International Childhood Cancer Cohort Consortium Workshop
September 28–29, 2005
Doubletree Rockville Hotel
Rockville, MD, USA

This meeting was held in conjunction with the National Children's Study, which is led by a consortium of federal agency partners: the U.S. Department of Health and Human Services (including the National Institute of Child Health and Human Development [NICHD] and the National Institute of Environmental Health Sciences [NIEHS], two parts of the National Institutes of Health, and the Centers for Disease Control and Prevention [CDC]) and the U.S. Environmental Protection Agency (EPA). This meeting was also supported by the Office of Rare Diseases, NIH, DHHS.

Welcome
Peter Scheidt, M.D., M.P.H., NICHD, NIH, DHHS

Dr. Scheidt, director of the National Children's Study, welcomed participants on behalf of the National Children's Study, the EPA, and the Office of Rare Diseases. The National Children's Study represents one cohort of the International Childhood Cancer Cohort Consortium (ICCCC). It will be a multi-stage nationally representative sample of 100,000 participants derived from 101 locations throughout the country. Eight vanguard sites will soon begin to implement the sample selection and data collection, and one clinical coordinating center will oversee the sampling procedures, data collection and analysis.

According to Dr. Scheidt, the National Children's Study holds the potential to:
- Define genetic and biomarkers that might be useful in answering questions about childhood cancers
- Validate exposure measures for case-control studies
- Serve as a control group for larger childhood cancer studies and registries.

Despite this potential, researchers recognized that a sample size of 100,000 is too small for reliable exposure-outcome relationships for childhood cancer. An international consortium such as the ICCCC offers the potential to overcome some of the inherent limitations of the National Children's Study as well as those of other cohort studies around the world. Dr. Scheidt emphasized that the ICCCC meeting is a working meeting. If the conclusion of the meeting is that the consortium is feasible, the expectation of the meeting is to identify possible:
- Realistic hypotheses
- Usable independent variables
- Mechanistic studies
- Next steps.

Key Hypotheses for the Etiology of Childhood Cancers
Terence Dwyer, M.D., M.P.H., Murdoch Children's Research Institute

According to Dr. Dwyer, case-control studies have been the principal strategy used to examine the association between environmental exposures and childhood cancer. However, case-control

studies are limited in terms of the reliability of data and their ability to draw relationships between exposures and outcomes that have occurred years prior to diagnosis of childhood cancers.

Recall Bias. Case-control studies of childhood cancer often rely on data derived from parental recall—the reliability of which is uncertain. One German study found that recall varied from what was reported at birth to the time the interview was conducted, suggesting that recall decreases with time. A study based on parental recall both before and after a sudden infant death (SIDS), demonstrated that parents assigned more retrospective importance to variables they saw as possibly contributing to their child's death. This potential bias could change the odds ratio (OR). Very large cohort studies have the potential to overcome recall bias.

Power of Studies. ICCCC will have the statistical power to detect small-to-moderate risks associated with the following certain exposure-disease associations (Table 1). As Dr. Dwyer noted, there is no single cohort study currently that has the power needed to reliably measure many exposure-disease associations. For an exposure affecting 5 percent of the population, more than 1 million participants would be necessary to attain the power needed to detect associations. For a more common exposure (15 percent) the sample would have to include more than 400,000 participants. Table 1 summarizes the cohort size that would be needed to reliably detect associations with acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML).

Table 1. Sample Size Needed to Detect Associations with Acute Leukaemia (ALL and AML)

Age Adjusted Incidence per 100,000 Person Years	Incidence for Cohort Follow-Up from Birth to 14 Years	Exposure %	Minimum OR Detectable	Power %	Number Required for 80% Power	Power %	Number Required for 90% Power
4.6	64.4	5	1.5	80	1,180,059	90	1,635,361
4.6	64.4	15	1.5	80	446,633	90	613,158
4.6	64.4	30	1.5	80	277,781	90	376,372
4.6	64.4	5	2.0	80	328,992	90	467,041
4.6	64.4	15	2.0	80	125,813	90	175,358
4.6	64.4	30	2.0	80	79,594	90	108,289

Source: Garcia-Closas M, Lubin JHH. Am J Epidemiol 1999; Age-adjusted SEER cancer incidence rates, USA 1975–2002

Table 2 summarizes data available through proposed ICCCC cohorts. Combined, the cohorts have more than 700,000 participants—a sample size sufficient to analyze more common exposures.

Table 2. Summary of Cohorts

Study	ICCCC Representative	Years of Recruitment	Cohort Entry Criteria	Age at Cohort Entry	Study Sample Size (Number Recruited at Cohort Entry)
Avon Longitudinal Study of Parents and Children (ALSPAC)	Jean Golding	1990–1992	Pregnant women resident in the geographical area of Avon with expected date of delivery April 1, 1991–December 31, 1992	Prenatal	14,541 pregnancies resulting in 3,988 children surviving the first year of life
Bradford Babies: Growing up in Bradford, UK	Nicola Symons	2006–2008	Born in Bradford	At birth	10,000
Canadian Childhood Cancer Surveillance and Control Program	Les Mery	1995–2001	Diagnosed and treated in a pediatric cancer centre in Canada	0–19 years	6,000
China Family and Children Cohort Study	Li Zhu	2006–2007	Newly married women	Preconception, prenatal	300,000
China-U.S. Collaborative Project on Birth Defects and Disabilities Prevention	Li Zhu	1993–1995	Newly married women	Preconception, prenatal	245,000

Danish National Birth Cohort (Better Health for Mother and Child)	Jørn Olsen, Sjurdur Olsen, Marie Louise Østerdal	1996–2002	All who spoke Danish well enough to take part in the interviews and who intended to carry the pregnancy to term	Pregnancy	101,000
French Study on Environment and Children's Health (EFESE)	Joëlle Le Moal	2008–2009		Pregnancy, at birth	20,000
Infancia y Medio Ambiente (INMA), Spain	Nuria Ribas-Fito	2001–2005		Week 12 of gestation	4,200
Jerusalem Perinatal Cohort Study	Susan Harlap	1964–1976	Population-based, total population, based on mother's residential address at the time of birth	At birth	91,458 live births
National Children's Study, USA	Peter Scheidt	2007–2011	National probability sample for pregnant women in their first trimester, women attempting pregnancy, and women of childbearing age	Preconception, pregnancy	100,000

Norwegian Mother and Child Cohort Study	Andrej Grjibovski	1999 onward	All pregnant women in Norway have been invited to participate. Invitation three weeks prior to routine ultrasound in week 17–19 of pregnancy	Approximately 17 gestational weeks	100,000 (50,000 as of March 2005)
Tasmanian Infant Health Survey (TIHS)	Terence Dwyer	1988–1995	Perinatal profile indicting higher risk of sudden infant death	Postnatal age 4 days	10,627

Potential Hypotheses. There is sufficient ecological variation in incidence of childhood leukemia associated with geography or socioeconomic factors to suggest that environmental factors are likely to be important. Moreover, there is evidence that chromosomal translocations present at birth—probably occurring during fetal life—are etiologically important. Other hypotheses that might warrant testing in a combined cohort include:

- **Birth weight.** Research suggests that higher birth weight increases the risk of ALL and AML. Possible causal pathways involving birth weight include maternal diet preconceptionally and during pregnancy, insulin-like growth factor 1 (IGF1) (shown to be important for breast and pancreatic cancers), paternal genomic imprinting of IGF2, and socioeconomic-related factors (for example, childhood infection).
- **Folate intake.** Low maternal folate intake periconceptionally has been shown to increase the risk of ALL, possibly via chromosomal translocations. Polymorphisms of the methylenetetrahydrofolate reductase (MTHFR) gene reduce production of active folate. Populations with a higher prevalence of these MTHFR polymorphisms—for example, Hispanics in the United States—have a different incidence of ALL than other U.S. population subgroups. If the theory holds, researchers expect to see a decrease in incidence of ALL with the inclusion of folate in the maternal diet as standard of care. Evidence to date is mixed. In Australia, for example, there was no marked decline in incidence of ALL since folate was introduced. However, Canadian data based in a small genetic subset demonstrate a notable change following the introduction of folate.
- **Exposure to infectious disease.** Data on associations between exposure to infectious disease and childhood cancer is mixed. Chan and colleagues found that the timing of exposure to infection might be a major determinant of ALL. Infection early on was associated with decreased risk, whereas at later stages, it was associated with increased risk.
- **Gene-environment interactions.** The search for gene-environment interactions is likely to be important. By way of example, Dr. Dwyer noted research showing the interaction between MTHFR genotype and low red blood cell (RBC) folate. The children of women who are

homozygous for the 677C→T polymorphism in MTHFR and who have RBC folate in the lowest quartile show a higher risk of cancer.

A carefully planned international consortium may help answer some of these questions. The remainder of the meeting will provide greater detail on the questions scientists might ask, the methods for answering them, and the feasibility of answering them through an international consortium.

Overview of Cancer Consortia and Success Stories

Daniela Seminara, Ph.D., M.P.H., National Cancer Institute, NIH, DHHS

Dr. Seminara suggested that the interdisciplinary consortium in epidemiology presents a paradigm that, while not entirely new, challenges traditional research. The benefit of a consortium in epidemiology is that it can facilitate the rapid replication of findings and increase the sample size by pooling data. A consortium can also support the study of:

- Interactions with environmental exposures
- Complex multigenic effects
- Gene discovery
- Etiologic heterogeneity for tumor subgroups
- Prognostic factors.

NIH supports the Interdisciplinary Research Implementation Group and Interdisciplinary Research (IR) Centers. NIH planning grants will be awarded for IR programs that address significant and complex biomedical problems, particularly those that have been resistant to more traditional approaches. Planning activities will include approaches to overcoming traditional institutional barriers to IR, which are intended to lay the foundation and prepare investigators for submitting a subsequent application for support through an IR consortium.

According to Dr. Seminara, the Epidemiology and Genetics Research Program (EGRP) currently awards $225 million to consortia (compared to $740 million for non-consortia activity). The EGRP-supported epidemiology consortia provide flexibility of design and research that promotes:

- Gene discovery
- Gene characterization
- Gene-gene and gene-environment interactions
- Translational clinical genetics
- Screening, prevention, and treatment.

Emerging consortia are defined as three or more groups of investigators from different institutions planning to launch a joint initiative by combining resources from case-control, familial or cohort studies. EGRP criteria for evaluating emerging consortia include:

- Scientific rationale and justification of need
- Preliminary rationale that large numbers are needed to address research questions
- Outline of proposed internal leadership and organizational structure
- Outline of guidelines for sharing of data and specimen resources and publication policies
- Tabulation of similarities and differences in design, data variables, and specimen acquisition

and storage (if applicable) across studies
- Proposed plan to address informed consent issues
- Consortia challenges and possible solutions (see Table 3).

Table 3. Consortia Challenges and Possible Solutions

Challenges	Possible Solutions
Communication and coordination	Web site, portals, teleconferences and in-person meetings
Informed consent and variable behaviors of institutional review boards (IRBs)	Prospective consortia, re-consent, education of IRBs
Informatics and analytic support for collection, management and analysis of extremely large and complex datasets	Central informatics units, standardization of informatics platform (caBIG), "think tank for analytic challenges"
Rapid and continuous integration of cutting-edge genomic and other technologies	Centralized technology platforms, public-private partnership
Biorepositories: centralized versus local, large scale retrieval of tissue	Work toward maximizing bioresources (transformed cell lines, WGA, pooling, tissue microdissection, multiplex microarrays)
Integration of disciplines	Interdisciplinary training, integration of new knowledge and concepts as they arise, shift in academic culture triggered by multiple outcome funding approaches
Intellectual property rights	Carefully crafted agreements, involving all partners
Authorship and principal investigatorship (especially for young investigators)	Change in structure of funding mechanisms, tenure criteria, and publication credits
Access for the scientific community at large (data sharing)	Development of clear process and policies (for example, CFRs), NIH may help with cost of sharing data
Review process	Appropriate IRG, education of peer scientists; interdisciplinary science requires interdisciplinary peer-review
Interdisciplinary research teams take time to assemble and require unique resources	Appropriate criteria for evaluation and measure of productivity taking in account planning and time to establish Infrastructure. Evaluate core activities and tools developed

Dr. Seminara listed several challenges faced by emerging consortia:

- Tighter pay lines for NIH grants
- Limits to programmatic discretion
- Added scrutiny for expensive studies
- Need for extensive coordination
- Rapidly evolving technology
- Need for meaningful data sharing models.

Discussion. Participants discussed several issues relating to ICCCC and consortia in general, including authorship issues, distribution of findings, past experience, and strategies for consolidating resources.

- **Authorship.** A common way to publish is as a consortium, but with a list all those involved somewhere in the article. Alternatively, researchers with a particular interest can take the lead on the appropriate scientific questions. Dr. Seminara noted that only major findings using pooled data are published by the consortium. Moreover, journals are becoming more sensitive to the need for multiple authors and sometimes allow equal ranking to all authors. Nonetheless, a very large consortium is likely to produce multiple papers. As the consortium matures, these issues are likely to work themselves out.
- **Public distribution of findings.** Generally, one or two participants provide content review for public information materials. The EGRP Web site provides a large space dedicated to the community. In some instances, the dissemination of research findings is supported by patient groups (as is the case for brain cancer).
- **Prior experience.** Deborah Winn, Ph.D., National Cancer Institute (NCI), NIH, DHHS, observed that ICCCC attendees have an advantage in that NCI has experience in developing successful consortia and has templates available from other consortia. Martha Linet, M.D., M.P.H., NCI, NIH, DHHS, agreed, adding that the ICCCC could learn from similar consortia, such as International Lymphoma Epidemiology Consortium (Interlymph). This grass roots effort arose in response to the global rise in lymphoma. Just 3 years since its inception, Interlymph has its third publication in press.
- **Consolidation of resources.** Stephen J. Chanock, M.D., NCI, NIH, DHHS, suggested that if ICCCC is launched, it will be unnecessary to conduct genotyping at six centers—one or two may be sufficient. This is a means to consolidate, control cost, and help concordance. Interlymph provides a good example of how this can be done. Interlymph investigators maintained a 99.9 percent concordance between centers and a 0.4 percent rate of error—equivalent to the error of the technology itself. Dr. Chanock noted, however, that there might be political issues with transporting samples out of the country.

Environment and Childhood Cancer: Challenges and Prospects

Andrew Olshan, Ph.D., University of North Carolina, Chapel Hill

Dr. Olshan noted that his presentation focused primarily on the state of the science for environment and childhood cancer, with particular attention to how an international consortium might enable researchers to surpass the capabilities of case-control studies.

The earliest studies of childhood cancer were based on vital statistics, primarily birth certificates linked to death certificates that listed cancer as the cause of death. This information was later

linked with other data, such as pesticide use. Case-control childhood cancer studies have since evolved in several ways:

- Increasing size (early studies had small samples that often grouped disparate childhood cancer types; now studies often have more than 1,000 cases)
- Use of medical records for validation of parental reports of medical conditions, treatments, family history of cancer
- Case subtypes (molecular markers, histologic type of tumor, age, and so on)
- More sophisticated questionnaires
- Polymorphisms of carcinogen metabolism, nutrients, single nucleotide polymorphism (SNP) pathways, DNA repair
- Exposure substudies (electromagnetic fields, pesticides, and so on).

Enhanced Occupational Questionnaires. There have been more than 40 epidemiologic studies since 1974 on parental exposures and childhood cancer. Most positive associations are with *paternal* occupation. Generally, these exposure assessments were based on interviews with the mother, and included:

- Source
- Job title/industry
- Exposure checklists
- Generic job-exposure matrices (JEMs)
- Methods based on industrial hygiene principles.

Occupational questionnaires have evolved to include job-specific questions related to work activities, environment, and exposures. Substudies and external databases have fine-tuned exposure evaluation. As a result, researchers are faced with the challenge of making the findings more precise to match the more detailed evaluation.

Exposure Assessment. As Dr. Olshan noted, although the connection between cancer and pesticide exposure has been well studied, controversy continues about the carcinogenicity of some pesticides as well as mutogenicity versus carcinogenicity. A challenge has been to identify the routes of exposure, sources of exposure (including residential, agricultural, and other sources) as well as different mechanisms (mother, father, *in utero*, and so on). Exposure assessment has taken many different forms, including residential evaluations and exposure checklists. Researchers have looked at general exposure, garden sources, general pesticide use, farm residences, and more. A variety of studies have ranked exposures from low to high. Some studies have examined subgroupings of childhood cancer for example by histologic subtypes. Some researchers have examined single nucleotide polymorphisms (SNPs), particularly NQ01—a SNP associated with pathways that may offer protection from the toxicity of certain compounds (such as benzene). In theory, individuals who have not inherited certain polymorphisms of enzymes related to metabolism involving NQ01 would have higher incidence of leukemia.

Research demonstrates suggestive, if mixed, associations for farm residences, professional extermination, garden use of pesticides, and use of pet flea/tick sprays (for prenatal exposure only). On the whole, there is a pattern of increased cancer risk associated with pesticide exposure—a pattern that becomes more impressive with detailed assessments. Despite

refinements to the study design and the increasing sophistication of analyses, methodological issues limit the value of case-control studies. These issues include:

- Power/sample size (n=45–504)
- Etiologic heterogeneity (in relation to differing histologies, molecular markers, age)
- Confounding bias (unknown risk factors)
- Controls/selection bias
- Exposure misclassification (many based on self reporting using less than desirable instruments).

Research Bias. A mixture of biases makes it difficult to interpret data and to identify false positives that are due to chance or selection bias. In identifying exposures, a lack of specificity limits the certainty of findings. Many studies have only been able to look at exposures in very broad categories based on recall. A cohort study with biomarkers and other measures holds promise of overcoming this limitation. Recall bias also limits the outcomes. Parents tend to be more motivated once there is a diagnosis of cancer. However, a growing body of literature suggests this bias may not be as pronounced as once thought.

At the same time, there is also a growing body of evidence that suggests that for sensitive items (pesticide use, abortion, illicit drugs), there is a greater potential for case mothers to link that with their child's disease. Timing represents another limitation. Case-control studies are dependent upon a parent's recall of the exposure window. Even in instances in which the study collects what might be deemed historical evidence (such as vacuuming a household carpet or floor) these measures are imprecise in identifying the exposure window. Moreover, frequency and duration of use is difficult to ascertain in hindsight. Finally, when the subsample is very small—even when derived from a large study—the confidence interval is very wide and imprecise. Significance does not necessarily imply a clear link.

Opportunities for childhood cancer research from having an international consortium include:

- Larger studies
- Ability to better understand certain exposures due to greater exposure variation (for example, substances present in higher levels or banned in the United States can be examined elsewhere)
- Emphasis on etiologic heterogeneity (subdividing cases by molecular markers, tumor type)
- New methodologies for and increased use of exposure assessments
- Validation substudies
- Exploration of gene-environment interaction.

Discussion. Participants discussed how many protocols can be realistically added to existing cohort studies, the implications for the inclusion of biological samples, and the relative benefits and drawbacks of a very large study. Specific points included:

- Sholom Wacholder, Ph.D., NCI, NIH, DHHS, expressed concern that if the consortium seeks to answer too many research questions, researchers will overwhelm participants. Compromises will be necessary, and ICCCC members will need to be very exact about what it is they want to study. Patricia Buffler, Ph.D., M.P.H., University of California, Berkeley, concurred, adding that it will be important to design the study in a way that focuses consortium resources on the questions that only this type of study can answer.

- Dr. Dwyer noted that the inclusion of biological samples gives a slightly different comparison between the average cohort study and the average case–control study.
- Dr. Linet challenged the notion that researchers will be driven to carry out larger and larger studies. Growing recognition of a wide variety of subtypes of exposures, disease, molecular markers and so forth does not suggest that investigators should conduct larger studies. Rather, this complexity suggests that researchers should approach the questions differently. A combination of mechanistic, cohort, and case-control studies may be needed. For instance, certain types of pathophysiology (such as inflammation) may be important in the etiology of several types of cancer, cardiovascular disease, and other types of chronic disease. There may be common pathways with different outcomes, and only a combination of studies of different types and designs can elucidate the factors involved.
- Dr. Chanock argued that a large study size is crucial not just to allow for subdivision, but also to track the timing of exposures and other factors that may be important. He suggested that researchers think of the consortium as a means of lumping rather than splitting. Differences in exposure may determine whether someone is at risk of developing Ewing's sarcoma or osteogenic sarcoma. Leukemia may be the result of a combination of one or more types of exposures among children with specific molecular defects; etiology may only be resolved by examining the relationship between genes and environment. There may be a gradient on both sides. Large studies not only give investigators the ability to focus on subsamples, they also help researchers to see the patterns and formulate hypotheses.
- According to Joseph Wiemels, Ph.D., University of California, San Francisco, *RAS* mutations cross several subtypes (ALL, AML and other cancers). There may be a common association with hydrocarbons. FLT3 mutations are independent of *RAS* and almost never occur together but seem to drive the same pathway. A large cohort study such as ICCCC is necessary to demonstrate that the same type of mutation may be linked to several subtypes of childhood cancer or that the same exposure or genetic polymorphism could cause different forms of childhood cancer.

Childhood Leukemia: Assessment of Time Windows of Exposure to Household Pesticides and Tobacco Smoke

Patricia Buffler, Ph.D., Professor, Department of Epidemiology, School of Public Health, University of California, Berkeley

Dr. Buffler, who was asked to provide examples of exposure assessment methods, discussed the use of time windows of exposure to tobacco smoke and household pesticides based on experience with the Northern California Childhood Leukemia Study (NCCLS). Study objectives were to examine the relationship between environmental exposures and childhood leukemia. Exposures included household pesticides/chemicals, infectious agents, diet and tobacco smoke during critical periods of the child's development (from 1 year prior to conception to 3 years of age). The study includes all childhood leukemias under the age of 15 years, as well as major molecular leukemia subtypes.

Important components of the study include:

- Representative Hispanic population (47 percent)
- Short interval between diagnosis and interview (increases recall of type and timing of environmental exposures)

- Molecular classification of leukemia
- Genotyping of cases, controls, and mothers of cases and controls
- Multi-disciplinary team including pediatric oncologists, epidemiologists, molecular biologists, nutritionists, toxicologists, and industrial hygienists and collaboration with California Department of Health Services (CDHS)
- Comprehensive and detailed exposure assessment with confirmation of self-reported pesticide exposures
- Strong genetic and molecular components—collection of biologic samples for 90 percent of study population (buccal cells, pretreatment blood and bone marrow, archived newborn blood)
- Ability to evaluate environmental and genetic factors simultaneously.

NCCLS genetic and molecular components include:
- Backtracking to birth of chromosome translocations using pretreatment blood and archived newborn blood specimens from cases
- Classification of leukemia into molecular subgroups using pretreatment blood and bone marrow specimens from cases
- Genetic polymorphisms using DNA from buccal cell scrapings from cases, controls and their mothers
- Studies of DNA methylation and *RAS* mutations
- Use of proteomics to study gene expression
- Use of protein adducts as a biomarker of exposure.

Study Design. The original study sample of 9 centers in the San Francisco Bay area was extended to 9 hospitals to include the 18 additional counties of the Central Valley of California to achieve a larger Hispanic population base as well as a broader range of levels of pesticide and other environmental exposures. CDHS played a central role, and the Survey Research Center on the Berkeley campus served as the central data collection point. Within 24–48 hours of diagnosis, researchers were informed of a newly diagnosed case. Hospitals provided peripheral blood and bone barrow for the biologic studies and access to medical records shortly after the acute diagnostic and treatment period. CDHS supplied archived newborn blood (Guthrie cards) as a source of DNA and as a source for back-tracking the diagnostic translocation to see if the genetic event was present at the time of birth. Each case was matched with two control births residing in the same area born on the same day. The buccal cells, peripheral blood, urine, and epidemiologic data were collected for cases and controls and mothers of cases and controls. Data presented were based on 776 incident cases enrolled from 1996 through June 2005. To be eligible for the pesticide measurement component of the study, subjects had to be a newly diagnosed case of leukemia residing in the 35 county study area, less than 15 years of age, living in the 35 county study area at the time of diagnosis, and have a biologic parent able to be interviewed in English or Spanish.

Pesticide Exposures. Sources of a child's exposure to pesticides are numerous in home environments. Previous studies suggest increased risks of childhood leukemia associated with *in utero* and postnatal pesticide exposures, although limited by non-specific exposure assessment, potential selection and recall bias, and small numbers. Preliminary findings from the NCCLS and

a Canadian study with similar design report increased leukemia risks with use of indoor pesticides pre- and postnatally.

Researchers took a three-step approach to assessing environmental exposures:

1. In-person interview. What products do people report using in the home or outside the home from preconception to age 3?
2. Household survey for reliability conducted 6 months after initial interview. What products are again reported and is the product (compound) present in the home? If so, information recorded from container, such as EPA registration number.
3. Measurement. What levels of relevant compounds are present in the home environment as determined form collection of household dust (carpets and window swipes) and air samples?

The study team conducted a thorough inventory of products in the home, scanning the bar codes to identify active agents. Urine and blood samples are also taken to provide a source of environmental biomarker data. Researchers tested the reliability of parental reports by analyzing samples collected from vacuumed debris and window sills. In addition, data were collected on commercial pesticide use in the vicinity of the participants' homes. Reliability studies based on a subset of 133 subjects indicated no significant differences in the reliability of reports by controls and cases. Specifically, researchers saw no evidence of over-reporting in the case population.

While there was strong evidence that *in utero* and post-natal exposures to indoor use of insecticides were critical in the development of childhood leukemia, there was no association with preconception use. Similar analyses conducted for outdoor herbicides showed increased risks with pre- and postnatal exposures. However, numbers are still limited for evaluating the separate roles of preconception and *in utero* exposures. Although the numbers of cses by immunophenotype were small, there were no differences observed in risk by histologic type (ALL and AML) or by ethnic group.

Researchers will continue to refine assessments of environmental exposures by:

- Integrating other sources of pesticide exposure, such as drift from agricultural applications or "take home" chemicals from workplaces of parents
- Measuring levels of selected pesticides in house dust and pesticide metabolites in maternal urine samples.

Additionally, researchers will strive to identify genetic polymorphisms involved in the metabolism of pesticides (such as the PON1 gene polymorphism). Finally, by increasing the sample size with continued data collection up through 2008, researchers will be able to analyze data by type of pesticide, by histologic and molecular subtype of leukemia, and by ethnic group.

Tobacco Smoke Exposure. Investigators tested the hypothesis that: (1) chromosome abnormalities are often the first or initiating events in childhood leukemia, occurring prenatally during fetal development, and (2) additional postnatal events are required for the development of childhood leukemia. Conditional logistic regression with adjustment for household income was performed to estimate the relative risk of childhood leukemia associated with parental smoking.

Paternal preconception smoking was associated with a significantly increased risk for AML,

while the increased risk for ALL was suggestive but not statistically significant. In contrast, maternal smoking during any time period including breastfeeding was not associated with an increased risk of childhood leukemia. However, paternal preconception smoking when combined with maternal postnatal smoking was associated with a greater risk of ALL than being exposed only to paternal preconception smoking.

Discussion. Participants asked for clarification on how samples were collected and how soon after diagnosis sample collection occurred. Dr. Buffler explained that ideally, household samples were collected within 5 months of diagnosis; however, some were not collected until 12–18 months post-diagnosis. Biological specimens from children were collected at the time of diagnosis, and clinical research personnel at the nine pediatric hospitals in the study area faxed the rapid ascertainment form to the UC Berkeley study office within 48 hours. To comply with the HIPPAA requirements, specimens were collected immediately, but the specimens were not released until consent was obtained.

Other topics of discussion included:

- **Timing of maternal pesticide exposure.** Dr. Buffler reported that the exact time of exposure was not known, but it was most likely in the first trimester. Data are available by trimester of pregnancy, but they have not yet been analyzed.
- **Composition of solvents.** Because manufacturers are not obligated to disclose solvents, the data obtained by scanning the bar codes for household products is incomplete. Future use of blood and urinary assays for environmental exposures may provide clues to some of the solvent exposures.
- **Strategies for working with Hispanic populations.** Based on the experience derived from researchers with the California study, particularly the minority supplement that increased Hispanic participation, Dr. Buffler suggested that when working with Hispanic study participants, researchers avoid using a state logo on letterhead; asking about the diagnosis of a family member or friend (disease is stigmatic); using any indicator of authority; or having male interviewers for the in-home personal interview. Additional considerations are the importance of appropriate Spanish translations for all study materials and adequate financial compensation for the time required for study participation, as many potential study subjects are hourly employees.
- **Data for outdoor exposure.** Joëlle Le Moal, M.D., Institut de Veille Sanitaire, asked how researchers obtained data on outdoor pesticide exposures. According to Dr. Buffler, the geographic areas are mapped by types of crops grown and data are obtained through the California Pesticide Use Reporting System, where every commercial pesticide application is registered. Other environmental variables to be included in study analyses are traffic emissions using a metric based on the number of cars over a certain time.

Birth Weight and Childhood Leukemia

Julie A. Ross, Ph.D., University of Minnesota, Department of Pediatrics and the Cancer Center

Based on CDC data, more than 55 percent of the population in the United States is either overweight or obese, and there has been a steady increase in this number in recent years. The prevalence of obesity (body mass index (BMI) greater than 30) is on the rise in Europe as well. According to Dr. Ross, this trend toward obesity is mirrored by a trend toward bigger babies. In

the United States, the number of neonates born with a birth weight greater than the 90th percentile is on the rise. Denmark and Germany show the same trend.

While the causes of adult obesity are understood, the trend toward high birth weight babies is not as well understood. Predictors of high birth weight include:

- Later gestational age
- Male sex
- Higher parity
- Maternal height and pre-pregnancy weight
- Weight gain during pregnancy
- Diabetes
- Hypertension.

The two predictors of high birth weight thought to be most important are maternal BMI and weight gain during pregnancy.

High Birth Weight and Leukemia. At least 30 studies have reported on birth weight and childhood leukemia. (High birth weight is generally defined as greater than 4,000 grams at birth.)

- Eight studies have reported results for "infant" leukemia (less than 2 years at diagnosis, somewhat more extended period than the usual definition of less than 1 year).
- Four studies have reported a significant positive association with increasing birth weight in Wilms tumor. Normally, IGF2 cells express only paternal alleles. In children with Wilms tumors, however, there appears to be a relaxing of this paternal imprint.
- One study reported an association with increasing birth length rather than birth weight.
- One study of high birth weight and risk for leukemia found an increased risk for girls and a decreased risk for boys.
- For other cancers, there are either no or inverse associations with birth weight and cancer, including neuroblastoma, which shows no association. However, low birth weight is a risk factor for hepatoblastoma.

There is a consistent strong correlation between higher levels of IGF1 and higher birth weight. In theory, proliferative stress is due to IGF1. A pre-leukemia clone arising *in utero* is the initial transforming event, followed by a second genetic even. This combination of factors may lead to a high birth weight baby with leukemia. Studies of other genetic polymorphisms include:

- IGF1 (variants 192 base per allele versus -192 bp allele) and low birth weight (Vaessen, Lancet 2002)
- G-protein β3 subunit (maternal) and low birth weight (Hocher, Lancet 2000)
- Peroxisome proliferator-activated receptor (PPAR) gamma (Pro12Ala) and high birth weight (Pihlajamaki, Obes Res, 2004)
- Human leukocyte antigen G (HLA-G) (14bp sequence) and high birth weight (Hviid, Hum Immunol 2004).

The consortium holds the promise not only of helping to answer several questions concerning the association between high birth weight and childhood cancer but for adult malignancies as well. Cohort studies can help explain risk associated with high birth weight both early and later on in life. Moreover, biospecimens collected in cohort studies can be used to identify potential

biomarkers and help to understand relationships with high birth weight. Investigation of possible sexual dimorphism and genomic imprinting will also be possible.

Discussion. Participants asked questions about the design of the studies that drew the connection between low birth weight and liver cancer. Specifically, how did researchers control for maternal conditions that cause *in utero* growth retardation? How did researchers control for the treatment of premature babies? How are researchers exploring factors that may not be associated with childhood cancer but are associated with increased risk later in life? Dr. Ross explained that research is in progress to clarify these issues, but that a large cohort such as the ICCCC may be what is needed to pinpoint the interplay of specific genetic and environmental factors.

Observing that the incidence rates for leukemia have not been rising as quickly as obesity, Sjurdur Olsen, Ph.D., Danish Epidemiology Science Center, Statens Serum Institut, asked if obesity might be an incidental or secondary factor. Dr. Linet explained that dietary factors, which may be the direct link, were very difficult to study. However, because the Danish cohort includes nutritional factors, this cohort may offer some insight.

In response to a question concerning outcomes and high levels of folate, Dr. Ross noted one study that showed changing folate administration changed the coat color in mice. It is unclear what the implications of this may be in humans.

Infections and Childhood Leukemia: Clusters, Clues, and Conundrums

Martha Linet, M.D., M.P.H., Chief, Radiation Epidemiology Branch, NCI, NIH, DHHS

Dr. Linet provided an outline of the early reports supporting the role of infections in the etiology of childhood cancer and described the general outcomes of studies in indirect and direct measures as well as biological measures.

Researchers have examined indirect measures between childhood cancer and infections through a variety of study types:

- Descriptive epidemiological studies (space, time, or space-time clustering; seasonal variation)
- Ecological studies (geographical population mixing, including areas of population growth, such as construction of new towns; parental occupational contact levels; social isolation/population density studies; relationship between socioeconomic status/deprivation and incidence of childhood leukemia)
- Other studies based on indirect measures, including daycare attendance (11 case-control studies show mostly decreased risks with no increased risks but socioeconomic status may confound these results); breastfeeding (19 studies have produced mixed results although 11 studies showed a decreased risk for longer duration of breastfeeding); birth order (in 42 case-control studies, a few showed decreased or increased risk, but most showed no relation).

Direct measures, which are based primarily on questionnaires, include:

- **Maternal infections during pregnancy**. Thirteen case-control studies in leukemia showed mostly no or modest increases. These were based primarily on interview data and in a few instances, medical data, and measured different organisms. A few small cohort studies

showed an increased risk for all cancers.

- **Childhood infections.** Of 25 case-control studies, 12 showed decreased risk, 6 suggested an increased risk, and 7 showed no effect.
- **Vaccinations.** In three trials of BCG vaccination for tuberculosis, there was no consistent relationship between cancer and infections. Studies of other vaccinations have also been inconsistent.

A small number of studies have analyzed serologic measures of a number of different types of organisms in mothers during pregnancy or, for a few, in children postnatally, including:

- Influenza
- Varicella
- Rubella
- Mumps
- Epstein-Barr virus (EBV)
- Human herpesvirus (HHV) 6
- B-19 parvovirus.

Case-control studies of serologic measures have been limited by the use of non-standardized tests and because the temporal sequence unknown. Other researchers have searched for viral genomic sequences in DNA using:

- JC, BK, SV40 polyoma viruses
- EBV, HHV-6, HHV-8 herpes viruses
- TT virus
- Mycoplasma pneumonia.

Samples evaluated for the presence or absence of viral sequences in DNA have included leukemia cells, bone marrow, and Guthrie cards.

While researchers have been able to suggest some correlations between exposure to infection and childhood cancer, these conclusions have been limited by the small number of samples (n=20–94 patients evaluated), different testing procedures (only two were pre- and post-treatment), and the unknown temporal sequence.

Evidence suggesting an infectious cause(s) of ALL includes:

- *In utero* translocations are an insufficient cause of childhood ALL making postnatal event(s) necessary.
- A growing body of evidence supports a delayed/reduced infection or rare abnormal immune response to common infections during infancy/early childhood (decreased risks of childhood ALL are associated with daycare attendance, breastfeeding; space-time clustering studies consistent with delayed infection hypothesis).

Based on the evidence suggesting a link, though falling short of firmly establishing a causal connection, Dr. Linet suggested that a cohort study could be helpful in providing data to answer questions about ALL and infection:

- Could *in utero* translocations be initiated by infection? Prospective studies of pregnant women that document occurrences of infection and studying chromosomal aberrations in

cord blood could answer this.

- Are additional epidemiological studies required to examine daycare, breastfeeding, vaccination and ALL? Prospective mechanistic studies might be helpful to determine the inter-relationship of infections, modulation of immune system function, and potential carcinogenic promoter effects in exposed versus unexposed children.
- Could an infection be the event that leads to cancer in children who have an AML-1 translocation? One percent of all children have an AML-1 translocation. This is one hundredfold larger than the incidence of childhood leukemia. Could an infection be a factor that determines which children with AML-1 translocations go on to develop cancer?

Discussion. Several participants suggested there may alternative theories than those posited in the literature. For example, an abnormal or programmed response in gene repair may play a larger role than the infection. Anne Louise Ponsonby, Ph.D., Murdoch Children's Research Institute, suggested that researchers consider the temporal sequence of events and natural history. For example, she noted that people with multiple sclerosis (MS) are more likely to be serum positive for EBV, which is uncommon in infancy, but less likely to be serum positive for herpes simplex virus 1, which is commonly acquired in infancy. Perhaps the evolutionary sequence of environmental viral exposure also requires consideration. People at risk of MS might be exposed to infant infections and therefore primed to cope with other adverse infections. Dr. Linet relayed that preliminary data from the large nationwide UK childhood cancer study showed more quasi-serious early infections. These were often unusual infections, such as fungal infections—evidence that the immune system may be suboptimal.

Participants discussed the expense of collecting and analyzing biological samples. Dr. Sjurdur Olsen noted that cytokine, chemokine, and other immune function measures related more strongly to TH1 or TH2 are being collected in a subset of the population to determine asthma risk. Similar use of subsets might be used for select alleles. Dr. Linet noted that this would be an area for discussion. An increasing fiscal emphasis on consortium-based research suggests that more funding may become available. Dr. Wiemels added that the individual cohorts appear to be collecting appropriate samples, allowing follow-up to childhood cancer translocations types of studies. He suggested that adding a genetic component may actually be cheaper than collecting some of the other data.

Genetics of Complex Diseases in Pediatrics: Aren't They All?

Stephen Chanock, M.D., NCI, NIH, DHHS

Observing that the face of genetics has changed in the past 4–5 years with the availability of the complete genome sequence, Dr. Chanock suggested that there remains tremendous untapped potential to better understand pediatric cancer. Examining genetics is, in many ways, easier than examining environmental factors. The purpose of the presentation was to clarify the role that genetic analysis can play in an international cohort study.

There is a wide spectrum of germ-line genetic variation, from common diseases with common variants (such as diabetes, heart disease, breast and prostate cancer) to more rare genetic variants occurring in common diseases and rare variants observed in rare diseases. It is now known that there are approximately 8–10 million SNPs across the genome with a frequency of roughly 3

percent or greater in one or more populations. The vast majority of these are silent. There are approximately 50,000–250,000 SNPs that change either the protein itself or, more important, the actual expression. It is becoming more evident in other diseases (for example, diabetes) that the important factor is not the structural polymorphisms that change the structure of a protein; it is the polymorphisms that change the amount of the protein. The same is likely to be true of cancer.

There have been two parallel approaches to examining genetic determinants of disease. One approach looks at markers to try to understand the genetic variation patterns of a particular region (for example, SNP maps). The other approach, which has been driven by gene-environment interactions, examines candidate genes that play an important role (for example, genes that are important in the metabolism of cigarette smoke). These two approaches are getting closer.

Common diseases have been the focus of SNP studies. With gene sequencing, researchers are just beginning to see how rare variants identified in multiply affected families contribute to common diseases. For example, in studying lipid genes, researchers have observed that there are many uncommon variants that appear to cluster in individuals who have particular lipid profiles. In thinking about pediatric cancer—a relatively rare disease—should researchers focus more on SNPs or on sequencing?

For larger studies such as the ICCCC, it is important to note that no SNP study stands alone. They are determinants of risk, requiring replication. Functional correlates provide plausibility, but this is looking at markers only. There are very few SNPs that have real functional data. There are many predictions, but in terms of demonstrating a functional attribute, researchers are only just beginning. Tools for association studies include:

- Annotation
- Db-SNP
- Re-sequence (SNP500 & Seattle SNP)
- HapMap (genotype only)
- Analytical approaches
- Single gene/SNP (multi-test challenge)
- Multi-locus analysis (preliminary).

Custom genome platforms include custom SNP platforms for single SNPs (TaqMan), multi-plexing (Sequenom, SNPlex), and high throughput (Illumina). Whole genome platforms are costly, ranging from $100,000 to $500,000. These include Affymetrix, Illumina, and Perlegen.

Osteosarcoma. To illustrate the potential for use of genetics, Dr. Chanock explained how researchers were applying some of the tools and knowledge to osteosarcoma, the most common malignant bone tumor in children and adolescents with a peak incidence during or shortly after an adolescent growth spurt. Researchers are beginning to understand the genetic complexity of this relatively rare disease. Studies of genetic variation and osteosarcoma risk have demonstrated an association with the Fok 1 polymorphism in the Vitamin D receptor (OR 1.78). There was no association with estrogen receptor or collagen 1 alpha 1 SNP. There is an association with the tumor necrosis factor (TNF)-alpha-238 allele but the TNF-308 allele was associated with a notably decreased risk (RR 0.18). In examining the relationship between growth-related genes

and osteosarcoma, researchers have found:

- Insulin-like growth factor pathway (mitogenesis; growth and development *in utero* through adolescence)
- IGF1 levels associated with risk for breast, prostate, colorectal, and lung cancer
- IGF2 over-expressed in many tumor cell lines
- IGF2R imprinting plays a role in many cancers.

These findings suggest there may be an association between genetic variation in growth pathway genes and osteosarcoma risk. Other possible associations might be explored through:

- Additional fine-mapping of the risk SNPs
- Further analysis of the gene for potential protection SNPs
- Follow-up on IGF1R and IGFALS

Breast and Prostate Cancers. Breast and prostrate cancers illustrate the genetic complexity of relatively common diseases. Steroid hormones and IGF pathways appear to be relevant to both types of cancer. These cancers will be studied through a whole genome SNP scan and will focus on "enhanced regions," 15 candidate regions in the family linkage study.

Future Considerations. Future considerations include:

- Genetic analysis (SNPs versus sequence)
- Combined cohorts (lumping versus splitting)
- Population genetics.

Discussion. Discussion points focused on clarification of specific points, the implications of SNP versus sequencing studies, and the prohibitive costs of collecting biological specimens. Dr. Chanock suggested that if the study is designed properly, the assay should only have to be done once. He emphasized that as the technology matures, the costs are likely to decrease. What may seem prohibitive now may be reasonable in 2 years. Dr. Winn noted that information from the breast and prostate cancer consortium will be posted on the Web site.

Perspectives on the Day

In a general discussion of the day's session, participants discussed the importance of deciding early on if a cohort is a good way to study childhood cancer. Because many of the cohorts do not have protocols for cancer studies, it will be important to decide how to introduce new protocols. Other details that will need to be worked out include:

- How should relevant information be collected (for example, case finding, tissue samples, or case-control studies to collect specific information)?
- Many existing consortia are further along or at different stages. What can be done to accommodate these differences?
- Are there other conditions that might benefit from the consortium approach?

Dr. Scheidt maintained that it may require 2–3 meetings to determine whether to join forces or to clearly understand potential benefit to any study/organization specifically. The hypotheses discussed at the meeting were a sample of many possible questions that could be evaluated, but they do represent scientific questions that have resisted resolution.

Dr. Scheidt made the following observations about the National Children's Study:

- There are 101 locations of the National Children's Study that are the first stage of more than 1,000 primary sampling units. Statisticians at the National Center for Health Statistics will conduct a multi-stage, clustered probability sample.
- The second stage will involve oversampling and school or geographic boundaries as determined by local areas.
- Study centers are the 40 sites that will carry out studies in two or three places; these may include health maintenance organizations, hospitals, health departments, and so on, or teams of these types. They will be awarded through a competitive contract to organizations that are able to engage the community and retain study participants.
- The National Children's Study is charged with looking at exposures early in pregnancy. Women will be enrolled as early as possible in pregnancy with a sample selected from households who are trying to or at risk of getting pregnant.

Data Available Through Individual Cohorts. Participants reported on the status of their individual cohorts and the data that might be of use to the consortium.

- Although the childhood leukemia cases that arise in the Chinese cohort may lack a precise leukemia subtype diagnosis, there are rich data on the entire health continuum from preconception on. This includes the mixed socioeconomic north and south provinces. Moreover, the cohort is large and longstanding with well-trained staff. Plans are underway to enroll 300,000 newly married couples. Researchers are considering a new cohort addressing exposure and outcomes.
- An Asian consortium has been exploring working to establish standardized instruments. A series of meetings have been held, and the Korean government promised support for a 25,000 person cohort, but 100,000 may be possible. It will be interesting to hear how other cohorts have worked through issues with their respective governments and funding organizations. Follow-up, controls, and case ascertainment may all be different in different countries.
- Not all cohorts are collecting DNA.
- Guatemala, France, and Jerusalem, and China have multi-generational studies. Guatemala focuses on diet and follows emigrants. Many U.S. cohorts are not sufficiently multi-ethnic.
- Approximately 20,000 children are enrolled in the French health and environment study. This study did not collect cancer data, but this might be possible within a cohort consortium environment.
- The UK cohort did not plan on including cancer, but it could be added. Half of the births in the UK Bradford cohort are expected to occur in Asians, especially Pakistani.

Study Design and Methodological Issues. Several participants made recommendations about specific aspects of the study design.

- Dr. Sjurdur Olsen noted that it would be important to develop a core protocol that can accommodate studies that are at different stages. It might be most efficient to select one cancer type and work through that protocol. He added that in the Norwegian study, researchers used a Danish food frequency questionnaire as a basis for developing their own questionnaire. He asked if China could collect a compatible dietary questionnaire. Li Zhu, M.D., M.P.H., Peking University Health Science Center, noted that there is a food frequency question that is in use.

- As Jenny Pronczuk, M.D., World Health Organization, observed, developing countries are at a certain disadvantage despite having well-trained scientists. A cohort study is feasible, and the environmental theme is valuable. Technologies are reaching developing countries. Governments change and priorities can change quickly. An international focus may help ensure continued support. Cancer classification is very standardized worldwide compared to asthma and wheezing definitions, making it feasible to examine childhood cancer.
- Danuta Krotoski, Ph.D., Ph.D., NICHD, NIH, DHHS, reported that she has been working with an international group to develop a common protocol for developing countries. This may be useful for ICCCC. There have been a number of countries involved in thinking about this.
- Dr. Linet noted that in addition to the core protocol suggested by the group, ICCCC members might also think about broad possible mechanisms (for example, obesity). She further suggested that senior investigators involve more junior investigators.
- Dr. Seminara suggested that information on existing cohorts and the protocols therein be assembled. Some cohorts may be interested in adding components from other cohorts. If there is a group of common protocols, the individual cohorts can select based on their interests.
- Dr. Olshan suggested that ICCCC members prioritize research items. What are the most important questions that can only be answered by the consortium?

Scientific Questions Arising from Day 1 Presentations and Discussion

Deborah M. Winn, Ph.D., Division of Cancer Control and Population Sciences, NCI, NIH, DHHS

Dr. Winn outlined the prospective research components based on the previous day's discussion.

Prospective diet/energy balance measurements that might be included in ICCCC are:
- Correlations with familial obesity other parental factors
- Folate
- Correlation with child obesity/overweight
- Antioxidants and oxidative stress.

Prospective measurement of prenatal, perinatal, childhood environmental factors include:
- Child care
- Breastfeeding
- Immunization, infections, and sequential measurements of serologic markers of infection
- Underlying immunity
- High-throughput genetic characterization to understand polymorphisms in subgroups.

Significant benefits to childhood cancer cohort consortium include:
- Multi-generational studies
- Emerging opportunities for whole genome association studies
- Opportunities for sharing of issues, challenges, and solutions in executing cohort studies, and learning from each other
- Major exposures potentially related to childhood cancer can best be measured prospectively and less well in case-control studies such as measures of immune function, infections,

chemical/physical environmental exposures, carefully examining specific windows of exposure and susceptibility
- Potential for better understanding immunity, DNA repair, and other processes and factors that may lead to a range of outcomes or explain why a range of infections, for example, may influence risk of childhood cancers
- Potential in some of the cohorts with data collection in the early stages or about to begin include a childhood cancer component based on a core/common/standardized protocol (this must not adversely affect the primary goals and priorities of the cohort)
- Exposure variability and genetic variability across countries will pinpoint environmental risks
- Potential for a complementary relationship between case-control and cohort studies—they can inform each other
- Tremendous potential for examining translocations evident at birth, with the future potential for cancer-related screening
- Useful to have cohorts from the developing world: there are differences in types of translocations for example (India versus other countries)
- Could be a model for other rare diseases
- Provides an excellent opportunity for examining pathways shared by many diseases
- Genetic admixture in the various cohorts may help inform patterns of risk.

Factors suggesting the potential success of the consortium include:
- There is strong experience among some of the meeting participants in consortia development and a number of successes to date.
- Cancer classification is well-standardized, especially compared to many childhood diseases like asthma which must rely on poorly defined symptoms such as wheezing.
- Childhood cancer is a rare disease—a consortium is needed.

Teaching an Elephant to Dance: Studying Childhood Cancer in the Collaborative Perinatal Project

Mark A. Klebanoff, M.D., M.P.H., NICHD, NIH, DHHS

Dr. Klebanoff opened his presentation with a disclaimer that the Collaborative Perinatal Project (CPP) was never intended to study childhood cancer. Rather, it was a large cohort study of neurological disorders. In fact, CPP did not specifically record childhood cancer occurrence.

Vitamin K. The impetus to examine the data came from published reports of an association between vitamin K and childhood cancer. Concern that these papers would spark a vitamin K "panic" as happened with spermicides, Bendectin®, and the pertussis vaccine, Dr. Klebanoff pursued the CPP data, which spanned the era when neonatal vitamin K use began to become standard practice in the United States, as it remains to this day. The CPP data contained meticulous studies of all aspects of the cohort's care. The study included data on all drugs administered to mothers and children as well as medical records, patient evaluations, and parent interviews.

Researchers looked closely at any file that noted neoplasia, followed by a notation of the specific cancer based on some form of biopsy. There were 52 probable cases found in live births. Among

live births, life-table probability of cancer by 90 months of age equals 1.1 per 1,000, a figure virtually identical to Surveillance Epidemiology and End Results (SEER) data (1–96 months=1.2 per 1,000), further indication of the reliability of the data.

To confirm exposure to vitamin K, Dr. Klebanoff consulted original microfilms, which recorded vitamin K use. This was supplemented by information from original treatment hospitals. Researchers found no association between vitamin K and any childhood cancer generally, or leukemia specifically. Several subsequent studies in Europe confirmed the lack of association.

Maternal Smoking. Dr. Klebanoff and colleagues found no association between childhood cancer and maternal smoking during pregnancy. This may be due to increased mortality in the smoking group of fetuses who would have developed cancer had they survived, but there is no way to ascertain this in CPP data, or probably in any other data either.

SV40 Virus and Cancer. Simian virus 40 (SV40), a monkey virus that contaminated polio vaccines prior to 1963, has been reported to cause malignancies in laboratory rodents. Some studies, though not all, have found SV40 DNA sequences in some human tumors, including pediatric brain tumors. CPP spanned the time from when polio vaccine was contaminated to when it was not. It was also common practice to give women a dose of polio vaccine at their initial prenatal visit. The children of women with no vaccine exposure, or who received vaccine after 1963, had no increased risk of cancer, but for the contaminated period, there was a rise in two types: hematological malignancies and neural tumors. Moreover, the pre-1963 vaccine was associated with conversion by virus-like particle (VLP) assay but not plaque assay (VLP assay cross reacts with other viruses). The conclusion was that receipt of pre-1963 vaccine was not strongly associated with seroconversion and this, in turn, was interpreted as not supporting a role of SV40 in childhood cancer. The reasons for the association between vaccine receipt and cancer remain unclear.

Neonatal oxygen supplementation and cancer. Several case-control studies have found an association between neonatal oxygen exposure and childhood cancer, mainly leukemia and hepatoblastoma. In clinical trials comparing use of 100 percent oxygen to use of room air for resuscitation of newborns, even a few minutes of oxygen exposure caused increased markers of oxidative stress at 28 days of age. The CPP study included delivery room observers with stopwatches and no clinical responsibilities to record all events. Based on 48 cases of cancer diagnosed after the first week of life, it was determined that the hazard ratio for more than 3 minutes of oxygen administered in the delivery room was 2.87 (1.46–5.66) and the hazard ratio for cases after 1 year of age was 2.00 (0.88–4.54).

Advantages and Disadvantages of Cohort Studies. Because exposure data are collected before outcomes are known, cohort studies are less susceptible to information bias than case-control studies. Moreover, the population is defined at the beginning, making it less susceptible to selection bias than case-control studies. Prospective cohort studies also offer the ability to collect information beyond that readily available (an advantage over both case-control and historical cohort studies where researchers are limited to what is in the records) as well as the ability to tailor biospecimen collection to specific hypotheses. Finally, the prospective cohort study offers the best design for the natural history of a disease.

Cohort studies hold certain disadvantages, most notably, their inefficiency. There are tremendous resources expended to collect and store data and specimens from participants, most of whom will never develop an outcome of interest. The less common the outcome of interest, the more inefficient a prospective cohort becomes. Cohort studies may also miss acute, short-latency effects of an exposure (this may explain why the Nurses Health Study may have overstated the cardiovascular benefits of estrogen). Biomarkers of exposure or susceptibility (particularly genetic polymorphisms), a reported advantage of cohort studies, can also be measured in non-cohort studies. Simple, inherited polymorphisms in the child could easily be detected by accessing stored metabolic screening blood spots (Guthrie cards). Polymorphisms in the parents could be obtained at any time in a specifically designed case-control study.

Despite these limitations, cohort studies also offer insight into factors that would be difficult to report retrospectively about pregnancy or early childhood after a lag of several years, including:

- Minor illnesses
- Psychosocial stress
- Over-the-counter medication use (medications available only by prescription are probably noted in records)
- "Alternative" medicines and treatments
- Home pesticide use, other "trivial" exposures
- Events early in the child's life (for instance, minor illnesses) that would not be noted in pediatric records.

In exchange for this added information, there is a tremendous cost in both study efficiency-resource use and limited power. If the ICCCC can acquire the data needed with minimal addition of data collection, and if the power is acceptable, it may be worthwhile. However, designing a cohort study specifically for childhood cancer, or even listing it as a primary objective of other cohorts, may be unrealistic. It may well make a valuable secondary outcome.

Discussion. Noting that in the CPP study there were 50,000 births, Dr. Dwyer asked if the National Children's Study population of 100,000 would be insufficient. Dr. Klebanoff indicated that this was not necessarily the case. If additional items can be added without exhausting the sample, it is worth doing as a secondary outcome. Approximately 100 cases would be the minimum to make the addition worthwhile.

Dr. Linet pointed out that Dr. Klebanoff's research on vitamin K demonstrated an important contribution that large cohort studies can make. Because a large data set was available, it prevented a public panic. Having cohort data available can prove invaluable to quickly respond to suggestive research findings from case-control studies.

Statistical Issues in Cohort Studies Assessing Postulated Risk Factors for Childhood Cancer

Sholom Wacholder, Ph.D., NCI, NIH, DHHS

Dr. Wacholder observed that while cohorts have substantial advantages, they also pose important constraints and tradeoffs. Case-control studies offer:

- Tremendous economy for rare diseases
- Ease of collecting disease-related biospecimens (fresh tissue, molecular pathology)
- Dedicated questionnaires with no competition with etiologic factors for other diseases that can be far more focused on putative etiologic factors.

Cohort studies offer:
- Prediagnostic biospecimen collection and possibly identification of predisease
- Serial biospecimen collection providing suggestions for early detection and screening and showing how biomarkers change over time
- Prediagnostic questionnaire collection reducing non-differential misclassification closer to time of exposure and serial collection for time-dependent exposures
- Ability to capture changes in exposure prediagnosis and closer to real time for better accuracy
- Reduction of differential misclassification (no rumination bias or distortion of reporting caused by disease)
- Ability to evaluate multiple diseases.

Lumping and Splitting. Cohort and case-control studies can be thought of in terms of "lumping" and "splitting," respectively. Broad categorization of diseases (lumping) requires a smaller sample size. Splitting allows for investigation of etiologic heterogeneity. Splitting within a cohort study is possible if there are sufficient details available for each cohort; splitting is advantageous if there are sufficiently etiologic heterogeneity specified subsets. Increased power to detect an association from splitting is realistic if the etiologic factor is strong in the subset *and* the etiologic factor is unrelated—or very weakly related—to other subsets. Chatterjee et al. have developed a method to cross classify disease in a multivariate way to identify etiologic heterogeneity one disease classification margin at a time. This could be of value in cohort studies if it is possible to collect detailed clinical and molecular data on each case.

Nested Case-Control Studies. Nested case-control (or case-cohort) studies offer an economical compromise. There are no shortcuts in terms of case ascertainment to this approach; however, it does offer the cost savings of only doing assays on identified cases. Such biospecimens may be previously collected or newly collected as cases accumulate.

A controversy in epidemiology is whether case and matched controls need to be in the same batch for molecular assays. Here again, there is a trade-off between bias and cost. There is a much greater cost for matching.

Data Missing by Design. Dr. Wacholder suggested that "data missing by design" might be a way to reduce costs. In explaining this concept, Dr. Wacholder said that data missing by design:
- Can be questionnaires, biospecimen collection, and biomarker evaluation
- Involves oversampling and undersampling for study components with probability of inclusion based on collected information and adheres to a sampling plan
- Enables missing data to be filled when appropriate
- Saves money and lessens burden
- Can be a "two-phase design," which, with proper statistical analysis, can assure lack of bias
- Requires careful design and fieldwork and special statistical analysis (see Weinberg and

Sandler for a simple example)
- Assumes that it is better to have best exposure assessment on a sample than poor exposure assessment on everyone.

In a nested case-control study in which data are missing by design, the sampling is based on disease status. This saves costs and time because not all participants are asked all questions, and biospecimens are not collected from every participant. Similarly, the partial questionnaires involve asking all participants broad questions and asking more detailed questions to a subset usually determined without regard to response to other questions. This reduces the burden on participants. A variant design can overcome the inability to collect elaborate details from everyone in the master questionnaire. Only a fraction of the "exposed" (for example, the smokers or homeowners who applied pesticides in their kitchen) are asked to provide additional details on timing or details of exposure (age started smoking; frequency, brand, mode of administration, name, brand, mode of administration, amount, frequency of pesticide applied). Because of random assignment, the results from this subsample can be extrapolated to exposed cohort members who were not asked about all the details.

Factors which can be used for determining who will be asked certain questions at all or who will be asked questions in more detail include:
- Demographics, like place of residence or age
- Family history of disease
- Social class
- Crude predictor of pesticide exposure from questionnaire
- Ever smoked 100 cigarettes
- Bottom 10 percent of birth weight.

Intense Subgroup for Serial Biomarker Follow-Up. Perhaps based on likelihood of exposure or endpoint, a subset of the cohort can be identified for more intense follow-up, with shorter intervals between consecutive questionnaire and, more important, biospecimen collection. For example, a cohort to study determinants of low birth weight (gestational cohort) could be followed with monthly biospecimen collection and clinical evaluation.

Consistency of Definitions and Techniques. This is vital and includes disease, exposure, and molecular assays and may require round robins for new biomarker methods.

Combining and Preserving Data. An obstacle to sharing data after the fact is the difficulty of creating combined data sets across cohorts. Carefully documented informatics is useful for:
- Study management
- Sharing
- The future (commitment to long-term biorepositories).

In order to ensure the long-term preservation of biospecimens, Dr. Wacholder stressed the importance of using biospecimens wisely. Specifically, he recommended that researchers avoid using biospecimens:
- To validate assays
- With assays that are not yes validated

- With assays of poor quality.

Discussion. In response to a participant question, Dr. Wacholder suggested that counter matching is a special case of two-phase design, and that some counter-matching analyses do not make use of all available information. A question was raised about assays based on pooled biospecimens, which can save research costs. Dr. Wacholder said that pooling for genotyping can be problematic. There are issues of quantitation (making sure that there were equal amounts of DNA from everyone). He acknowledged that there is work in pooling for other kinds of assays that may prove helpful.

Dr. Jørn Olsen, M.D., Ph.D., suggested that it would be possible to acquire blood samples from fathers. DNA from both parents has been successful in looking at genes related to infertility, spontaneous abortion, and very early perinatal events using the transmission disequilibrium test (TDT). If fathers are not available, data can be filled in with siblings.

Epidemiological Studies of Childhood Cancer Methodological Issues

Martha Linet, M.D., M.P.H., NCI, NIH, DHHS

The goals of studies of childhood cancer are to characterize and quantify childhood cancer risks associated with exogenous, endogenous, and genetic factors, alone and combined. The ultimate goal is risk reduction.
The key issues in exposure assessment include:

- Study design (case-control or cohort; retrospective, prospective, or combined)
- Exposure measurements (feasible, direct versus indirect, repeated, long-term)
- Validity (compare exposure assessment with the gold standard)
- Reproducibility (repeatability over time, geographic region, and across subjects).

There are two major types of data collection strategies used in retrospective exposure assessment: interviews and post-diagnosis measures. The assumptions are that post-diagnosis measures correlate with prediagnostic exposures. This is a major limitation of case-control studies. Retrospective exposure assessments have to be restricted to children living in the same home during exposure windows and post-diagnosis. This requirement can lead to sampling bias because people who move tend to be of a lower socioeconomic status than those who do not move.

Dr. Linet suggested that the childhood cancer consortium would offer the opportunity to push beyond the clear limitations of case-control studies as described below.

- **Long-term follow-up.** The Japanese atomic bomb survivors cohort study, which included a follow-up for more than 55 years, quantified excess relative and absolute risks by age at and time since exposure. A notable finding was that the highest risks were seen in those exposed at the youngest ages. It is the largest study of radiation-exposed children to quantify lifetime risks after single acute exposure. Slightly more than one half of the population is still alive, and it is estimated that the largest number of radiation-related cancers will be occurring over the next 20 years. This suggests the importance of lifetime follow-up.
- **More precise data.** There is an established association between pediatric CT scans and the risk of childhood malignancies. A cohort study offers repeated and validated exposure data.

Exposures can be estimated based on radiation doses.

Key considerations in analyzing multi-factorial carcinogenesis include:

- **Initiators and promoters.** Case-control studies make it easier to study promoters while with cohort design, it is easier to assess early and late carcinogens. A combination of studies may be needed to understand the roles of different variables.
- **Time windows.** Measurements obtained closer to postulated temporal occurrence are more reliable.
- **Endogenous changes.** Studies should account for changes over time in physiology (immune function), hormonal milieu (puberty), organ size (bones), size, and function (brain).
- **Prenatal chromosome rearrangements.** Characteristic leukemia chromosome rearrangements present at birth (Guthrie card study) is necessary but not sufficient to cause cancer as evidenced by the high percentage of evident translocations that do not have an outcome of cancer. (Cord blood has one hundredfold higher translocations at birth than incidence of pediatric leukemia.) Studies have postulated a variety of promoters (such as infectious agents, nutrients, chemicals, and radiation) but some of these are weak effects.
- **Confounding biases.** Defined as circumstances in which an estimate of exposure effect is distorted when mixed with effect of extraneous factor, confounding biases can modify a risk estimate upward or downward. It is difficult to distort risk estimates substantially because most confounders produce minor effects. However, a confounder may be more important for a weak effect.
- **Effect modification.** Effect modification—defined as a change in magnitude of an effect due to a third variable (after exposure and disease)—can combine with exposure to increase or decrease risk in whole populations or subgroup. By way of example, Dr. Linet pointed to leukemia following Down syndrome versus *de novo* childhood leukemia. Subgroups may have different etiologies or may be more sensitive to the same risk factors. A nested case-cohort study may be a way to tease out these differences.
- **Variation by subgroups.** Risk factors for childhood cancer vary by subgroup (such as age, sex, race, ethnicity, or genetic variables). The challenge lies in determining if subgroup differences reflect a real factor or are simply a chance finding. For rare subgroups, case-control studies may not provide sufficient numbers; however, combining cohorts may provide a way forward.

In conclusion, Dr. Linet suggested that when combining cohorts, research should consider positive components (for example, prediagnostic information and temporal sequence) and negative components (for example, limited numbers of childhood cancers and biases not eliminated). The unique methodological features of cohort investigations enable an exploration of risk factors and mechanistic clues from case-control studies as well as economical nested case-control or case-cohort investigations (such as lab studies of fewer controls).

An Example of Pooling Child Cohort Data to Examine Early Life Environmental Determinants of Disease

Anne Louise Ponsonby, Ph.D., Department of Epidemiology, Murdoch Children's Cancer Center

Dr. Ponsonby outlined a preliminary example of cohort data pooling and analysis, using data from two studies that were not designed in tandem. Data sources included the Avon Longitudinal Study of Parents and Children (ALSPAC) (from J. Golding 2005, personal communication) and

the Tasmanian Infant Health Survey (TIHS). Sample information from the Avon and Tasmanian studies as well as the combined data are summarized in Table 4.

Table 4. ALSPAC and TIHS Combined Cohort Study

	ALSPAC Sample	TIHS Sample	Combined Sample
Sample	Residents expecting a baby 1 April 1991 to 31 Dec 1992	Tasmanian live births at higher risk of SIDS (top quartile)	
Enrollment period	1990–1992	1988–1994	
Pregnancies	n=14,541		Perinatal variables: n=24,770
Postnatal	Postnatal survivor past year 1: n=13,988	Postnatal (early): n=10,627	Postnatal (early) variables: n=23,923
Childhood asthma status and other variables	Child asthma status: n=8,342	Child asthma status: n=863 (subgroup follow-up)	Child (6–7 yrs) variables: n=9,165

There were 94 variables in selected ALSPAC subset of requested data, 555 variables in TIHS data, and 18 common variables.

ALSPAC had data on infants known to be in a smoker's presence at age 6 months, while TIHS data were on how often others smoked in the same room as the infant at one month of age. These were not directly comparable. The common variable was if the infant had ever been exposed to tobacco smoke. Is postnatal smoke exposure associated with increased child asthma independently of maternal smoking during pregnancy? Both studies suggest a non-significant tendency for infants in a smoke-free postnatal environment to be at lower risk of asthma after adjustment for antenatal maternal smoking. In the merged data set, the larger sample size has increased the statistical power to detect an effect of postnatal smoke exposure, independent of maternal antenatal smoking. As for childhood leukemia, child asthma is a heterogeneous disorder. It is likely that subgroups differ not only in phenotype but etiology. Some examples of this issue were shown.

Participants raised several points:

- In response to a question about how frequently parents were asked about infant exposure to smoking, Dr. Ponsonby noted that the ALSPAC study asked during pregnancy and at 6 months. For TIHS, data were collected on pregnancy at day 4 and infancy at 1 month postnatal age. Data are also available for subsequent periods but have not been analyzed.
- Dr. Jørn Olsen noted that there are two ways of looking at combined data: random effect model or fixed effect model. In this example, all data were pooled. Dr. Ponsonby observed that previous pooled adult cohorts had used matched analysis for case-control samples nested in cohorts, fine stratification with pooled odds ratios, or proportional hazard modeling. Because the analysis here is preliminary, researchers have not yet established definitive

strategies for analysis. Dr. Dwyer noted that in order to understand the data sets, pooling is necessary.

- Dr. Sjurdur Olsen observed that the Norwegian and Danish cohorts have done considerable pooling. When there are many cohorts, it is necessary to conduct an estimate basis for each cohort.
- Dr. Wacholder made several methodological observations. The Tasmanian study used a design that could be considered "two-phase." In selecting the individuals, it was clear that there was a deficit of females. Researchers enrolled those at greatest risk: boys and low birth weight babies. By using the "missing by design" strategy, it will be possible to reconstruct cohort characteristics based on estimates from the samples. This will not be as precise as a random sample but it will be valid nonetheless.

Possible Mechanistic Studies for the ICCCC

R. Julian Preston, Ph.D., National Health and Environmental Effects Research Laboratory (NHEERL), Office of Research and Development (ORD), EPA

Dr. Preston explained that his presentation would provide a framework for predicting risk based on recent EPA guidelines that emphasize the use of mechanistic data. Elements of risk assessment include:

- Hazard identification (such as those from environmental exposure)
- Dose-response for hazard (cancer)
- Characterization of risk
- Identify gaps or uncertainties
- Conduct research studies to improve risk assessment.

Key Events. Key events are developed for animal modes of action (MOA)—a general process that might include the ability of a chemical to react with DNA or the ability of a chemical to react with a receptor. Key events are a set of measurable parameters thought to underlie a specific MOA. As Dr. Preston noted, however, most MOAs and corresponding key events for chemicals have been based on studies with animals. Moreover, children may respond differently to MOAs and key events than adults. In fact, there is a requirement in the cancer risk assessment guidelines to consider sensitive subpopulations (children in particular) to specific exposures. Supplemental guidance for early life exposures states that up to age 2, there is a tenfold adjustment factor; for ages 2–15, the adjustment factor is 3, and for 15 and older the factor is 1.

A starting place is to consider whether children are different from adults in their likelihood of producing one or more key events in response to environmental carcinogens. For the present discussion, Dr. Preston noted that he would focus on DNA-reactive carcinogens. (The same set of key events could provide mechanistic data for other outcomes.) The top five key events include:

- **Exposure of target cells to DNA-reactive species:** This key event frequently requires metabolism, for which there are differences between children and adults.
- **DNA damage and repair:** There is some evidence that links deficiencies in DNA repair with an increased risk of childhood cancer from radiation exposures. These types of studies need to be expanded and shown to be repeatable. Underlying mechanisms need to be characterized. The model for childhood leukemia proposes that characteristic chromosomal

translocations occur *in utero* and a second hit occurs postnatally, possibly caused by environmental factors and inherited susceptibility from polymorphisms in DNA repair genes.

- **Misreplication and misrepair to produce mutations:** The frequency of misreplication on a damaged DNA template or misrepair of DNA damage will determine mutation frequency. Data are not available for establishing this error rate for early life stages but are clearly needed.
- **Mutations in critical genes in tumor target cells:** There is a developing database that provides suggestions of links between specific genetic alterations and childhood tumor formation, in particular for leukemia. Investigators need to establish if these same markers could be used to assess risk from environmental exposures.
- **Additional mutations**: As a result of DNA damage and errors of replication/repair together with enhanced cell replication, additional critical mutations can occur resulting eventually in tumors. Enhanced cell replication in early-life stages could provide a mechanism for enhanced sensitivity of young.

Future Studies. Dr. Preston suggested that future studies might:

- Compare mechanisms underlying key events in animal models (or people) at different life stages following exposures to different chemicals acting through a DNA-reactive MOA
- Consider key events for other MOAs (for example, receptor-mediated, cytotoxicity, and mitogenicity) and whether these could be induced to different degrees in different age groups
- Conduct well-designed studies to investigate tumor induction in young animals for comparison with data from "traditional" 2-year bioassay (there are almost no data on animals, let alone children in early life stages)
- Develop risk models for early-life stages based on these types of mechanistic data.

Discussion. Discussion focused on particular points of mechanistic studies.

- Literature has been based on known strong mutagens and adult tumors with limited *in utero* research. Little has been published explaining paternal exposures based on animal models.
- Dr. Ross suggested that participants consult a review by Chris Wild that lists genes that have not been studied, but that should be studied. Moreover, some of the findings of case studies have looked at genes that are not expressed during fetal development; therefore, it may not be important to look at them.
- Some common childhood leukemias diagnosed at ages 1–5 have hemoglobin rearrangements that appear very immature. This may have occurred during liver hematopoiesis before it moved to the bone marrow. In infants, however, the hemoglobin rearrangements appear very mature and may have originated in the bone marrow during the third trimester of pregnancy. For infants, prenatal events associated with leukemia may actually occur later in pregnancy than they do in other childhood cancers.
- Immediate endpoints may be easily assessed in a cohort study that includes cord blood for pregnancy exposures.
- The cohorts may provide an opportunity to pinpoint childhood origins of adult cancers, such as breast cancer.
- Revised assays have a better predictability than early assays.
- Dr. Linet noted that there are policy implications for the research. If there is a way to identify susceptible subgroups, should these groups be treated differently with regard to allowable exposure thresholds? It is important to remember that for an increasing number of potential

carcinogens, there is no safe dose. Everyone is at risk.

Establishing an International Childhood Cancer Cohort Consortium

Deborah Winn, Ph.D., NCI, NIH, DHHS

Dr. Winn outlined the following important steps required to form an international consortium:

- Development of a written statement of scientific hypotheses is a critical first step. This defines scientific rationale for the existence of the consortia. This statement plan can and should evolve over time (a strategic scientific plan).
- Conduct a preliminary assessment of feasibility of addressing scientific questions.
- Develop a leadership policies and procedures document that describes name of consortia, mission, goals, principles, and governance. Governance issues can include criteria for membership, data sharing policies, publications policies, and process issues. It is important to have a written structure and set of agreements, and stated principles are very important. Stated principles might include helping younger researchers evolve. The document should not be cumbersome or unnecessarily complicated
- Create a name. Examples include the International Head and Neck Cancer Epidemiology Consortia (INHANCE) and Interlymph.

Dr. Winn provided a general model that might serve as a starting point for the consortium. The leadership roles and responsibilities might include:

- Steering committee to provide scientific leadership, rationale for the study, and to set the scientific agenda, and in some instances, to set general policies
- Advisory committee to provide independent scientific advice to the consortium and funding organizations
- Working groups that are small enough to get the work done.

Steering committee composition might include:

- Approximately one to three leaders of specific cohorts (or case-control studies) and representation from a funding organization and/or a data coordinating center
- All leaders of specific cohorts (or case-control studies) and representation from a funding organization and/or a data coordinating center
- Government or funding members
- Working group chairs
- Representation from outside agencies or groups (can be an important component).

Subcommittee types include those that are:

- Based on fundamental administrative functions (for example, publications committee)
- Based on scientific content areas of interest (to develop a specific project)
- Based on technical issues (pathology, data analysis).

Communication strategies include:

- Concentrating work and information exchange within working groups
- Conference calls for policy updates and communication across working groups
- Web site and portals
- In-person meetings to maintain momentum

- Methods to keep funding agencies abreast of progress and needs (funds beget funds).

As a final note, Dr. Winn noted that NCI can and will provide examples of governance documents. Based on experience, the consortium will succeed if it has:

- A commitment from leaders with expertise in specific studies
- A committed, small steering committee composed of persons dedicated to creating the consortium and making it work
- A feeling of enthusiasm and trust among those who make a commitment to join and a willingness to work out issues together.

Based on NCI's experience, funding for consortia meetings has come from a number of sources, including NCI, other NIH offices, and charitable/patient organizations. Specific consortium hypothesis-driven projects are eligible to apply for regular NIH grant funding. Funding from one source can sometimes be used to leverage additional funding from another source.

Discussion. Dr. Linet added that in her experience with Interlymph, researchers have had initial success in acquiring small grants and are currently applying for larger grants. This was in part due to demonstrated progress in their respective research agenda, and in part due to their affiliation with a larger research agenda. As a group, Interlymph has also enjoyed success in acquiring funding to support meetings from NCI, NATO, and the Leukemia Research Fund in England. There have also been expressions of interest from private organizations in the United States.

Participants suggested that soon after the consortium is organized, researchers should distinguish between short- and long-term projects in order to ensure that data are published before the end-points are ready for analysis. For example, researchers could study the children of cancer victims. Dr. Linet suggested studying risk factors (such as age of initiation of smoking) as well as outcomes. There is no need to wait for cohorts to mature. Dr. Dwyer added that the impetus for establishing the ICCCC was that it enhanced existing cohorts without distracting from the individual cohort's agenda. There are many possible studies that could be carried out within the ICCCC, and it will be easier to publish based on an international consortium. One participant suggested that a paper be published announcing the formation of the international consortium and outlining its research design.

Implications of Establishing an International Childhood Cancer Cohort Consortium on Harmonization of Hypotheses

Danuta Krotoski, Ph.D., NICHD, NIH, DHHS

Dr. Krotoski outlined implications for establishing an international consortium in developing countries. International cooperation is extremely important because it can:

- Extend studies to include a broader range of research questions and approaches
- Provide leverage for individual studies by including common measures in longitudinal cohort studies of the environment and children's health
- Provide the evidence base for the development of effective health policy.

Because ICCCC is an international project nested in ongoing and new long-term studies on

children's health with an emphasis on environmental interactions, the collaboration of cohorts from multiple populations will:

- Provide a wider range of exposures (both physical and cultural)
- Provide a basis for identifying genetic and cultural factors in children's health
- Address other health conditions
- Provide a model for pooling data from ongoing and new cohorts for other outcomes
- Aid in the identification of common outcomes, measures, and hypotheses.

Other international collaborative groups include:

- **The National Children's Study's International Interest Group (IIG).** Established in 2002, IIG collaborated on the WHO Working Group on long-term studies of the impact of the environment on children's health in developing countries. IIG identified the feasibility of long-term cancer studies in developing countries and establishing common core protocol for all international long-term cancer studies. Currently, IIG is working with parallel cohorts in Europe and planned studies in Mexico, Thailand, and South Africa. An NICHD/NCI/EPA initiative is underway to pool data from new and existing long-term studies to identify causes of rare conditions in children, such as childhood cancers.
- **WHO Working Group Long Term Studies on Environmental Threats to the Health of Children in Developing and Industrialized Countries.** This group is working to stimulate studies that will identify environmental threats to children's health and to ensure comparability of data collected across countries through a common protocol.
- **WHO informal consultations co-sponsored by the National Children's Study's IIG, CDC, and EPA.** The first consultation identified the feasibility of undertaking long-term studies in developing countries, as well as the challenges and benefits for countries, health care systems and the children. Since then, there have been three subsequent consultations to identify key issues and common hypotheses; identify core hypotheses on respiratory effects, pregnancy outcome, neurodevelopment, growth, birth defects, and cancer; and to develop a preliminary set of measurements as a matrix to be used in the preparation of the core protocols for studies to be undertaken in low- and middle-income countries. This matrix will be made available on the IIG Web site.

Recommended collaborations include:

- A multi-country approach to provide sufficient size to facilitate the investigation of the less common conditions
- Internationally agreed systems for data collection, sampling and storage; analytical and measurement methods; maintaining copies of data and results centrally as well as in each individual country; considering all ethical issues involved
- A public education system that will cooperate with assessing and recording pupils' competence and behavior.

Issues for consideration for long-term studies in low- and middle-income countries include:

- Measurements
- Terminology definitions
- Protocols
- Informatics support
- Capacity building

- Biostatistics
- Ethical aspects
- Equipment, laboratories, questionnaires, instruments, and biometrics
- Diseases, conditions, signs and symptoms
- Common and country specific
- Consent and community involvement.

A broad range of countries and organizations have been involved with these efforts.

Discussion. Dr. Pronczuk stressed that those involved with international cohort studies could not impose a protocol on developing countries. However, it is possible to develop a shared protocol from which each country can select elements of interest. This at least ensures compatibility of collected data. There are a number of WHO activities that give the basis for developing cohorts and enhancing chances of success. One of these is the capacity building component. Since 2000, WHO has been developing a series of training modules for health care providers. Forty-two modules are in use. A small pilot in Argentina is evaluating the use of a "green page"—a checklist to record environmental factors the child has been exposed to be included in the medical record of each child.

Wrap-Up and Next Steps in the Development and Implications of an ICCCC

Terence Dwyer, M.D., M.P.H., Murdoch Children's Research Institute

Dr. Dwyer outlined several next steps outlined for the group involved with the steering committee for the ICCCC.

1. Revise the steering committee to include representation from the planned cohorts, particularly the very large cohorts from Scandinavia and China. Dr. Dwyer suggested that Dr. Li represent the China cohort on the steering committee and a senior investigator from Scandinavia. Dr. Linet suggested that these initial assignments might be temporary. Rotation can be an important part of the process. The current steering committee includes representation from NICHD, NCI, and EPA. One of the activities the steering committee will be involved with is establishing working groups. The steering committee will convene via conference calls between in-person meetings for the consortium as a whole.
2. Develop protocols for cancer ascertainment and key exposures. These issues are unlikely to be decided in conference calls. Some may be topics for the next meeting.
3. Develop topics/hypotheses for ICCCC. Dr. Winn highlighted some of these in her presentation. This preliminary list will need to be refined.
4. Develop policies and procedures. Dr. Seminara and Dr. Winn provided a starting point for the policies and procedures to be expanded upon by the group.
5. Establish a portal and Web site. This can be accomplished fairly quickly.
6. Assemble data on relevant details of all studies. Preliminary work has been done to establish a database for recording details of the studies and exposure measures. This will help ascertain where additional work is needed.
7. Obtain information on ethics requirements of each cohort.
8. Set a date for the next meeting. This may be in conjunction with another major meeting.

Discussion. In a general discussion of the next steps, participants made the following points:

- Dr. Linet asked what the expected timeframe was for various aspects of the ICCCC. Dr. Dwyer indicated that it took approximately 1 year to launch the first meeting. Ideally, the second meeting will not take as long, perhaps 3–6 months, although this will largely depend on the effort put forth by the group. Protocol development will vary, but shorter-term projects can begin soon. A draft of the meeting report will be ready within approximately 2 weeks.
- Dr. Jørn Olsen suggested that a review process be implemented.
- Dr. Sjurdur Olsen noted that in the Danish and Norwegian collaboration one potential point of immediate collaboration was to look at folic acid in pregnancy. He suggested looking at critical points in time to look at exposure distributions of folic acid, breastfeeding, and other variables. It may be possible to compare the Danish/Norwegian data with the China cohort to draw comparisons and evaluate the power of combining these three studies. Some data suggest that women who take folic acid during pregnancy have an increased risk of developing breast cancer. This may be another area where the three cohorts can be combined. This could be something that could be done within 1 year. Dr. Winn noted that it would be possible to get childhood cancer incidence data from Cancer Incidence in Five Continents data collected and updated every 5 years by the International Agency for Research on Cancer.
- Susan Harlap, M.B., B.S., suggested that a relatively uncomplicated study would be to test the hypotheses that paternal age is a factor for both ALL and AML and that this effect is mediated by high birth weight.
- Dr. Jørn Olsen suggested that the French study start with pregnancy rather than birth.
- Dr Wiemels suggested a follow-up of all infants with chromosomal translocations relevant to childhood ALL and AML to determine their association with leukemia and other outcomes.

In conclusion, Dr. Dwyer noted that combined cohort data to identify critical stages would be a plausible short-term project. Another smaller collaboration might be to examine adolescents through adulthood. Researchers will continue to identify projects as the studies continue.

Participants

Jose Fernando Arena, M.D., Ph.D., NCI, NIH, DHHS
Kenneth Bart, M.D., M.P.H., Office of Public Health and Science, Office of the Secretary, DHHS
Arthur M. Bennett, B.E.E., M.E.A., NIH, DHHS
Ruth A. Brenner, M.D., M.P.H., NICHD, NIH, DHHS
Rebecca C. Brown, M.P.H., M.E.M., National Center for Environmental Assessment (NCEA), Office of Research and Development (ORD), EPA
Patricia Buffler, Ph.D., M.P.H., University of California, Berkeley
Richard Callan, M.P.H., National Center for Environmental Research (NCER), ORD, EPA
Stephen Chanock, M.D., NCI, NIH, DHHS
Terence Dwyer, M.D., M.P.H., Murdoch Children's Research Institute
Suzanne E. Fenton, Ph.D., National Health and Environmental Effects Research Laboratory (NHEERL), ORD, EPA
Nigel Fields, NCER, ORD, EPA
Alexa Fraser, Ph.D., Westat
Andrej Grjibovski, M.D., Ph.D., Norwegian Institute of Public Health

Susan Harlap, M.B., B.S., Columbia University
Daehee Kang, M.D., Ph.D., NCI, NIH, DHHS
Carol H. Kasten, M.D., NCI, NIH, DHHS
Andrea S. Kim, Ph.D., NCEA, ORD, EPA
Carole A. Kimmel, Ph.D., NICHD, NIH, DHHS
Mark Klebanoff, M.D., M.P.H., NICHD, NIH, DHHS
Sarah S. Knox, Ph.D., NICHD, NIH, DHHS
Danuta Krotoski, Ph.D., NICHD, NIH, DHHS
Joëlle Le Moal, M.D., Institut de Veille Sanitaire
Martha Linet, M.D., M.P.H., NCI, NIH, DHHS
Errol Mazursky, M.P.H.
Pauline Mendola, Ph.D., NHEERL, ORD, EPA
Les Mery, M.Sc., Public Health Agency of Canada
Jørn Olsen, M.D., Ph.D., Department of Epidemiology, University of California, Los Angeles
Sjurdur F. Olsen, Ph.D., Danish Epidemiology Science Center, Statens Serum Institut
Andrew Olshan, Ph.D., University of North Carolina, Chapel Hill
Marie Louise Østerdal, M.Sc., Statens Serum Institut
Montira Pongsiri, Ph.D., U.S. Department of Energy
Anne Louise Ponsonby, Ph.D., FAFPHM, FRACP, Murdoch Children's Research Institute
Monica Pourrat, M.D., Children's National Medical Center
R. Julian Preston, Ph.D., NHEERL, ORD, EPA
Jenny Pronczuk, M.D., World Health Organization
Nuria Ribas-Fito, M.D., Ph.D., Institut Municipal d'Investigació Mèdica
Cécile Ronckers, Ph.D., NCI, NIH, DHHS
Julie A. Ross, Ph.D., University of Minnesota
Rajni Samavedam, M.P.H., Booz Allen Hamilton, Inc.
Peter C. Scheidt, M.D., M.P.H., NICHD, NIH, DHHS
Sherry G. Selevan, Ph.D.
Daniela Seminara, Ph.D., NCI, NIH, DHHS
Nicola J.H. Symons, M.D., Ch.B., F.R.C.P., C.H., Bradford Hospital Trust
Ritu Tuteja, M.P.H., NCER, ORD, EPA
Sholom Wacholder, Ph.D., NCI, NIH, DHHS
Joseph Wiemels, Ph.D., University of California, San Francisco
Deborah Winn, Ph.D., NCI, NIH, DHHS
Li Zhu, M.D., M.P.H., Peking University Health Science Center