While the relatively high risk of leukemia for children with Down syndrome supports studying this population, it is not known whether the pathogenesis of leukemia in these children is similar to that for common ALL in the general childhood population. Dr. Ross noted that the immunocompromised nature of many children with Down Syndrome may allow this group of subjects to be informative, although it was also noted that immunodeficiency may not play the role in leukemia susceptibility that the genetic abnormality (trisomy 21) does.

Dr. Malcolm Smith then directed discussion to address the following questions: Are there recommendations that this group might have for investigations of the etiology of the cluster of childhood leukemias in Fallon and for information that should be collected in Fallon related to population mixing? Are there other similar towns near large military or industrial facilities that should be studied? What large studies should be conducted in the future, and how?

Dr. Roman provided a short presentation to recapitulate the last 20 years of study on the Windscale cancer cluster in the United Kingdom. Many diverse types of epidemiologic studies have been conducted, from ecological to case-control to cohort, but very little has been uncovered to conclude an etiology of the original excess risk of leukemia and lymphoma. Only the population mixing hypothesis remains as a possible explanation for this cluster. Dr. Roman urged research on etiology rather than “cluster-busting,” which has been generally unproductive.
Regarding Fallon, she had no specific ideas for particular studies or approaches to recommend and, indeed, suggested that no efforts be made beyond those ongoing by the CDC. She encouraged careful review of data from studies currently underway or recently completed in order to identify potential leads for future studies. Dr. Martha Linet agreed and cautioned against initiating new large studies prior to careful analysis of these prior studies. She believes much is yet to be learned from ongoing large studies such as the United Kingdom childhood cancer study, which is currently undergoing multiple analyses. The group felt that any new large epidemiological studies should incorporate methods to address the population mixing hypothesis and should try to investigate the underlying biology of disease.

The group next focused on recommendations that could be made for the large cohort lifetime study of children from pregnancy onward currently being planned jointly by NICHD, EPA, and CDC. Particular emphasis was placed on acquiring immune response measures. Dr. Pat Buffler is willing to be a contact person to forward ideas to the working group. It was recommended that a longitudinal study be considered focusing on immune response development and on children in whom in utero translocations had occurred. Another opinion was voiced that smaller studies may be able to address certain questions in a more timely fashion than the big cohort study being planned.

Dr. Linet next described the encouragement provided by Dr. Rick Klausner, former NCI Director, to develop international consortia. In one endeavor, investigators who have completed or are conducting case-control epidemiologic studies of NHL have developed a consortium termed “InterLymph.” NCI wants to support joint collaborative efforts of this type of activities by developing new funding initiatives.

Particular recommendations for the Fallon situation included collecting biological samples. Identification of cancer clusters occurring in similar small towns in rural areas may also be informative and may be a way to accumulate larger numbers of subjects who may share similar exposures and/or etiology.

For current and future large case-control studies of childhood leukemia, meeting participants were asked to suggest and prioritize a list of potential genes that may modify the risk of developing ALL, as has been compiled by InterLymph members. Drs. Chanock and Morgan will forward the list to the group. It was felt that an overlap of priorities among groups evaluating childhood leukemia clusters will add cohesiveness to the various studies.

Dr. Heath expressed his opinion that a cancer cluster should not be described by numbers alone and that more questions about the social aspects to a cluster can and should be asked. He emphasized that space-time clustering of cancer could occur due to chance.

Discussion revolved around ways to investigate cancer clusters and the practical and legal issues associated with storing and testing biological samples. A consensus was reached on the need to define urban and rural areas. A third term, suburban, was also forwarded for definition. The suggestion was made to include the concept of diversity along with that of population density in defining these terms.
VI. Summary and Discussion of Action Items

Dr. Malcolm Smith concluded the meeting by extending the appreciation of the organizers to the meeting attendees for their active participation. In summary, he noted that clusters should be studied with caution, especially in regard to public perception; that the population mixing hypothesis offers a plausible explanation for increased incidence of ALL in special situations; that the underlying biology for the increased ALL risk that may be associated with population mixing is important to study in the future, especially with regard to the immune response; and that common definitions are difficult to establish but will be important to progress in the field.

Action Items
- Circulate a list of genes to target in epidemiologic studies of childhood cancer [Steve Chanock and Gareth Morgan].
- Prepare initial draft of definitions of rural, suburban, and urban for circulation to meeting participants.
- Convey suggestions to Dr. Pat Buffler about items to include in data collection for the National Children’s Study (longitudinal cohort study).

VII. Background References Related to Population Mixing

Childhood Cancer Clusters and Clustering:

In utero Initiation of Childhood ALL:


**Population Mixing References:**


2) Fear NT, Roman E, Reeves G, Pannett B. Are the children of fathers whose jobs involve contact with many people at an increased risk of leukaemia? Occup Environ Med 1999; 56(7):438-442


**Population Mixing and Childhood Diabetes:**


**Possible Role of Infectious Agents in Childhood ALL Etiology:**


VIII. Meeting Participants

Freda Alexander, Ph.D.
Professor, Department of Public Health Sciences
University of Edinburgh Medical School
Teviot Place
Edinburgh EH8 9AG Scotland
United Kingdom

Martin Belson, M.D.
ORISE Fellow
Centers for Disease Control and Prevention
1600 Clifton Road, NE, Mailstop E-23
Atlanta, GA 30333

Roumiana S. Boneva, M.D., Ph.D.
Medical Epidemiologist
Centers for Disease Control and Prevention
NCID, Mailstop G-19
1600 Clifton Road, NE
Atlanta, GA 30333

Patricia A. Buffler, Ph.D., M.P.H.
Professor of Epidemiology
Division of Public Health Biology and Epidemiology
University of California at Berkeley
714-F University Hall
Berkeley, CA 94720-7360

Stephen Chanock, M.D.
Senior Investigator
Pediatric Oncology Branch, CCR
National Cancer Institute
Advanced Technology Center
8717 Grovemont Circle
Gaithersburg, MD 20877

Osborn Bryan Eden, M.B.B.S., DR.C.O.G.
Academic Unit of Pediatric Oncology
Christie Hospital NHS Trust
Wilmslow Road
Withington, Manchester M20 4BX
United Kingdom

Carolyn A. Felix, M.D.
Associate Professor of Pediatrics
Division of Oncology
The Children's Hospital of Philadelphia
Abramson Research Building, Room 502B

3516 Civic Center Boulevard
Philadelphia, PA 19104

James J. Goedert, M.D.
Chief, Viral Epidemiology Branch, DCEG
National Cancer Institute
Executive Plaza South, Room 8003
6120 Executive Boulevard
Rockville, MD 20852

Frank Groves, M.D., M.P.H.
Department of Biometry and Epidemiology
Medical University of South Carolina
135 Cannon Street, Room 302-H
Charleston, SC 29425-0835

James W. Hanson, M.D.
Acting Branch Chief
Clinical and Genetic Epidemiology Research Branch, DCCPS
National Cancer Institute
Executive Plaza North, Suite 5134
6130 Executive Boulevard
Bethesda, MD 20892-7393

Clark W. Heath, Jr., M.D.
retired (American Cancer Society)
675 Plantation Point Road
Woodbine, GA 31569-2124

Adrianne Holmes
Project Officer
Centers for Disease Control and Prevention
1600 Clifton Road, NE, Mailstop E-23
Atlanta, GA 30333

Peter Kaatsch, Ph.D.
Institute for Medical Biostatistics, Epidemiology and Informatics
German Childhood Cancer Registry
Langenbeckstr. 1
55131 Mainz
Germany

Wendy E. Kaye, Ph.D.
Chief, Epidemiology and Surveillance Branch
Agency for Toxic Substances and Disease Registry, CDC
1600 Clifton Road, NE, Mailstop E-31
Atlanta, GA  30333

Will D. King, Ph.D.
Associate Professor, Department of Community Health and Epidemiology
Queen's University
2nd Floor Abramsky Hall
Kingston, Ontario K7L 3N6
Canada

Leo Kinlen M.D.
Professor, Cancer Epidemiology Unit
University of Oxford
Gibson Building, Radcliffe Infirmary
Oxford OX2 6HE
United Kingdom

Graham Law, Ph.D.
Leukaemia Research Fund
Epidemiology and Genetics Unit
University of Leeds
Margaret Smith Building
Leeds LS2 9JT
United Kingdom

Martha S. Linet, M.D.
Chief, Population Studies Section
Radiation Epidemiology Branch, DCEG
National Cancer Institute
Executive Plaza South, Room 7054
6120 Executive Boulevard
Bethesda, MD  20892-7238

Peter McCarron, M.D., M.P.H.
NICR Consortium Fellow in Cancer Epidemiology, DCCPS
National Cancer Institute
6116 Executive Boulevard, Suite 504
Bethesda, MD  20892-8316

Lee Moore, Ph.D., M.P.H.
Investigator
Occupational Epidemiology Branch, DCEG
National Cancer Institute
Executive Plaza South, Room 8118
6120 Executive Boulevard
Bethesda, MD  20892-7240

Gareth Morgan, M.D.
Professor, Department of Hematology
Institute of Pathology
University of Leeds
Great George Street
Leeds LS2 9JT
United Kingdom

Nancy Nelson, M.S.
Science Writer, Mass Media Branch
Office of Communications, OD
National Cancer Institute
Building 31, Room 10A19
Bethesda, MD  20892

Siobhan O'Connor, M.D., M.P.H.
Assistant to the Director of the National Center for Infectious Diseases
Centers for Disease Control and Prevention
1600 Clifton Road, NE, Mailstop C-12
Atlanta, GA  30333

Peggy Reynolds, Ph.D.
Chief, Environmental Epidemiology and Geographic Information Section
State of California - Department of Health Services
1515 Clay Street, Suite 1700
Oakland, CA  94612

Lynn Ries, M.S.
Health Statistician, SEER Program
Cancer Statistics Branch, DCCPS
National Cancer Institute
6116 Executive Boulevard, Suite 504
Bethesda, MD  20892-8316

Leslie L. Robison, Ph.D.
Professor of Pediatrics
Department of Pediatrics
University of Minnesota
Mayo Mail Code 422, Room 554 CCRB
420 Delaware Street, S.E.
Minneapolis, MN  55455

Eve Roman, Ph.D.
Leukaemia Research Fund
Epidemiology and Genetics Unit
University of Leeds
Margaret Smith Building
Leeds LS2 9JT
United Kingdom

Julie A. Ross, Ph.D.
Associate Professor
Department of Pediatrics
University of Minnesota
Mayo Mail Code 422, Room 554 CCRB
420 Delaware Street, S.E.
Minneapolis, MN  55455

Carol Rubin, D.V.M., M.P.H.
Chief, Health Studies Branch
Division of Environmental Hazards and Health Effects
Centers for Disease Control and Prevention
1600 Clifton Road, NE, Mailstop E-23
Atlanta, GA  30333

Joachim Schuz, Ph.D.
Institute for Medical Biometrics, Epidemiology and Informatics
University of Mainz
D-55101 Mainz
Germany

Thomas H. Sinks, Ph.D.
Associate Director for Science
Division of Birth Defects, Child Development, and Disability and Health
National Center for Environmental Health
Centers for Disease Control and Prevention
4770 Buford Highway, NE, Mailstop F-29
Atlanta, GA  30341-3724

Malcolm A. Smith, M.D., Ph.D.
Head, Pediatric Section
Clinical Investigations Branch
Cancer Therapy Evaluation Program, DCTD
National Cancer Institute
Executive Plaza North, Room 7025
6130 Executive Boulevard
Bethesda, MD  20892

Martyn T. Smith, Ph.D.
Professor of Toxicology
Division of Environmental Health Sciences
University of California at Berkeley
School of Public Health

216 Warren Hall
Berkeley, CA  94720-7360

Logan Spector
Doctoral Candidate
Department of Epidemiology
Emory University
1518 Clifton Road, NE
Atlanta, GA  30322

Brooke Steele, Ph.D.
Cancer Surveillance Branch
Division of Cancer Prevention and Control
National Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention
Atlanta, GA  30341-3724

Malcolm Taylor, M.D.
Reader in Child Health and Medical Genetics
Immunogenetics Laboratory
Department of Medical Genetics
St. Mary's Hospital
Manchester M13 0JH
United Kingdom

Daniel Wartenberg
Environmental and Occupational Health Sciences Institute
170 Frelinghuysen Road
Piscataway, NJ  08854