Third International Childhood Cancer Cohort Consortium Workshop
November 16–17, 2009
International Agency for Research on Cancer
Lyon, France

This workshop was sponsored by the U.S. National Institutes of Health (NIH) (including the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development [NICHD] and the National Cancer Institute [NCI], Bethesda, Maryland; the U.S. National Children's Study[1]; the Murdoch Childrens Research Institute (MCRI), Melbourne, Australia; and the International Agency for Research on Cancer (IARC) of the World Health Organization (WHO), Lyon, France.

Day 1

Welcome and Charge for the Meeting

Christopher P. Wild, Ph.D., Director, IARC, France
Martha S. Linet, M.D., M.P.H., NCI, NIH, USA

Dr. Wild welcomed the participants to the Third International Childhood Cancer Cohort Consortium (I4C) Workshop. Under his directorship, IARC has placed more emphasis on the relationship of early life exposures and the risk of childhood and adult cancers. From a scientific perspective, this relationship has become increasingly important because of the accumulating evidence that exposures very early in life influence cancer risk. The combination of mother–child cohorts, advances in the understanding of mechanisms of carcinogenesis, and technological advances now allow questions to be asked in population studies about the relationships among exposures, their biological effects on cancer pathways and disease, and risk of disease. Advances in knowledge of carcinogenesis can be translated into the clinical setting and can inform the population sciences. This two-way translation of basic science can contribute to the prevention and control of cancer. Although there are benefits of large-scale partnerships and cooperation, there are barriers to creating the interdisciplinary environment for such partnerships and cooperation. Continuous funding for prospective mother–child birth cohort studies can be challenging. The IARC welcomes partnership with the I4C if it has a role in supporting the development of the mother–child cohort studies. The IARC will propose partnership with the I4C when it submits its medium-term strategy to its governing body in May 2010.

Dr. Linet thanked Dr. Wild and the IARC for hosting the workshop. She welcomed returning colleagues as well as new colleagues. The I4C evolved from the recognition of the lack of progress in identifying consistent risk factors for childhood cancer through case-control studies, which can be confounded by the various forms of biases such as recall and maternal differences (case versus control). There is also recognition of the need to understand the mechanisms of carcinogenesis in children. The First I4C Workshop was held in 2005 to determine the interest in

[1] The U.S. National Children's Study is led by a consortium of federal partners: the U.S. Department of Health and Human Services (HHS) (including the NICHD and the National Institute of Environmental Health Sciences of the NIH and the Centers for Disease Control and Prevention), and the U.S. Environmental Protection Agency.

using the increasing number of child cohorts to study the etiology and mechanisms of carcinogenesis of pediatric cancer. Workshop attendees recognized that the largest individual cohort studies do not have adequate sample size to study childhood cancer on their own, but a consortium of cohort studies would. At the Second I4C Workshop in 2007, key hypotheses were identified for evaluation and discussion at the Third I4C Workshop. Workshop attendees discussed the potential of high-throughput studies of nonrandom chromosomal translocations that are seen in childhood leukemias to determine whether there are differences among populations in the occurrence of chromosomal translocations. Since the 2005 workshop, participating cohorts have provided data and data dictionaries in order to identify the variables that have been collected and to determine which variables overlap and which have the potential to be combined to test hypotheses. In addition, Jean Golding, Ph.D., Sc.D., has developed a comprehensive strategy for launching new child cohort studies as well as improving cohort studies already underway. The major focus of the Third I4C Workshop was to identify new ideas for future work.

I4C Cohorts: Summary of Data Collected

Terence Dwyer, M.D., M.P.H., MCRI, Australia

Previous I4C workshops focused on etiology (for example, chromosomal translocations) and exposures that may be of interest. This Third I4C Workshop has expanded this focus to include epigenetics, population genetics, and genome-wide association studies (GWAS). The scientific approach now focuses on the environmental exposure, questionnaire, and biospecimen data that have been collected and compiled. Dr. Dwyer reviewed the measures collected by the six "foundation" I4C cohorts and addressed the issues of the whether measures are comparable and whether they can be pooled. The six cohorts have been sharing data through the coordinating center (the Murdoch Childrens Research Center), which has been receiving, compiling, and attempting to pool the data.

The six foundation cohorts and their respective countries and number of participants are:

- Tasmanian Infant Health Study (TIHS)—Australia, 10,628
- Avon Longitudinal Study of Parents and Children (ALSPAC)—United Kingdom, 14,062
- Birth Defects Surveillance System for the Collaborative Project China (BDSS-China)—China, 245,336
- Danish National Birth Cohort (DNBC)—Denmark, 96,841
- Norwegian Mother and Child Study (MoBa)—Norway, 108,407
- National Children's Study—United States, 100,000.

None of the cohorts were set up to specifically study cancer, but their measures may allow for analyses of the etiology of childhood cancers. The cohorts have many common measures that are of interest to the etiology of childhood cancer. The foundation cohorts have collected the following cancer-relevant maternal measures (primarily questionnaire data): demographics, folic acid intake, smoking history, alcohol intake, parity, prenatal infections, pesticide exposures, aerosols exposures, paint exposures, heavy metal exposure, and radiation (x-ray and other) exposure. Child measures include gestational age, birth anthropometry, feeding methods, growth measures, infections (birth to 36 months), child care, and immunizations. Maternal biospecimens included saliva, hair and toenail clippings, buccal swabs, vaginal swabs, urine, and blood. Child

biospecimens include meconium, placenta, hair and toenail clippings, milk teeth, cord blood, and blood. Not all cohorts have collected all of the aforementioned biospecimens. Direct environmental measures—although not common across all six cohorts and not for all subjects within a cohort—include baby's room temperature and humidity; indoor and outdoor air pollutants (NO_X and CO); noise levels; magnetic radiation; maternal exposures in first and second trimester such as indoor air samples (heavy metals, total carbon, volatile organic compounds, carbonyls, NO_2), dust samples, tap water samples, and global positioning.

Three new I4C cohorts are the Japan National Children's Study (60,000 pregnancies), the China Children and Families Cohort Study (300,000 babies), and the Taiwan Birth Cohort Study (24,665 live births). These cohorts will have measures and collect biospecimens comparable to those of the foundation cohorts. Longitudinal maternal–child studies in other countries (for example, Israel, Brazil, Germany, Spain, Italy, Jamaica, France, and the United Kingdom) may eventually join or contribute to the I4C cohort.

I4C Data Management

Luke Stevens, MCRI, Australia

Mr. Stevens described the I4C data management workflow at the coordinating center (the MCRI). The workflow was developed for the TIHS and ALSPAC source data sets and data dictionaries. The final output of the data management process is pooled data for analysis.

As the I4C moves forward, an exchange of information via the Internet would be beneficial. Although a Web portal has not been developed, the following content ideas have been proposed:

- Publicly accessible information
  - General information: organization, research aims
  - Consortium members: foundation, current, and prospective
  - Links to publications
  - Acknowledgements: funding, collaboration efforts
- Consortium member access
  - Secure file exchange
  - Contacts: consortium members
  - Progress reports: current research activity
  - Future research: areas of prospective future analyses
  - Summary data: contributed data sets, pooled data
- Separate public and private areas
  - Public site
    - General information
    - Contacts
    - Links to publications
  - Private site
    - Consortium members
    - Progress reports
    - Research activity
    - Summary data.

Ascertainment of Childhood Cancer Cases and Initial Pooling Efforts: I4C Foundation Cohorts

Terence Dwyer, M.D., M.P.H., MCRI, Australia

The six cohorts involved in testing the folate hypothesis have been used to ascertain childhood cancer cases. Four of the cohorts (TIHS, ALSPAC, DNBC, and MoBa) use population-based cancer registries. Information in the registries is linked to ascertain cancer cases. This process has not been difficult with the TIHS and ALSPAC cohorts. The BDSS-China and U.S. National Children's Study cohorts will use hospital medical records, supplemental medical record data, and parent and/or caregiver reports to identify cancer cases.

For the TIHS, 31 total cancer cases and 4 leukemia cases were identified. For ALSPAC, 24 total cancer cases and 5 leukemia cases were identified. The observed numbers of leukemia cases in these two cohorts were lower than expected. By mid 2001 the numbers of expected leukemias for the DNBC, MoBa, and BDSS-China cohorts were, respectively, about 60, 44, and 139.

Exposure prevalence measures were compared across the TIHS, ALSPAC, BDSS-China, DNBC, and MoBa cohorts, and a number of common measures were identified. Maternal factors were:

- Demographics (age at time of baby's birth, completion of university degree, marital status)
- Prenatal folic acid supplementation (use, weeks of gestation when used, use by trimester)
- Prenatal alcohol consumption
- Prenatal smoking
- Prenatal insecticide exposure.

Child measures were gender, birth weight, body length at birth, head circumference at birth, and breast feeding during the first 6 months. Paternal demographics included age at time of birth and completion of university degree.

Dr. Dwyer described the initial pooling efforts. Cohort information from questionnaires, data dictionaries, and some publications are compiled. The information is abstracted to identify key domains and time periods of data collection. Questions within domains are compared to identify common elements. The data types and formats are then compared. Subsequently, data for case-cohort analyses are pooled.

The key parental I4C domains for examining the etiology of childhood cancer are as follows:

- Demographics (age at time of child's birth, parity, education, marital status, income)
- Occupation (antenatal occupation/type of work, postnatal occupation/type of work)
- Smoking and alcohol consumption (antenatal smoking/drug use [includes alcohol], postnatal smoking)
- Diet (antenatal supplement intake [includes folic acid], antenatal fish and diet intake, antenatal yogurt intake)
- Exposures (radiation during pregnancy, pesticide/chemical during pregnancy, passive smoking during pregnancy, sun exposure/vitamin D during pregnancy)
- Infections during pregnancy
- Biological specimens (blood, urine, hair, toenail clippings)
- Other (atopy/asthma).

The key child I4C domains for examining the etiology of childhood cancer are as follows:

- Anthropometry (gestational age, length, weight, head circumference, and so on)
- Diet (feeding methods [breast, bottle], fish and diet intake, yogurt intake)
- Exposures (passive smoking, pesticide/chemical exposure during first year of life, radiation exposure in first year of life, mixing with other people/siblings and child care)
- Infections during first year of life
- Biological specimens (blood, urine, and so on)
- Other (atopy/asthma).

By examining common elements and time points, pooled TIHS and ALSPAC data have been able to examine:

- Whether infants were exposed to household pesticides during the first few months of life
- Frequency of exposure to household pesticides by mother during infant's first few months of life
- Whether the mother smoked during pregnancy
- Trimester during which mother smoked
- Number of cigarettes smoked per day.

Other data examined for pooling include cancer outcomes; sex, birth weight, and ages; supplement intake (binary only); breastfeeding (at 4 weeks); maternal and paternal education; maternal substance use during pregnancy; passive smoking; and exposure to early childhood infections (child). So far, biospecimen data have not been able to be pooled.

Purposes, Pitfalls, and Practicalities of Undertaking a Birth Cohort Study and the Case for a Coordinating Center

Jean Golding, Ph.D., Sc.D., University of Bristol, UK

The ultimate aim of a birth cohort study is to identify factors influencing health and development so that preventive interventions can be put in place. In planning such a study, fundamental questions about the purpose of the study, the type and number of study participants, and study time frame must be addressed.

It is important to determine both environmental/exposure and outcome measures. The most statistically powerful environmental and outcome measures are continuous measures. Environmental measures need to include diet and psychosocial factors (both mother and child) as well as measures of pollutants. Because of problems in interpretation and comparison, outcome measures should include not only anthropometric measures and diagnoses but also signs and symptoms. It is important to that all measurements be accurate.

Biological samples are important in order to assess environmental exposures (especially to persistent chemicals), provide outcome measures, identify intermediate phenotypes, and provide a source of DNA. Genetic samples are important in order to help identify causal pathways, identify genetic effects that are contingent on the environment, and determine epigenetic effects.

Data sources include interviews/questionnaires to parents and study child, hands-on physical

examinations, hospital and other health records, case registers (for example, type 1 diabetes), education records, vital records (births and deaths), results of chemical/biochemical assays, genetic variants, and geographic information system (GIS) data concerning the neighborhood.

With regard to preparing, piloting, and validating, it is important to allow several years, involve scientists from many fields, include representatives from local areas and cultures, pilot all measures for acceptability and feasibility, and validate as many of the measures as possible.

Starting a birth cohort is a huge burden with regard to choosing the right protocols, gaining experience in piloting and validation, and reviewing literature. A coordinating center is important to the success of new birth cohort studies. A coordinating center should be able to advise on common protocols for data collection. Common protocols enable data to be combined to look at rarer outcomes and allow comparison of results. Common protocols allow data pooling and the comparison of results in different areas. If the results are similar in different areas, the findings are strengthened. Different results may indicate an interacting environmental or genetic effect. The Malaria Genomic Epidemiology Network (MalariaGEN) is an example of an effective coordinating center.

Collecting comparative information in a birth cohort study can be more efficient if there is a coordinating center. The staff of the coordinating center, including the director, should have experience in running cohort studies. The coordinating center should be appropriately funded. The coordinating center should advise on ethics; information technology; data organization; statistical analyses; collecting, storing, and processing biological samples; and building capacity. In addition, the coordinating center should provide training, resources, and literature reviews. Staff from the coordinating center should check on back-translations before going into the field; make site visits, especially during piloting and preparation; be available to advise and mentor research students; assist in writing grants; and provide a repository for the documentation, training manuals, and data.

There is a good scientific case for newly planned cohorts to work together with a coordinating center to enable the study to be designed in such a way that comparative studies can be mounted. Such a center can only be viable if funded appropriately. Staff of such a center should have experience with various aspects of running birth cohort studies. The aims of the coordinating center should include building capacity in the constituent centers and providing advice, training, and resources for researchers.

Reports on Progress of Studies Underway

Paternal Age and Childhood Leukemia Study

Jørn Olsen, MD, Ph.D., University of California, Los Angeles, USA

Male germ cells have their origin in embryonic life. They are dormant until puberty, at which time they produce an average of 100 million sperm per day. Because cell divisions continue in the male germ cells, there may be an accumulation of mutations with age, which may affect children. Female germ cells complete their cell divisions during fetal life. There are about 7 million germ cells at 6 months of gestational age, 2 million at the time of birth, 400,000 at

puberty, and 20,000 at the age of 30. About 400 eggs are needed to cover the time of reproduction. The expectation of accumulations of mutations with age in females is less than that for males.

Although there is evidence that the risk of miscarriage and some congenital malformations increase sharply with maternal age, evidence is limited for an effect of increasing male age. A number of studies have shown associations between paternal age and spontaneous abortions, some congenital malformations, preterm birth, low birth weight, perinatal and infant mortality, epilepsy, schizophrenia, and some childhood cancers. Recent studies have shown an association between paternal age and some childhood leukemias and central nervous system cancers. Most of these studies were case-control and registry-based studies, which are limited in their ability to control for potential confounders and biases.

Age effects may be due to biological aging, cumulative environmental exposures, or confounders and biases. Isolating the effect of paternal age from an effect of maternal age is difficult because the two ages are often closely correlated. Most information will come from parents with a large variation in age. However, parents with a large variation in age are not randomly selected from all parents.

Cohort studies can be used to examine confounding factors such as differences in lifestyle (and mutations), infections, and medicine use in couples with large variations in age. Understanding confounders may allow for greater statistical control in order to examine more accurately the effects of parental age on childhood cancer.

Using pooled data, the I4C will test the hypothesis that parental age is associated with childhood leukemia. The study aims to examine differences in lifestyle factors associated with differences in parental age in different cohorts, examine the association between parental age and childhood leukemia diagnoses before the age of 5, and adjust these associations for differences in lifestyle factors and origin of data. Investigators expect to see a moderate association between advanced parental age and leukemia. They expect part of this association to be driven by lifestyle factors, and they expect differences in age profiles of parents and the associated lifestyle factors to be highly culturally specific. Participating cohorts are TIHS, ALSPAC, BDSS-China, DNBC, MoBa, and the Campinas Infant Health Survey.

Prenatal Folate and Childhood Leukemia Study

Terence Dwyer, M.D., M.P.H., MCRI, Australia

There is good evidence that chromosomal translocations present at birth—probably occurring during fetal life—are important etiologically. Studies have shown that chromosomal translocations present at birth were present in about 40 percent of childhood leukemia cases. However, of the children who had the translocations at birth, only about 1 percent were later diagnosed with leukemia.

Folate and related genes are central to DNA metabolism. Changes in the synthesis of DNA have the potential to cause differences in the construction of DNA that might lead to translocations. Therefore, folate is worth investigating in relation to childhood cancer and, in particular,

childhood leukemia. Dietary folate is involved in metabolic pathways leading to chromosomal translocations frequently found in childhood leukemia. Folate alters the rate of synthesis of DNA and the level of DNA methylation. Polymorphisms of the methylenetetrahydrofolate reductase (MTHFR) gene have been associated with a reduced risk of acute lymphocytic leukemia (ALL) and acute myeloid leukemia. Epidemiological evidence from case-control studies suggests that folate supplementation during pregnancy could have a protective effect against the development of leukemia.

Pooled I4C data from the six foundation cohorts will eventually be used to test the hypothesis that prenatal maternal folic acid intake is associated with ALL development in children. Testing this hypothesis will require about 300,000 children. A pooled cohort analysis of TIHS and ALSPAC cancer registry data (n = 25,000) was conducted as proof of principle.

Data required for proof of principle are (1) identification of ALL cases in the two cohorts, (2) maternal prenatal folic acid supplementation intake, and (3) confounders and covariates, which include birth weight, parental age, socioeconomic status, maternal smoking and alcohol intake, passive smoking in early life, and parental exposure to herbicides, pesticides, and solvents. The questions from the maternal prenatal folic acid supplementation questionnaires were examined to determine the timing of supplement intake during pregnancy.

The I4C research plan is to use a case-cohort design drawing on pooled data from the six I4C foundation cohorts. However, the TIHS/ALSPAC proof-of-concept analysis used a case-control design.

Preliminary results from pooled cohort analysis (TIHS and ALSPAC) indicated an increased association with location. There was no association with male versus female births. There was a slightly higher risk among preterm births, which was not significant. There was an increased risk association with birth weight, but the increase was not significant. There was no association of risk with maternal and paternal age; confounders may play a role in this finding. Mothers having a degree is associated with a decreased risk. Folate was not protective. It is important to note that these results are based on only 56 cases.

Genetic Measures

Measuring Epigenetic Changes in Cohort Studies

Zdenko Herceg, Ph.D., IARC, France

The term *epigenetics* refers to changes in phenotype (appearance) or gene expression caused by mechanisms other than changes in the underlying DNA sequence. These changes may remain through cell divisions for the remainder of the cell's life and may last for multiple generations. Because there is no change in the DNA sequence, nongenetic factors cause differences in gene expression. Epigenetics plays an important role in the effects of early life exposure and susceptibility to childhood cancer and to diseases, including cancer, in adulthood. Epigenetic changes are profoundly altered in cancer and are likely to be "drivers" of cancer development and progression.

In embryonic development, cell fate regulation is based on selection between preexisting, intrinsically robust fates. The phenotypic state of a cell at any time is indicated by the position of the cell on the "epigenetic landscape." However, little is currently known about this epigenetic landscape and the environmental factors that influence this landscape and affect a cell's phenotypic state.

Epigenetic code extends and modulates DNA code. All cells in a given organism have the same genome. But it is the organization of the genetic information that allows different epigenomes. The epigenome defines the totality of epigenetic marks in a cell type. Cancer cells will have a distinct epigenome.

The three main epigenetic mechanisms are DNA methylation, histone modifications, and RNA-mediated gene silencing. These mechanisms are influenced by internal and external factors and involve an interface between the genome and the environment. Epigenetic mechanisms can buffer the influence of environmental factors but can also modulate responses to these factors. Because epigenetic changes are reversible, there is an enormous potential in terms of therapy and cancer prevention. There are several fundamental roles of epigenetic events in cancer:

- Mechanisms of cancer development and progression
- Biomarkers, diagnosis, and prognosis
- Cancer therapy
- Cancer risk assessment and prevention.

Recent advances have resulted in exciting opportunities for cancer epigenetics in understanding the causes of cancers and development of strategies for cancer prevention and control, including:

- New emerging concepts involving epigenetic mechanisms in critical cellular processes and disease
- Remarkable technological advances in epigenetics and epigenomics that allow powerful screening of large series of samples
- Availability of large case-control studies and population-based cohorts.

These advances allow testing of hypotheses on the impact of maternal prenatal diet and early life exposure on epigenetic states and susceptibility to childhood cancer and diseases in adulthood. These advances will also allow the discovery and validation of different biomarkers including epigenetics. Biological hypotheses and perspectives involve the following:

- Reconfiguration of DNA methylation in early embryos and windows of vulnerability
- Impact of maternal diet and early life exposure on epigenetic states and susceptibility to disease (timing of exposure: preconception, prenatal, or postnatal)
- Relationship between dietary supplementation, methyl-related intermediary metabolites in mothers, and epigenetic changes in offspring
- Identity and regulation of metastable epialleles (currently only a few genes with metastable epialleles have been identified)
- MicroRNA as biomarkers of early life exposure
- Epigenetic events and stem/progenitor cells
- Transgenerational epigenetic inheritance and determining which parts of the epigenome are not cleared between generations.

GWAS in Cohorts: Tobacco-Related Cancers

James McKay, Ph.D., IARC, France

Analysis of the Swedish, Utah, and Iceland Cancer registries revealed an elevated familial relative risk (*RR*) for common cancer types in first-degree relatives of cases:

- *RR* = 1.5–2.0 for breast, rectal, bladder, brain, cervix, and stomach cancers and non-Hodgkin lymphoma
- *RR* = 2.0–2.5 for prostate, colon, lung, ovary, pancreas, and kidney cancers; melanoma; and myeloma
- *RR* > 2.5 for lip, thyroid, testicular, and larynx cancer; lymphocytic leukemia; and Hodgkin's lymphoma.

Families and syndromes linked to susceptibility to tobacco-related cancers indicate genetics play a role in susceptibility (for example, the *P53* and *TERT* genes). Although genes are involved in lung cancer susceptibility, these genes have not been identified.

Investigators have been challenged in finding genetic elements that may be involved in susceptibility to different types of cancer. Ten years ago, they were examining rare variants and conducting linkage analyses, using families and family-based approaches to map out susceptibility genes and focus on actual causative genes. These approaches were not successful for a number of reasons. Today, investigators are using association-based studies, and in some contexts, this approach has been successful. Recently, IARC has been focusing more on GWAS.

The GWAS approach studies all common genetic variation in the genome in a number of cases and controls and looks at the frequency of alleles in cases and controls to determine those variations that are associated with a given outcome. A key advantage of the GWAS approach is that no prior hypothesis is needed. However, GWAS are prohibitively expensive to conduct on a large scale. To cut costs, GWAS are generally divided into "discovery" and "replication" phases. In the discovery phase, the entire genome is studied in a portion of cases and controls. Genotypes are determined for all cases and controls for a panel of more than 500,000 single nucleotide polymorphisms (SNPs). In the replication phase, the most "interesting" SNPs are identified and replicated in second independent sample (and third, and so on). This second phase can confirm the variants that are found to be associated with risk.

GWAS have a number of limitations or caveats. GWAS work best for detecting common genetic variations (that is, frequent SNPs). Rarer variations are more difficult to detect, primarily from the lack of statistical power. A common variant that is prevalent within a population is thought to be, in an evolutionary sense, relatively old, which implies that there is not much selection pressure on that variant. Overall, the GWAS effects are generally weak, which may require large samples to detect. GWAS rely on population heterogeneity (that is, the same variant being present in all study populations). A log additive inheritance model is most applicable.

Several IARC cancer GWAS have focused on lung, head, and neck cancers. One study confirmed a strong association between 15q25 (*CHRN*) and 5p15.33 (*htert*) gene variants and lung cancer. In the discovery phase of this study, the sample included four studies with 3,259 cases and 4,159 controls. In the replication phase, the sample included four studies with 2,899

cases and 5,573 controls. A GWAS of upper aerodigestive tract cancer genes included 2,091 cases and 8,334 controls from 2 studies in the discovery phase and 6,454 cases and 7,892 controls from 14 studies in the replication phase.

The IARC lung cancer GWAS revealed two loci of particular interest, one on chromosome 5p15 and one on chromosome 15q25. On 5p15, two independent variants are located near *TERT* and CPTM1L. On 15q25, the variants are located near nicotinic acetylcholine receptor genes (*CHRNA5* and *CHRNA3*). 15q25 has been confirmed as a lung cancer susceptibility locus, and the nicotinic acetylcholine receptor genes are strongly implicated. An addiction effect is present, but it is small and cannot explain the lung cancer effect. Clarifying the role of acetylcholine receptors in both addiction and in tumor development may suggest chemoprevention or treatment possibilities for addiction or lung cancer or both.

Use of Neonatal Blood Spots in Epidemiologic Studies

Carol H. Kasten, M.D., NICHD, NIH, USA

Neonatal blood spots are blood specimens collected from heel-pricks of newborns. The blood specimens—only a few drops—are placed on a card with absorbent filter paper. These cards are sometimes called Guthrie cards. Once the blood soaks through the paper, it is allowed to dry. The dried specimen and a collection form are placed in a special envelope and mailed within 24 hours to a newborn blood screening lab.

Neonatal blood spots are collected to screen for "inborn errors of metabolism." The general rule for disease screening is that the disease is treatable and that early diagnosis will prevent life-long disability. Examples of such diseases are congenital hypothyroidism, phenylketonuria, galactosemia, and sickle cell disease. Disease screening using DNA mutation panels have recently been added to blood tests. Cystic fibrosis is one of the diseases that be detected with this method. More than 30 diseases may be screened depending on regional resources and social viewpoint.

Neonatal blood spots are collected in all developed countries as well as many developing countries. The U.S. Centers for Disease Control and Prevention (CDC) has a laboratory quality assurance plan in collaboration with about 60 countries. CDC neonatal blood spot analytes include albumin, C-reactive protein, cortisol, thyroid hormones, and progesterone. Albumin is a measure of nutritional status, which could be used in conjunction with other variables to study maternal nutritional status and some outcomes. C-reactive protein is elevated in inflammation and infection and has the potential to be used as a primary screen for the occurrence of an infection and an outcome such as cancer. Maternal hyperthyroidism and neonatal hypothyroidism may play roles in premature births. Progesterone is one of many reproductive hormones that have been studied for their role in pathogenesis of breast cancer.

Neonatal blood spots have the potential to answer epidemiologic questions. Global neonatal blood spot repositories—particularly population-based collections—would provide an extraordinary resource for epidemiologic studies. Extracted, nonamplified DNA can be used for most studies of childhood disorders (for example, autism, developmental delay, and schizophrenia) as well as cancer studies. Neonatal blood spot DNA for GWAS can now be used

on high-density microarrays. Newer DNA extraction techniques from neonatal blood spots clearly demonstrate DNA yield is more than adequate for current purposes.

Neonatal blood spots can be analyzed for biomarkers of environmental exposures including metals (for example, lead, mercury, and cadmium) and maternal smoke. However, their use in analyzing environmental exposures requires more research and assay validation.

Use of neonatal blood spots for research purposes, while extremely valuable, poses numerous ethical issues:

- Parental consent for use in research
- Procedures if a serious genetic disease is identified (that is, whether permission is needed to recontact or whether parents should be contacted at all)
- Procedures if the disease is treatable versus diseases that are not treatable
- Confidentiality and data protection
- Whether research should wait until the subject can assent.

The I4C could lead the way in addressing ethical issues for epidemiologic studies of childhood cancer, and a conference on ethics of neonatal blood spot use in cancer research may be appropriate.

Environmental Exposures and DNA Methylation Patterns in Neonatal Blood and Leukemia Cells

Joseph L. Wiemels, Ph.D., University of California, San Francisco, USA

Epigenetics involves the heritable changes in gene function that cannot be explained by changes in DNA sequence. Changes such as DNA methylation are heritable. DNA methylation patterns can be passed on to successive cells, but the cells' genetic sequence is not altered. Although the pattern is relatively stable throughout the life history of a cell, cells generally become progressively more methylated as they age. Epigenetic changes cannot easily be altered but are potentially reversible.

DNA methylation is an exciting new area of biology because of its relationship to nongenotoxic environmental effects, developmental anomalies, and stem cell biology. In the case of cancer, there are possibilities for therapeutic intervention by changing DNA methylation in cancer cells. The DNA methylation component is readily adaptable for epidemiologic and clinical studies.

DNA methylation is the covalent addition of a methyl group to the fifth position of cytosine. Methylation does not change the coding function of cytosine. It is largely confined to CpG dinucleotides (that is, cytosines next to guanines). DNA methylation is catalyzed by DNA methyl transferase, and it is accessible for measurement. Most cytosines at CpGs are methylated, except for gene promoter regions. DNA methylation in promoter regions is associated with gene silencing.

Although there are "waves" of methylation and demethylation, DNA methylation is metastable during development. DNA methylation is a part of normal cell life span. It is important for gene imprinting and for specialization of cells. At important times in life, there is the potential for

aberrations in DNA methylation that increase future risk of disease. Genetic and epigenetic aberrations can occur in early development and can contribute to cancer. Understanding the aberrations may lead to prediction of disease and risk stratification of populations.

Environmental exposures at early developmental stages can play a role in aberrations of DNA methylation. For example, paternal prepregnancy smoking and not maternal smoking (before or after birth) has been shown to be associated with DNA methylation patterns in leukemia cells, and a history of parvovirus B19 has been shown to be associated with DNA methylation patterns in leukemia. Current studies of methylation in neonatal blood spots are looking at the effects of maternal folate nutrition, parental smoking, and pesticide exposure.

An intensive environmental assessment is crucial to understanding the role that environmental exposure plays in aberrations of DNA methylation. Cohort studies such as the Northern California Childhood Leukemia Study have an advantage for such assessments by conducting more precise environmental measures, tracking infections and chemical exposures in the child prospectively and serologically, tracking translocations and DNA methylation events prospectively, and using prior evidence from case-control studies to build hypotheses.

Genetic Measures: Panel and Audience Discussion

Panel members: Paul Ekert, Ph.D., M.B.B.S., FRACP, MCRI, Australia; Carol H. Kasten, M.D., NICHD, NIH, USA; Sharon Savage, M.D., NCI, NIH, USA; Camilla Stoltenberg, M.D., Ph.D., Norwegian Institute of Public Health, Norway

The genetics panel members and other meeting participants discussed the following issues:

- Strategies for genetic analysis, both short-term and long-term
- Identification of intermediate outcomes, endpoints, biomarkers, and measurements
- Development and prioritization of hypotheses
- Biospecimen types and times of collection
- Rapidly developing technologies and implications for long-term strategy
- Controls for childhood cancer cases
- Ethnicity as a confounder in cross-cohort comparisons
- Ethical issues of identifying phenotypes
- Selection and response bias in genetic studies
- Family-based studies as the foundation for studying cancer genetic risk factors
- Identification of alleles and SNPs associated with childhood cancers
- Identification of cancer precursors such as chromosomal translocations and DNA methylation
- Determinants of suspected risk factors such as high birth weight
- Data harmonization, pooling, and analysis.

Summary of Day 1

Martha S. Linet, M.D., M.P.H., NCI, NIH, USA
Christopher P. Wild, Ph.D., IARC, France

Drs. Linet and Wild summarized the Day 1 presentations and noted the following issues:

- Distinction between the purpose of the cohorts to look at childhood cancer and the additional

complementary opportunities to understand early life exposures in relation to cancer in adulthood
- Impact of very early life exposures on immediate changes in biology of the child
- Impact of early life exposure on the predisposition to cancer later in life
- Impact of environmental exposures on patterns of epigenetic changes and gene expression
- Transfer of knowledge and experience and stimulus for new studies in low- and middle-income countries
- Support of activities in other cohorts
- Coordinating center's role in
  - Pooling and analysis of data
  - Developing best practices and providing methodologic support and technical advice to new investigators and new cohorts
  - Training and building capacity
- IARC's role in facilitating partnerships/pairings between existing cohorts and new cohorts
- I4C activities that are different but complementary to case-control studies
- Notifying other cohorts and investigators of special scientific and funding opportunities.

Day 2

Environmental Measures

Epidemiological Approaches in the Investigation of the Environmental Causes of Childhood Cancer

Manolis Kogevinas, M.D., Ph.D., Center for Research in Environmental Epidemiology, Spain, and National School of Public Health, Greece

Some of the major problems in environmental cancer research in children are low exposures, mixed exposures (for example, air pollution and pesticides), and difficulties in obtaining information such as physical measurements and biological samples. These problems are in addition to all of the other complications of cancer research. Therefore, it is important that case-control studies of environmental exposures be adequately powered to prevent nondifferential misclassification of the effects of specific exposures on outcomes.

Major advances in environmental cancer research include the use of large studies and consortia, improved exposure assessment and better tools, new biomedical technology, and better understanding of mechanisms and development of disease, which allow testing of new hypotheses. In addition, new biomedical technologies are being used in epidemiological studies.

There are several advantages for exposure assessments in the I4C. Compared with case-control studies, cohort studies are more accurate and unbiased in measuring exposures than are other designs. However, accuracy and lack of bias are conditioned on the availability of similar and detailed exposure information in large cohorts. Some environmental exposures can be evaluated later using existing environmental databases and biomarkers.

The I4C is interested in many exposures associated with childhood cancers, such as:
- Air pollution (NO_2, particulate matter)

- Water contaminants (disinfection by-products, arsenic)
- Food contaminants (dioxins, acrylamide, polycyclic aromatic hydrocarbons)
- Pesticides
- Persistent organic pollutants (polychlorinated biphenyls, DDT) and novel contaminants (benzo(*a*)pyrene diolepoxides, perfluorooctane sulfonate)
- Radiation (extremely low frequency, radiofrequency, medical)
- Hormone disrupters (bisphenol A, dioxins, phthalates)
- Metals (arsenic, lead, cadmium, mercury)
- Solvents (benzene)
- Other chemicals (volatile organic compounds, paints)
- Lifestyle related exposures (second-hand smoke)
- Parental occupational exposures.

Several recent methodologic developments and new analytical tools have improved exposure assessments: GIS-based modeling methods, statistical modeling, improved deterministic models, and remote sensing.

Misunderstanding the exposure mechanism can lead to extreme exposure misclassification. Basic knowledge on mechanisms can change the view of how exposure assessments are conducted and how results are interpreted. Examples include low-level dietary dioxin exposure, evaluation of routes of exposure of drinking water disinfection by-products, and evaluation of global exposure of hormone disrupters, rather than evaluation of components.

Use of Biomarkers and "Omics" Technologies in Relation to Population-Based Studies

Christopher P. Wild, Ph.D., IARC, France

Environmental exposure assessments are important for several reasons. Most major common diseases have an environmental etiology. Currently exposure measurement is problematic in many areas, leading to misclassification. The success of large prospective cohort studies (for example, the UK Biobank) is predicated on the availability of accurate exposure assessment. This kind of investment suggests that there needs to be parallel improvement in exposure assessment. Exposure biomarkers can contribute to several areas in addition to elucidating disease etiology.

The "exposome" is defined as all environmental exposures from conception onward (including exposures from diet, lifestyle, occupation, and endogenous sources)—a quantity of critical interest to disease etiology. There are challenges to characterizing the exposome: It is large scale, complex, and dynamic. Unlike the genome, the exposome changes over time, and there may be critical windows of exposure (for example, in early life). However, even partial characterization can bring major benefits to epidemiologic studies.

Recent advances in exposure assessment include biomarkers, GIS, personal and environmental monitoring, and sophisticated questionnaires (for example, for dietary assessments). Exposure biomarkers in population studies can help define etiology through improved exposure assessment and reduced misclassification, by identifying susceptible individuals or subgroups, and by

contributing to biological plausibility of an exposure–disease association. Exposure biomarkers can be used for evaluating interventions as well as hazard and risk assessment.

There are several reasons for increased interest in studies of early life exposure and cancer risk:
- Observational studies linking early life exposures to disease later in life
- Fetal programming, adaptive response, and indications of alterations in the epigenome
- Vulnerability of children to environmental exposures
- Reported rise in childhood cancer rates.

Opportunities for associating early life exposure and cancer risk include the availability of mother–child birth cohorts and studies of mechanism-based biomarkers that relate exposure to disease. An example of a mechanism-based biomarker is aflatoxin–albumin adducts in blood as a measure of exposure to dietary intake of aflatoxin. Validation is an important step for biomarker studies. Application to population studies should be considered during the development of methods such as biomarkers.

First-generation exposure biomarkers tended to focus on a classical mutagen–carcinogen model of carcinogenesis (for example, metabolites, adducts, chromosomal alterations, somatic mutations). New biomarkers in relation to other mechanisms of carcinogenesis are focusing on epigenetic changes (for example, promoter methylation, histone acetylation, microRNA) and altered gene expression. Epigenetic biomarkers can be applied to population studies through quantitative analysis of DNA methylation and detection of stable microRNAs in plasma and serum. These biomarkers have a potential application to "exposures" such as nutrition, obesity, and physical activity.

There are a number of questions about whether "omics" help improve exposure assessment:
- Do specific exposures, or categories of exposure, alter the expression of specific groups of genes, proteins, or metabolites ("exposure fingerprint")?
- How do such alterations relate to dose?
- How stable are the alterations over time?
- How do potential confounding factors affect the association between exposure and "omics" biomarkers?

Although studies have begun to answer these questions, there are problems in comparisons of "omics" data in poorly designed studies, including unmeasured confounding by lack of information on age, sex, and other exposures; bias through differences in sample processing; selection bias through sampling procedures; and high costs leading to one-off or small-scale studies.

Challenges to use of biomarkers and "omics" technologies in relation to population-based studies include:
- Strengthening collaboration across mother–child birth cohorts internationally, with associated biobanks
- Investing in exposure biomarkers to complement genetic analysis for prospective cohort studies to fulfill their promise
- Applying new methodologies (for example, metabonomics) and knowledge of mechanisms

(for example, epigenetics) to population-based investigations of environment, lifestyle, and cancer
- Prioritizing studies of biological plausibility in establishing etiology, particularly in cases of modest risk elevation such as diet and cancer.

Environmental Data To Be Collected in the U.S. National Children's Study

Michael Dellarco, Dr.P.H., NICHD, NIH, USA

The National Children's Study (the Study) will examine the effects of environmental influences on the health and development of 100,000 children across the United States, following them from before birth until 21 years of age. The Study defines "environment" broadly, taking a number of natural and man-made environmental, biological, genetic, and psychosocial factors into account. The Study will serve as a platform for high-quality data acquisition for integration with additional studies. The Study will be of unprecedented scope and complexity, with two related phases: a Vanguard Study to guide the content and operations and a Main Study. The Vanguard Study, which began in January 2009, will enroll 1,000–2,000 children. The Main Study will begin when sufficient data are collected and analyzed from the Vanguard Study.

The Study's priority exposure areas are physical environment, chemical exposures, biologic environment, genetics, and psychosocial milieu. The priority health outcome areas are pregnancy outcomes, neurodevelopment and behavior, injury, asthma, and obesity and physical development.

Study visit measurement methods include questionnaires, environmental samples, physical examination, and biological specimens. There are several considerations for Study visit measurements: numbers of samples and specimens; costs for sample collection, storage, and analyses; burden on Study participants; and quantity and stability of environmental samples and biological specimens for future analyses. Most samples will be stored until needed for analysis. The three requirements for Study visit measurements are acceptability, feasibility, and cost. Each Study visit measurement method will undergo an evidence-based assessment for value, efficiency, and economy.

Environmental measurements will include:
- Indoor air (residence, child care locations)—particulate matter, NO_2, O_3, volatile organic compounds, aldehydes, ketones)
- Outdoor air (community-level)—particulate matter, NO_2, NO_X, SO_2, O_3
- House dust—allergens, endotoxin, mold, metals, pesticides (plus archive for future analyses)
- Potable water—disinfection by-products, metals, coliforms, nitrate, pesticides
- Soil and food—metals, pesticides.

Biomonitoring for chemical contaminants will include:
- Blood—polychlorinated biphenyls, persistent and nonpersistent pesticides, benzo(*a*)pyrene diolepoxides, perfluorinated compounds, perchlorate; lead, mercury, cadmium; bisphenol A
- Urine—perfluorobutane sulfonate, alkyl phenols, mercury (inorganic), arsenic (speciated), perchlorate, halogenated phenols (phencyclidine), phthalates, atrazine, organophosphates, carbamates, pyrethroids, ethylenebis(dithiocarbamate)/ethylene thiourea, cadmium

- Breast milk—dioxins/furans; organochlorine pesticides; polychlorinated biphenyls
- Meconium—cotinine, organophosphate metabolites
- Nails—mercury (organic, inorganic)
- Hair—cadmium, cotinine, mercury, nicotine.

Other exposure data will be collected using observations/visual assessments, diaries, and questionnaires. Questionnaire topic areas include diet, housing characteristics, occupation/hobbies, consumer product use, and pesticide use.

The Study protocol defines 11 biological sampling points from preconception through 36 months of age. Samples will be collected from mothers, children, and fathers. There will be 12 primary sample types: blood, urine, hair, nails, buccal cells, saliva, breast milk, vaginal fluid, meconium, cord blood, umbilical cord, and placental tissue. There will be 42 different sampling events in this time frame.

The Study protocol defines 11 environmental sampling points from preconception through 36 months of age. Four environmental sources (air, house dust, drinking water, soil) comprise sample types such as particulate matter, metals, inorganic and volatile organic pollutants, allergens, mold, and disinfection by-products. There will be 56 different sampling events in this time frame.

Study visit assessments must be feasible, acceptable, reliable, and reproducible. The visit assessments must be of value, in terms of both information gained and cost to deploy. The assessments must not be redundant. Finally, the assessments must be informative. They should be able to address a question that has potentially important public health impact, requires a study of the Study's size and robustness to answer, and is unlikely to be answered in another context.

Environmental Data Available from Established I4C Cohorts

Terence Dwyer, M.D., M.P.H., MCRI, Australia

The development of more accurate and reliable environmental measures will allow better estimations of exposure and risk factors in epidemiologic studies. The I4C will continue to investigate how some of the most advanced methods for measuring environmental factors can be used to study the cohorts. Some measures can be adopted into studies already under way, even in situations in which the children are now older than when exposures should ideally be measured. For example, biomarkers can be analyzed in biospecimens that were collected years ago, and GPS-type measures and historical data can provide information about a child's area of residence early in life.

To fully realize the great strength of the I4C, new studies have to use measures comparable to those of the established cohorts. Comparable measures allow pooling of data and analyses across the cohorts. The I4C data coordinating center is in the process of developing core measures in order to pool existing data from the foundation cohorts. Core variables will emerge, which will in turn guide other investigators as they start new studies. Current measures of environmental exposures have been gathered primarily from questionnaires. Potential measures include GIS information, videotaping of behaviors, and personal/individual environmental monitoring.

The foundation cohorts have collected questionnaire data on the following maternal environmental exposures: smoking history, alcohol intake, pesticide exposures, aerosols exposures, paint exposures, heavy metal exposure, and radiation (x-ray or other) exposure. Direct environmental measures—although not common across all six cohorts and not for all subjects within a cohort—include temperature and humidity of infants' rooms, indoor and outdoor air pollutants (NO_X and CO), noise levels, magnetic radiation, indoor air samples (heavy metals, total carbon, volatile organic compounds, carbonyls, NO_2), dust samples, tap water samples, and GIS data.

The data coordinating center will continue to develop more environmental variables. The best core measures will be included in all new studies. If there is the capacity in terms of resources and subjects' availability and willingness to participate, the data coordinating center will add other measures (for example, sun exposure and vitamin D).

The *IARC Monographs*: Some Recent Evaluations of Childhood Cancer

Kurt Straif, IARC, Lyon, France

The *IARC Monographs* are considered the "encyclopedia of carcinogens." They identify environmental factors that can increase the risk of human cancer. These factors include chemicals, complex mixtures, occupational exposures, physical and biological agents, and lifestyle factors. Since 1971, more than 900 agents have been evaluated. Of these, 102 are carcinogenic to humans, 68 are probably carcinogenic to humans, and 245 are possibly carcinogenic to humans. National and international health agencies use the *IARC Monographs* as a source of scientific information on potential carcinogens and as scientific support for their actions to prevent exposure.

Volume 100 of the *IARC Monographs* was recently completed. The goal of the volume was to update the critical review for each Group-1 carcinogen and identify tumor sites and plausible mechanisms. With a recent amendment to the preamble, which came into effect in 2006, the *IARC Monographs* became more specific in terms of identifying susceptible populations and developmental stages. This volume will provide data for two subsequent scientific publications: *Tumour Site Concordance between Humans and Animals* and *Mechanisms Involved in Human Carcinogenesis*.

Volume 100E addresses increased risk of cancer in children born of parents who smoke preconception, during pregnancy, and/or perinatally. Four recent studies showed significantly increased risk of hepatoblastoma associated with parental smoking and sufficient evidence of carcinogenicity. A meta-analysis of 11 studies of childhood leukemia (in particular, ALL) showed increased relative risk due to parental smoking but limited evidence of carcinogenicity. Based on mechanistic data, the following findings are reported:

- Newborns of smoking mothers have increased frequencies of *HPRT* mutations, chromosomal translocations, and DNA strand breaks.
- Sperm of smokers has increased frequencies of aneuploidy, DNA adducts, strand breaks, and oxidative damage.
- Cigarette smoke causes germ-cell mutations in mice.

Volume 100D addresses ionizing radiation (radioiodines and X and gamma radiation). Two findings are reported:

- There is sufficient evidence in humans for the carcinogenicity of exposure during childhood and adolescence to short-lived radioisotopes of iodine, including iodine-131. Short-lived radioisotopes of iodine, including iodine-131, cause thyroid cancer. The evidence comes from studies of the Chernobyl accident.
- There is sufficient evidence in humans for the carcinogenicity of *in utero* X and gamma radiation for childhood cancers.

Volumes 98 and 100F address occupational exposure of painters—specifically, maternal exposure to paints and childhood leukemia. Of nine case-control studies, five studies showed significant positive associations for maternal paint exposure either before or during pregnancy adjusted for age and/or sex, race, social class, or other variables; three studies showed borderline significant positive associations and nonsignificantly elevated odds ratios. In two studies, there were significant exposure–response relationships, according to duration of maternal paint exposure. Evaluation of studies provided limited evidence of an association between maternal exposure to painting—before and during pregnancy—and an increased risk of childhood leukemia in the offspring.

Volume 80 (2002) addresses nonionizing radiation—specifically, residential extremely low-frequency (ELF) electric and magnetic fields and childhood leukemia. Pooled analysis of 24 studies revealed limited evidence in humans for ELF magnetic fields and childhood leukemia and inadequate evidence for electric fields and childhood leukemia.

Nutrition and Environmental Exposure

Isabelle Romieu, M.D., M.P.H., Sc.D., Instituto Nacional de Salud Public, Mexico

Studies have shown that many pollutants (for example, ozone) induce signaling pathways of oxidative stress and inflammatory responses. Micronutrients (for example, omega-3 fatty acid and antioxidant vitamins) can play a role in neutralizing oxidants and modulating inflammatory responses. Because air pollutants are strong oxidants that can lead to an oxidative stress response, antioxidant defenses are important to neutralize the effect. Because vitamin C in lung lining fluid and vitamin E in cell membranes help protect against oxidative stress, supplementation by antioxidant could modulate the impact of air pollution. A randomized controlled trial of children with mild-to-severe asthma showed that supplementation with vitamins C and E reduced changes in pulmonary function and cytokine levels in nasal lavage in response to ozone exposure. The results indicate that vitamin supplementation was protective against the effects of air pollution. In addition, antioxidant supplementation can modulate the genetic susceptibility to air pollution effects, and increased dietary intake of fruits and vegetables can modulate lung inflammatory responses to air pollution.

POSGRAD (Pregnancy Omega-3 Supplementation and Growth and Development) is an ongoing randomized controlled trial of 1,000 pregnant women randomly assigned to receive 400 mg docosahexaenoic acid (an omega-3 fatty acid) or placebo from 18 to 22 weeks gestation until delivery. The trial will evaluate several intermediate outcomes in children from birth to 5 years

of age, including growth, mental development, immunologic response, atopy and asthma, and lung growth. POSGRAD will also evaluate allergen exposure, exposure to pesticides and insecticides, exposure to traffic, and parental smoking.

TROHOC Studies

Leslie Thomas Stayner, Ph.D., University of Illinois at Chicago

Case-control studies are basically not different from cohort studies. The control group in a case-control study is essentially a subsample of the source population in a cohort study. Therefore, a well designed case-control study can be as good as a cohort study. There are pros and cons for both types of studies. Case-control studies are efficient for rare diseases (pro) but are subject to recall bias and selection biases (cons). In cohort studies, recall bias is not an issue (pro), but there may be participation bias as well as greater time and expense to implement (cons). One advantage of cohort studies is the collection of biospecimens and exposure information during the prenatal and perinatal periods. Depending on the hypotheses, it may be possible to pool data from case-control studies with data from cohort studies. Analyses in cohort studies are often nested case-control studies.

Reports from Other Cohorts

Israel

The Jerusalem Perinatal Study began in 1963 with the observation of ethnic differences in the rates of preeclampsia among the women in Jerusalem. From 1964 to 1976, the study recorded information on 92,408 live births and stillbirths, about 42,000 mothers, and about 40,000 fathers. Data were collected on maternal conditions, birth and pregnancy complications, and birth outcomes. The study generated about 60 publications during the study's initial phases. Subsequently, investigators used the cohort data to assess outcomes among military recruits at 17 years of age. Correlational studies included birth weight and weight at 17 years. In 1999, cohort data were linked to Israel's population death and cancer registries. The cancer registry was set up in 1961, was enforced by law in 1981, and is about 95 percent complete. The data are fairly complete from 1961 to 1981. Investigators have complete ascertainment of childhood cancers from the cohort. More recent studies of the cohort include associations of birth weight and congenital malformations with cancer, paternal age and breast cancer, familial cancers, and hospitalization for infection as a risk factor mainly for non-Hodgkin's lymphoma. The Jerusalem Perinatal Study can be used for multigeneration studies and familial cancer studies. There were about 125 cancer cases in the cohort children, about 1,000 cases among their offspring, and about 10,000 cases among the cohort parents. Other areas of interest are congenital malformations and parental cancer, and high birth weight as a risk factor for mothers' leukemia.

Brazil

The Brazilian cohort study is being conducted in a city (Campinas) with a population of 1 million. The study began with a single patient clinic and no funding. Today, it has the biggest specialized pediatric oncology center from Mexico to Argentina. The building has 20,000 square meters of space, with several laboratories, radioimaging technology, and radiation technology.

There are various cancer programs within the cohort study. One focus is multidisciplinary collaboration of childhood cancer prevention. Research areas include epidemiologic genetics, childhood colon cancer, teratogenesis and birth defects, basic carcinogenesis, environmental exposures, genetic polymorphism, pregnancy and childhood cancer, cancers at different developmental stages from birth to 1 year of age, leukemia in preschool children (2–6 years of age), Hodgkin's lymphomas in Hispanics, breast cancer in women 18–30 years of age, cost–benefit issues for the use of high-cost technologies, and education on prevention and control.

Germany

The Survey of Neonates in Pomerania (SNiP) is a population-based birth cohort study. Between 2003 and 2008, SNiP assessed just about all mothers at the time of birth (98 percent coverage) in the region. The study has about 7,000 mothers and about 7,100 children. About 75 percent of the mothers participated in interviews. The study collected routine data during pregnancy, perinatal hospital data, clinical interview data on parental chronic diseases, and questionnaire demographic and social behavior data. Biospecimens include cord blood (plasma and DNA) and placenta. The children will begin attending school in 2010, and there is a mandatory school exam at 6 years of age. So far, data analyses have been limited due to lack of funding.

Spain

The Spanish Environment and Childhood Research Network (Infancia y Medio Ambiente [INMA]) is composed of seven small birth cohorts, with about 3,000 births. Detailed environmental data have been collected. The outcomes of interest include growth, neurodevelopment, behavior, and respiratory health. The network may be able to contribute a few cancer cases to the I4C. INMA is coordinating with other European birth cohorts, and eventually there will be a detailed inventory of environmental exposure data and some outcome data. It is hoped that the European birth cohorts will be expanded into an international registry of birth cohorts.

Italy

NINFEA (Nascita ed Infanzia: Gli Effetti dell'Ambienta) is a Web-based multipurpose cohort set up to investigate the effects of exposures during pregnancy and early life on infant, child, and adult health. The objective of the study is to recruit and follow up a birth cohort of at least 7,500 newborns in Italy. Ongoing recruitment began in July 2005. The pilot phase ended in November 2008. Collection of saliva samples from mothers and children began in September 2009. There are currently about 2,500 participants. The existence of the study is advertised using "active" (offline) and "passive" (online) methods. Members of the cohort are babies born to women who become aware of the existence of the study during pregnancy, have Internet access, and volunteer to participate. Women participate through registration on the study Web site and complete online questionnaires. Questionnaires are given during pregnancy, at 6 months of age, and at 18 months of age. Infant saliva samples are collected at 6 months of age. Short follow-up questionnaires will be given at 4 years, 6 years, and up to 18 years. Childhood cancer cases will be ascertained through cancer and clinical registries.

Jamaica

The Jamaican birth cohort study began in 1986. The study's objective was to identify modifiable factors for maternal and perinatal mortality and morbidity. More than 10,000 women and their babies born in September and October of 1986 were recruited. About 94 percent of the live births that occurred in that year were identified and enrolled. Funding limitations have restricted follow-up of these children to a subset of between 1,500 and 1,700 children who attended schools in the Kingston metropolitan area. The children were evaluated at 11 and 15 years of age. About 900 children were evaluated at 18 years of age. About 85 percent of children's follow-up data have been linked to their birth records. Few biological samples were collected. Data on maternal health and pregnancy and perinatal factors were gathered mostly from interviews. The University of the West Indies–Mona has a tertiary referral hospital that operates a cancer registry, which provides an opportunity to identify diagnosed cases among the cohort. National hospital mortality and morbidity registries use ICD-10 codes, which gives the potential to identify admissions of cancer cases and cancer deaths. A Web site is being set up to reestablished contact with the cohort members outside of the Kingston metropolitan area. Population-based cancer epidemiology is an underdeveloped area of research. The University of the West Indies–Mona welcomes the opportunity for capacity development as a resource center for the Caribbean region.

France

The French Birth Cohort Study (Etude Longitudinale Française depuis l'Enfance [ELFE]) was established to explore questions about the effects of education, socioeconomic status, health, and environmental factors on perinatal disease. The cohort will include 20,000 births occurring on 12–16 days chosen in 4 different quarters of the year. There are face-to-face questionnaires at birth, 2 months of age, and 3 years of age. Phone interviews will be conducted at 1 and 2 years of age. Medical examinations will be conducted at 3, 6, and 10 years of age. Exposures of interest include lead, mercury, pesticides, organotins, phthalates, radon, medical radiations, and ultraviolet radiation. Health outcomes of interest include asthma, allergies, respiratory health, neurocognitive development, and endocrine dysfunctions. Biological analyses will include maternal urine, milk, and hair; cord blood; and placenta. The cohort will prospectively collect data on a variety of exposures such as diet, pesticides, and infection. It is estimated that the cohort will have about 40 cases of childhood cancer over a 15-year period, with about 12 cases of leukemia.

China

From 2006 to 2007, the China Children and Families Cohort Study (CCFC) recruited women during preconception and prenatal periods. The study plans to enroll 300,000 children and will include parents and grandparents. Environmental exposures will include air pollution, water pollution, and pesticides.

United Kingdom

The Centre for Longitudinal Studies is conducting three ongoing cohort studies: the 1958

National Child Development Study, the 1970 British Cohort Study, and the Millennium Cohort Study (MCS). The MCS is a longitudinal survey. It is following the lives of a sample of about 18,500 babies born in the United Kingdom in 2000 and 2001. Interviews were conducted at 9 months of age and at 3, 5, and 7 years of age. The study's focus has been primarily socioeconomic status and cognitive development, but there are physical measurements and saliva samples at 3 years of age. The study has consent to access medical records, maternity, hospital episodes, and general practice records. Cancer cases can be ascertained through hospital episode records. Another potential resource for cancer epidemiology is the Southampton Women's Survey, which recruited 12,000 women of reproductive age and took baseline measurements. Three thousand of the women became pregnant. Data collected on these women included fetal growth measurements, maternal diet, cord blood, and placenta. The Southampton Women's Survey is involved with other birth cohort studies, including one in Puna, India.

Ireland

The Growing Up in Ireland National Longitudinal Study of Children will take place over 7 years and will follow the progress of 2 groups of children; 8,500 9-year-olds and 10,000 9-month-olds. During this time the study will carry out two rounds of research with each group. The study is now in its third year. The main aim of the study is to paint a full picture of children in Ireland and how they are developing in the current social, economic, and cultural environment. The study collects interview data from the child, the primary caregivers, and schools. The first wave of data collection on the 9-year-olds has been completed. The data will be posted to a public archive in January 2010. The posting will include a data dictionary. This information will be used to assist in policy formation and in the provision of services that will ensure that all children will have the best possible start in life. The study initially did not collect any physical measures but has recently received permission to ask participants for saliva samples from the 9-month-old cohort when the children are 3 years old. Preliminary data indicate that about 11 percent of the 9-year-olds have chronic illness. The data also indicate a greater proportion of psychological, behavioral, and psychiatric disturbances in the 9-year-old males. This group appears to be falling behind in educational performance, compared with females.

Nutritional Measures

Maternal Diet and Childhood Cancer: The Rationale Behind and Plans for Coordinated Analyses of DNBC and MoBa

Sjurdur F. Olsen, Ph.D., Statens Serum Institut, Denmark

The objective of the European Union Sixth Framework Programme consortium—"Earnest" 2005–2010, WP 2.5—is to appraise the scientific potential of undertaking analyses of the combined DNBC and MoBa databases and to explore the potential for future combined analyses.

Epidemiologic evidence about the impact of maternal diet on childhood cancer comes from numerous case-control studies. Implicated dietary factors include vegetables, iron, nitrosamides, phytoestrogens, vitamins, proteins, cured meats, red meats, and dietary supplements. So far, there have been no (large enough) prospective studies with concurrent assessments of dietary intake during pregnancy. There is a great need for longitudinal studies regarding health effects of

maternal dietary exposures. The results of early case-control studies motivated collection of dietary data in DNBC. The need for evidence from longitudinal studies provided a good rationale for coordinated analyses between DNBC and MoBa on these issues.

Reasons for undertaking coordinated analyses include greater power for studying rare phenomena, confirming findings obtained in one cohort, and widening the exposure ranges. Issues for the DNBC–MoBa coordinated analyses are whether there will be sufficient power and the timing of opening the databases. Based on power calculations, the data analyses may be undertaken at the following times:

- When recruitment in MoBo has been completed: 2008
- When the first children in DNBC are older than 14 years: 2010
- When all children in DNBC are older than 14 yeas: 2015
- When all children in MoBa are older than 14 years: 2023.

DNBC and MoBa used food frequency questionnaires (FFQs) to assess prenatal maternal diet. Both FFQs are extensive, with more than 250 items. Many items are identical or nearly identical. In general, both FFQs have similar emphases (for example, very detailed questions about vegetables and fish consumption). The two FFQs have somewhat different time windows, but if general dietary patterns are the interest, this may not be a problem. They have comparable time scales, similar general structure, and many identical or near-identical items.

Power calculations suggest that coordinated analyses between DNBC and MoBa would be meaningful. Compatibility between employed designs and methods is relatively high. Particularly, comparisons of structures and underlying principles of the two FFQs suggest that coordinated analyses would be meaningful for several purposes. In addition, there is balance between study sizes (N ~ 70,000 for DNBC and N ~ 86,000 for MoBa), and they are concurrent studies (with 3–5 years in between them).

The Statens Serum Institut received a grant from the Danish Cancer Society to identify and specify maternal diet–childhood hypotheses, explore distributions of cancer relevant exposures, develop databases as needed, and develop analytic strategies for specific hypotheses. However, this grant provides only partial funding for the project. The Nordic Cancer Union may provide additional funding.

Comparison of Methods on Dietary Assessment in Chinese Population

Hao Ling, Ph.D., Peking University Health Science Center, China

The most commonly used methods of measuring diet in large-scale studies in Chinese are 24-hour recall, household food inventory, and FFQs. FFQs have been used only in the past 10 years. Because of the differences between Chinese diets (that is, the high proportion of mixed dishes) and Western diets, there are concerns about ability of FFQs to accurately assess Chinese diets. The methods were compared in three studies.

- For the Fourth National Nutritional Survey in 2002, the FFQ was considered good for measuring the intake of frequently consumed food items.
- For the Shanghai Women's Health Study in 1997, an assessment indicated that the FFQ reliably and accurately measured usual intake of major nutrients and food groups among

women in the study.
- In the Folate Status Investigation Among Chinese Adults, an FFQ assessed 138 food items and food groups over a 3-month period. The focus was on folate intake and the frequency and average amounts each time. The results showed a good correlation between folate intake estimated by FFQs and folate biomarkers.

There are limitations to the three dietary assessment methods. To capture seasonal variations, the 24-hour dietary recall must be administered more frequently, which will increase cost. The household food inventory can collect good data for a family or group, but it is not good for collecting individual diet information. FFQs do not provide good estimates of absolute nutrient intakes. FFQs may not be reliable with children. Gathering data on food intake at schools and day care centers may be challenging and may require alternate approaches.

The CFCS I is a prospective follow-up study of all the live born children and their mothers who participated in the folic acid intervention during 1993–1996. A diet feasibility study with 3,000 adolescents 13–15 years of age and 3,000 mothers 35–45 years of age from the CFCS I will be implemented in 2010–2012. One of the aims of the adolescent portion of the study is to collect data on the association between diet and nutritional status (from blood measures) during adolescence and sexual maturation and anthropometric indices, as risk factors for adult cancers. One of the aims of the maternal portion of the study is to collect data on the association between methylation status (blood measures) and diet, and the effect modification with the history of folate supplementation during pregnancy.

Reports from Breakout Sessions

Two breakout sessions were held to discuss (1) environmental measures to be added or analyzed for the next meeting and (2) genetic data and issues to be analyzed for the next meeting.

Environmental Measures

The breakout group identified the following measures:
- Chemicals
  - Pesticides (herbicides and other biocidal agents)
  - Solvents (benzene, including paints)
  - Metals (lead, mercury, arsenic, cadmium)
  - Other chemicals (persistent organic pollutants, polychlorinated biphenyls, perfluorooctane sulfonate, hair dyes, cosmetics)
  - Food contaminant (aflatoxins)
  - Air-pollution (indoor and outdoor)
  - Allergens
  - Water contaminants
  - Endocrine disruption
- Radiations
  - Ionizing radiation (natural environmental , gamma, radon decay)
  - Medical ionizing radiation (prenatal and postnatal exposures)
  - Nonionizing (extremely low frequency, radiofrequency, ultraviolet)

  - Emerging exposures (nanoparticles)
- Nutrition, infections, and lifestyle
  - Nutrition, obesity, and physical activity
  - Infections
  - Smoking (active and second-hand smoke)
  - Hormones and drugs
  - Vaccination
  - Population mixing
  - Stress.

Because of their consistent linkage with leukemia and other childhood cancers, the following initial priority areas were identified:
- Pesticides
- Solvents
- Other chemicals such as polychlorinated biphenyls and perfluorooctane sulfonate
- Medical ionizing radiation
- Radiofrequency radiation
- Nutrition, obesity, and physical activity
- Infections
- Smoking
- Hormones and drugs.

The breakout group proposed the following goals and timeline:
- Form six working groups and identify chairs
- Define contact procedures (conference calls)
- Evaluate what the specific hypotheses should be; what data are available; what measures should used, and feasibility
- Product is a brief proposal
- Proposals before November 2010.

Genetic Data and Issues

The breakout group noted that it is important to continue building on the unique features of I4C. The I4C's greatest resources are the biospecimens, environmental samples, and family data that have been collected prior to disease development. At this time, the I4C needs to develop a short-term strategy focusing on studies and activities that can be conducted in the next 1–3 years.

The breakout group made several general suggestions:
- Focus on intermediate phenotypes
- Conduct studies of available cases of childhood leukemia by 2011 or 2012
- Establish control groups from the different cohorts that can be used for multiple purposes, both with cases from the cohort or intermediate phenotypes from the cohort, and perhaps with cases from other sources.

The breakout group made several specific suggestions:
- Propose DNA methylation studies in children

- Focus on DNA methylation as both an intermediate phenotype and in available cancer cases in the cohorts
- Determine which chromosomal regions to focus on
- Focus on translocations that are associated with subgroups of leukemia in children, as an intermediate phenotype and possibly as a study in available cases
- Explore which cohorts could contribute RNA samples
- Study translocations in samples such as neonatal bloodspots
- Propose studies of telomere length in children
- Focus on the advantage of having sequential samples from different developmental times
- Consider studies of copy number variations, using primary DNA or GWAS data
- Collect GWAS data from preterm births, spontaneous abortions, and other phenotypes
- Focus on effects of environmental exposures (for example, infections) on genes involved in the relevant metabolic pathways.

The breakout group proposed that working groups be established for high priority genetic measures and that cross-links between the environmental working groups and the genetics working groups be examined. It was recommended that all prospectively enrolling cohorts collect neonatal blood spots. The collection of RNA samples was suggested, as was a joint WHO/IARC meeting at the Brocher Foundation in Geneva.

Conclusion and Next Steps

The meeting participants concluded the following:
- I4C data can be pooled.
- Biospecimens and environmental samples and measures have been collected and are available.
- I4C cohorts will collaborate.
- Other cohorts—even smaller cohorts—that could pool data need to get involved and maintain contact with the I4C coordinating center.

Participants

Annibale Biggeri, M.D., M.S., University of Florence, Florence, Italy
Karen Birmingham, University of Bristol, Bristol, UK
Valentina Bollati, Ph.D., University of Milan, Milan, Italy
Silvia R. Brandalise, M.D., State University of Campinas, São Paulo, Brazil
Marie-Noel Bruné, WHO, Geneva, Switzerland
Jia Chen, Sc.D., Mount Sinai School of Medicine
Jacqueline Clavel, M.D., Ph.D., Institut National de la Santé et de la Recherche Médicale (INSERM), Villejuif, France
Maria F. Costa, Campinas Health Secretary–Brazil, Valinhos, Brazil
Yayun Dai, International Agency for Research on Cancer, Lyon, France
Michael J. Dellarco, Dr.P.H., NICHD, NIH, HHS
Wei Deng, Ph.D., Fudan University, Shanghai, China
Terence Dwyer, M.D., Ph.D., MCRI, Parkville, Australia
Paul Ekert, Ph.D., M.B.B.S., FRACP, MCRI, Parkville, Australia

Ruth A. Etzel, M.D., Ph.D., WHO, Geneva, Switzerland
Li Feng, M.D., Ph.D., Shihezi University, Xinjiang, China
Nigel A. Fields, M.S.P.H., National Center for Environmental Research, EPA
Jean Golding, Ph.D., D.Sc., F.Med.Sc., University of Bristol, Bristol, UK
Yang Gonghuan, M.D., Ph.D., China Center for Disease Control and Prevention, Beijing, China
Bernd Grosche, Federal Office for Radiation Protection, Oberschleissheim, Germany
Neela Guha, Ph.D., M.P.H., IARC, Lyon, France
Shuxia Guo, M.D., M.P.H., Shihezi University, Xinjiang, China
Hind Hafsi, IARC, Lyon, France
Zdenko Herceg, Ph.D., IARC, Lyon, France
Hector Hernandez, IARC, Lyon, France
Clarisse Héry, IARC, Lyon, France
Margaret A. Honein, Ph.D., M.P.H., National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention (CDC), HHS
Huang Jin, M.D., Ph.D., Shihezi University, Xinjiang, China
Carol H. Kasten, M.D., NICHD, NIH, HHS
Ausrele Kesminiene, IARC, Lyon, France
Emmanouil Kogevinas, M.D., Ph.D., Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
Fabienne Lesueur, Ph.D., IARC, Lyon, France
De-Kun Li, M.D., Ph.D., Kaiser Permanente
Zhu Li, M.D., M.P.H., Peking University Health Science Center, Beijing, China
Martha S. Linet, M.D., M.P.H., NCI, NIH, HHS
Hao Ling, Ph.D., Peking University Health Science Center, Beijing, China
Somdat Mahabir, Ph.D., M.P.H., NCI, NIH, HHS
Manuela Marron, IARC, Lyon, France
Milena M. Maule, University of Turin, Turin, Italy
Affette M. McCaw-Binns, Ph.D., M.P.H., University of the West Indies–Mona, Kingston, Jamaica
John W. McDonald, Ph.D., University of London, London, UK
James McKay, Ph.D., IARC, Lyon, France
Franco Merletti, University of Turin, Turin, Italy
Domenico F. Merlo, Ph.D., National Cancer Research Institute, Genoa, Italy
Elizabeth Milne, Ph.D., M.P.H., Telethon Institute for Child Health Research, West Perth, Australia
Cynthia A. Moore, M.D., Ph.D., National Center on Birth Defects and Developmental Disabilities, CDC, HHS
Thomas O'Dowd, M.D., Trinity College, Dublin, Ireland
Jørn Olsen, M.D., Ph.D., University of California, Los Angeles
Sjurdur F. Olsen, Ph.D., Statens Serum Institut, Copenhagen, Denmark
Ora Paltiel, M.D.C.M., M.Sc., Hadassah Medical Organization, Jerusalem, Israel
Rafael Peris-Bonet, M.D., Ph.D., University of Valencia, Valencia, Spain
Shay Pintov, M.D., Academic College of Management, Kefar Bilu, Israel
Paula Pisani, University of Turin, Turin, Italy
Thais J.S. Pontes, M.D., IARC, Lyon, France
Xu Qian, Ph.D., Fudan University, Shanghai, China

Jiang Qingwu, M.D., Fudan University, Shanghai, China
Karina Ribeiro, D.D.S., Ph.D., Faculdade de Ciencias Medicas da Santa Casa de São Paulo, São Paulo, Brazil
Lorenzo Richiardi, University of Turin, Turin, Italy
Isabelle Romieu, M.D., M.P.H., Sc.D., Instituto Nacional de Salud Public, Cuernavaca, Mexico
Suna Sabuncuoglu, IARC, Lyon, France
Markku Sallmen, Ph.D., Finnish Institute of Occupational Health, Helsinki, Finland
Sharon Savage, M.D., NCI, NIH, HHS
Kenneth C. Schoendorf, M.D., M.P.H., NICHD, NIH, HHS
Chiara Scoccianti, IARC, Lyon, France
Remy Slama, Ph.D., M.Sc., INSERM, Grenoble, France
Leslie Thomas Stayner, Ph.D., University of Illinois at Chicago
Eva Steliarova-Foucher, Ph.D., M.Sc., IARC, Lyon, France
Luke Stevens, MCRI, Parkville, Australia
Camilla Stoltenberg, M.D., Ph.D., Norwegian Institute of Public Health, Oslo, Norway
Kurt Straif, IARC, Lyon, France
Gitanjali S. Taneja, Ph.D., NICHD, NIH, HHS
Jochen René Thyrian, Children's University Hospital of Greifswald, Greifswald, Germany
Gabriella Tikellis, MCRI, Parkville, Australia
Martine Vrijheid, CREAL, Barcelona, Spain
Joseph L. Wiemels, Ph.D., University of California, San Francisco
Christopher P. Wild, Ph.D., IARC, Lyon, France
Nicholas C. Wong, Ph.D., B.Sc. (Hons), MCRI, Parkville, Australia
Ming-Zhu Zhao, Ph.D., Pu Nan Hospital of Shanghai, Shanghai, China