This workshop was sponsored by the Centre for Research In Environmental Epidemiology, Barcelona, Spain; the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, HHS, USA; Murdoch Childrens Research Institute, Australia; the National Cancer Institute, National Institutes of Health, HHS, USA; and the National Children’s Study, USA.

### Day 1

**Welcome and Introduction to the Meeting**  
*Martine Vrijheid, M.D., M.P.H., Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain*

Dr. Vrijheid welcomed the participants to the Fourth International Childhood Cancer Cohort Consortium (I4C) Workshop. She gave a brief overview of CREAL. The mission of CREAL is to conduct high quality epidemiological research on environment and health and to provide scientific knowledge relevant for public health action. CREAL’s research is mainly focused on six program areas: respiratory, cancer, childhood, air pollution, water pollution, and radiation. CREAL is participating in the following four birth cohort studies:

- INMA—Infancia y Medio Ambiente, a Spanish birth cohort study on environmental pollutants and child health
- The Rhea Study—a mother-child cohort in Crete
- ENRIECO—Environmental Health Risks in European Birth Cohorts
- CHICOS—Developing a Child Cohort Research Strategy for Europe
- Born in Bradford—a birth cohort study in the United Kingdom.

**The I4C Initiative**  
*Terry Dwyer, M.D., M.P.H., MCRI, Australia*

**I4C Rationale.** The rationale for the I4C is to provide prospective data on potential causes of childhood cancer that are currently lacking and impeding progress in preventing cancer in children. Because childhood cancer is rare, a consortium is needed to obtain sufficient evidence on the relationship between childhood cancer and environmental exposures. For a cohort of 100,000 children followed from birth to 14 years of age, there are 242 cases for all types of cancer. Of the 242, there are 56 cases of acute lymphoblastic leukemia (ALL) and only 15 cases of acute myeloid leukemia (AML). In order to be sufficiently powered (80 percent with

---

1 The National Children’s Study is led by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development of the National Institutes of Health (NIH) in collaboration with a consortium of federal government partners. Study partners include the National Institute of Environmental Health Sciences of the NIH, the Centers for Disease Control and Prevention, and the Environmental Protection Agency.
minimum risk detectable of 1.5), the number of subjects needed to study ALL and AML would be 277,781 for 30 percent of subjects exposed, 446,633 for 15 percent of subjects exposed, and 1,180,059 for 5 percent of subjects exposed.

**Progress.** In the last 2 years, the I4C has (1) established expert working groups for key hypotheses, (2) developed a core protocol package to support new cohorts and encourage harmonization across cohorts, (3) obtained formal agreements to participate from key cohorts, and (4) transferred data from cohorts to enable the pooling process.

**Workshop Goals.** Dr. Dwyer listed the workshop’s goals:
- Inform participants about progress in data collection in existing cohorts and plans for new cohorts
- Inform participants about progress with data pooling and opportunities for hypothesis testing
- Provide information on scientific developments in the field
- Enable discussion in working groups about progress with analyses and next steps
- Discuss future directions for the I4C
- Decide on date and location for the next I4C workshop.

**Identifying Risk Factors for Childhood Cancer: Progress in the I4C Cohorts**

*Moderator: Dr. Dwyer*

**Risk Factors for Childhood Cancer: An Update**

*Martha S. Linet, M.D., M.P.H., NCI, NIH, USA*

**Characteristics and Descriptive Features.** In 2007, there were an estimated 10,400 total cancer cases and 1,545 deaths for children younger than 20 years (SEER Cancer Statistics Review 1975–2008). The highest incidence rates were 36 per million for leukemia, 26 per million for brain/central nervous system cancers, and 24 per million for lymphoma (SEER Cancer Statistics Review 1973–1998). The 5-year survival rate for childhood cancer is currently 80 percent.

Childhood cancer types vary by age, gender, and race. From 1975 to 2004, the mortality rates for total childhood cancer decreased. However, during that same period the incidence of total childhood cancer increased modestly, but has been stable since the early 1990s. It is estimated that 60 percent to 70 percent of childhood cancers are due to gene-environment interactions.

**Known Risk Factors.** The known risk factors for childhood cancers are as follows:
- Leukemia—high-dose single exposure to ionizing radiation (from the atomic bombings of Hiroshima and Nagasaki, although the association was more apparent for schwannoma than for gliomas); prenatal diagnostic x-ray exposures; chemotherapy (for example, alkylating agents and epipodophyllotoxins); and genetic and constitutional disorders (for example, Down syndrome, Bloom syndrome, ataxia telangiectasia, and neurofibromatosis type I)
- Brain tumors—high-dose single exposure to ionizing radiation (for example, atomic bombs); fractionated radiotherapy; and genetic and constitutional disorders (for example, neurofibromatosis, nevoid basal cell carcinoma syndrome, tuberous sclerosis, and familial Li-Fraumeni syndrome)
- Other pediatric cancers—Epstein-Barr virus (linked with Burkitt’s lymphoma); HIV, immunodeficiency syndromes, and organ transplant (non-Hodgkin lymphoma); mutation in
RB gene (retinoblastoma, hereditary); low birth weight (hepatoblastoma); and Beckwith-Wiedemann syndrome and idiopathic hemihypertrophy (Wilm’s tumor).

**Recent Findings.** Recent studies reported on associations between childhood leukemia and the following:
- Birth weight
- Folic acid, $MTHFR$ gene variants, and diet
- Maternal diagnostic x-rays
- Pesticides
- Chemicals in house dust
- Residential proximity to chemical sources
- House painting
- Allergies
- Timing of chromosomal translocations and with mutations
- Inherited susceptibility for B-precursors from GWA studies
- Biologic plausibility of variation in the $CDKN2A$ gene.

**How the I4C Can Contribute to Understanding and Validating Recent Findings.** A prospective birth cohort study design has several methodological advantages over a case-control design. The cohort approach yields new insights and can identify mechanisms. The cohort approach also may decrease selection bias. Compared with retrospective exposure assessment, the cohort approach can minimize recall bias. Limitations of observational studies include difficulty in identifying small risks. Findings require confirmation, but and associations identified using different designs (case-control and cohort approaches) increases the plausibility that such associations are real.

**Cancer Ascertainment through Cancer Registries**
*Tone Bjørge, M.D., Ph.D., University of Bergen, Norway*

There are four main concepts in assessing data quality:
- Comparability—the extent to which coding and classification procedures at the cancer registry, together with the definitions of recording and recording specific data items, adhere to agreed international guidelines
- Completeness—the extent to which all diagnosed neoplasms are included in the registry database
- Validity—the proportion of cases in the cancer registry with a given characteristic that truly have this attribute
- Timeliness—evaluated in terms of the time from diagnosis to registration, and the time from registration to the reporting of incidence data.

The main types of cancer registries are:
- Population-based registries—These describe the extent and nature of the cancer burden in the community (incidence, prevalence, mortality, and survival). They are used as a source of material for etiological studies, and they can help in monitoring and assessing the effectiveness of cancer control activities.
- Hospital-based registries—These contribute to patient care and provide data on treatment and results (data for administrative purposes and reviewing clinical performance).
- Cancer specific/clinical registries—These are extended registrations of diagnostic measures, therapy, and follow-up. In Norway, registries have been established for eight types of cancer.

The basic requirements for cancer registration are (1) a clear definition of the catchment population; (2) availability of reliable population denominators; (3) generally available medical care and ready access to medical facilities; and (4) easy access to case-finding sources such as hospitals, pathology departments, death certificates and other sources of clinical data. The main sources of data collection include information from treatment facilities, information from diagnostic services, and death certificates.

Cancer “incidence” can be defined as the date of first consultation at, or admission to, a hospital, clinic, or institution for the cancer in question; the date of first diagnosis by a physician or date of the first pathology report (for cases not admitted to or seen at a hospital); or date of death (for cases first diagnosed at autopsy, which were unsuspected during life or cases reported with a death certificate only). The most valid basis of diagnosis is of great interest in assessing the quality of the data. The minimum requirements of a cancer registry are to discriminate between tumors that were microscopically verified and those that were not and distinguish between neoplasms that were diagnosed on the basis of a clinical history plus other investigations. The recommended reference that cancer registries should use to code topography and morphology of tumors is the *International Classification of Diseases for Oncology (ICD-O)*.

Cancer registries should emphasize quality of data collected rather than data quantity. When evaluating data quality, cancer registries should carefully consider completeness (the extent to which every incident case of cancer is identified) and validity (the extent to which the information recorded on the different variables is true, or accurate).

Reporting results should include incidence (the number of new cases, average annual number of new cases, and age-adjusted incidence rates per 100,000 person-years), mortality (the number of cancer deaths), survival (the relative and conditional survival), and prevalence of cancer at different periods. Data can be stratified, for example, by primary site, sex, county, stage, age or age groups, period of diagnosis, and year(s).

Cancer registries can provide information on the distribution of cancer in a population and data on cancer occurrence in epidemiological studies. Cancer registry data can be used in a wide variety of areas of cancer control, including etiological research, primary prevention, secondary prevention, health care planning, and patient care. In planning for cancer control programs, the following should be considered: The annual number of new cases indicates resources needed for treatment, the number of prevalent cases describes the need for long-term follow-up, and cancer registry data are important sources for making predictions of cancer incidence.

Dr. Bjørge described the data quality of the Cancer Registry of Norway (2001–2005):
- Comparability—The coding and classification system follows international standards.
- Completeness—The overall completeness was estimated at 98.8 percent (2001–2005); there was a variable degree of underreporting for hematological malignancies and tumors of the
central nervous system.

- Validity—Of the cases, 93.8 percent were morphologically verified.
- Timeliness—The underreporting in 2005 due to timely publication is estimated at 2.2 percent overall, based on the number of cases received at the registry during the following year.
- Conclusion—The routines at the Cancer Registry of Norway yield comparable data that can be considered reasonably accurate, close to complete, and timely.

Harmonization of Key Risk Factors for Childhood Cancer from I4C Cohorts

Gabriella Tikellis, Ph.D., MCRI, Australia

At the Third I4C Workshop in Lyon, France, in November 2009, work was initiated as part of the folate hypothesis. The prevalence of key exposure variables across five cohorts—the Tasmanian Infant Health Study (TIHS), the Avon Longitudinal Study of Parents and Children (ALSPAC), the Danish National Birth Cohort (DNBC), the Birth Defects Surveillance System for the Collaborative Project-China (BDSS-China), and the Norwegian Mother and Child Study (MoBa)—with folic acid data was surveyed. A preliminary pooled analysis based on data from TIHS and ALSPAC was conducted to show “proof of principle.” Since then, the I4C has identified common variables among the cohorts, transferred cohort data to the I4C International Data Coordinating Center (IDCC), established a database management system, and begun analyzing the pooled data.

Harmonization of the variables at the I4C IDCC involves detailed examination of questionnaires from the cohorts and documentation of the details of the type of data available for factors associated with childhood cancer. The aim of this harmonization effort is to create a pooled data set for each hypothesis or research question being examined within the I4C.

As a result of the harmonization work, an I4C New Cohort Protocol Support Package (NCPS) was developed. The NCPS provides a support guide for the collection of data relating to key ‘environmental’ (in the broad sense) exposures associated with childhood cancer. It can be accessed from the NIH I4C Portal at https://communities.nci.nih.gov/i4c.

In addition to the TIHS and ALSPAC data, data from the Jerusalem Perinatal Study (JPS) and the Collaborative Perinatal Project (CPP) were transferred to the I4C IDCC in 2011. In September 2011, DNBC approved the transfer of data to be used for the folate proposal; approval for data to be used for the birth weight proposal is pending. The DNBC data set consists of 151 cancer cases (58 ALL) and a random selection of the cohort (10 percent, n = 8,803). Linkage to the MoBa cancer registry is completed. A data set consisting of 106 cancer cases (33 ALL) and a random selection of the cohort (10 percent, n ~10,000) is being compiled. Use of the data for the folate and birth weight proposals has been approved. As of September 2011, the total number of childhood cancer cases at the I4C IDCC was 528; of these, 171 were any leukemia and 134 were ALL.

I4C working groups have proposed several studies. Current hypotheses involve maternal prenatal folic acid supplementation and the risk of childhood cancer, paternal age and childhood cancer, birth weight and childhood cancer, and pesticide exposure and childhood cancer.
The I4C IDCC has identified 22 variables that are common amongst the hypotheses. The I4C working groups involved in these studies will expand the harmonization and data pooling efforts. In addition, a Web-based data pooling application is being developed at the I4C IDCC.

For the birth weight and childhood cancer hypotheses, data from four cohorts (TIHS, ALSPAC, CPP, and JPS) have been pooled. The analysis will be based on a case-cohort design. The data set includes 130 cases of childhood cancer, of which 33 are any leukemia and 25 are ALL.

The next steps for the birth weight study are to (1) complete the transfer of additionally requested data from ALSPAC and JPS to add additional covariates/confounders, (2) incorporate data from MoBa and DNBC, (3) complete the cleaning and harmonization of data, (4) consolidate the pooled data set with a data dictionary, and (5) begin formal analysis of pooled data from six cohorts. The epigenetics portion of the birth weight study is awaiting biospecimens. The maternal prenatal folic acid supplementation study recently received data from DNBC and has data pending from MoBa to be incorporated into a pooled data set. The pesticide occupational exposure study is working on standardization of a job matrix and identifying cohorts with relevant data.

**Childhood and Young Adult Cancer in the JPS**

*Ora Paltiel, M.D.C.M., M.Sc., Hadassah-Hebrew University, Israel*

The JPS research cohort was established between 1964 and 1976. The cohort consisted of all births to residents of West Jerusalem (92,408 children, 42,956 mothers, and 39,620 fathers). The JPS involved active surveillance of infant mortality, birth defects, and pregnancy complications within the cohort.

The JPS data set has been analyzed to determine the extent to which events during early human development or peripartum affect cancer incidence and mortality over a lifetime in the Israeli population. The results of several studies are as follows:

- **Birth weight**—In an analysis of birth weight and cancer, there was a consistent relationship between high birth weight and acute leukemia. Mothers of higher birth weight infants had a higher relative risk of leukemia.
- **Congenital malformations**—Preliminary results show a higher relative risk of cancer in children with severe malformations. Children with severe malformations had a higher relative risk of solid tumors, breast cancer, and sarcoma. In an analysis of the relationship of congenital anomalies to cancer in siblings or parents, there was no association in siblings and an association between cleft lip/palate and genitourinary cancer in parents.
- **Hospitalization for infection in infancy**—An evaluation of the relationship between hospital admission in the first year of life due to infectious disease and the risk of developing malignancy in childhood and early adulthood found a threefold increased risk of non-Hodgkin’s lymphoma after first-year hospitalizations due to infectious disease. Hospitalizations for infection were related to lower birth weight, higher birth order, lower socioeconomic status (SES), and less educated mothers.
- **Familial cancer**—In a study of 16 families, 25 percent of mothers whose children had cancer also had malignancy. In half of the families, cancer sites were concordant for lymphoma, breast cancer, ovarian cancer, testicular cancer, and melanoma.
The JPS provides a data set for a unique population and unique research opportunities. The unique findings so far, require confirmation in other populations and offer potential for collaborations. Study results to date indicate that exposures during the pre-, peri-, and postnatal periods have an impact on parents’ and offspring’s cancer risk.

Questions and Discussion

The following issues and topics were discussed:

- Accounting for differences in birth weight among different ethnicities
- What is considered “normal” birth weight among different ethnicities
- Accounting for ethnic differences in “normal” birth weight in data harmonization
- Using DNA biomarkers to inform ancestry, not whole genomes
- Accounting for effects of exposures during pregnancy on miscarriages and stillbirths
- Focus of cancer registries only on live births
- Future data collection on complete birth information, including early fetus loss, reproductive history at enrollment, and breastfeeding
- Measuring early childhood growth trajectory
- Considering birth weight for gestational age.

Use of Blood Spots for Research

Moderator: Zdenko Herceg, Ph.D., IARC, France

Investigating the Utility of Archival Birth Blood Spots for (Epigenetic) Epidemiology: An Overview

Richard Saffery, MCRI, Australia

Introduction and Definitions. Epigenetics is the study of factors “imposed” on DNA that alter gene expression without changing underlying DNA sequence. Altered gene expression is heritable through cell division, a so-called “cellular memory.” There are multiple types of epigenetic modifications. The epigenome is the total epigenetic state of a cell. Unlike the genome (that is, DNA sequence), there is no single epigenome of an individual, organ, or even tissue.

DNA Methylation. DNA methylation features very stable covalent modification, which can be measured in any DNA sample. It is the binary variable equivalent to a single nucleotide polymorphism (SNP) at specific CpG sites. Each cell has two copies of specific sites, each of which can have two states, methylated or not. Each gene can have many sites that act in a coordinated way to regulate gene expression. Each cell in a biospecimen has a specific profile that contributes to measurement output of a biological sample, which reflects the percentage of specific sites within the sample that are methylated.

Challenges for Epigenetic Epidemiology. There are multiple challenges for epigenetic epidemiology. Ethical and sample quality issues of archived clinical specimens need to be considered. Previously identified challenges include:

- Existing archived biospecimens are valuable for a limited range of epigenetic profiling.
- Because epigenetic profile changes over time, particularly in early development, cause and
effect cannot be inferred as per genetic epidemiology.

- Epigenetic profile varies between different cell and tissue types, which bring into question the relevance of “accessible” tissues to diseases of interest.
- Many incredibly rich cohort studies began years before epigenetics was considered relevant to complex disease etiology. However, most did not collect “predisease” biospecimens hindering the examination of the role of epigenetics in the causal pathway to disease.

Different Measures from Archived Neonatal Dried Blood Spots (DBS). Archived neonatal DBS may be used for analyses of proteins; vitamins, micronutrients, and other nutrients; steroids; RNA; and DNA. In addition, whole-genome amplified DNA from archived neonatal DBS samples may be used for reliable genome-wide scans and is a cost-efficient alternative to collecting new samples.

DNA Methylation Analysis from Archived Neonatal DBS. Considerations for DBS DNA methylation analysis and data interpretation include:

- The age of DBS correlates with the level of DNA degradation.
- The storage environment (including temperature) is important.
- The method of DNA extraction is critical.
- Blood composition is variable and may associate with pregnancy-associated factors (for example, maternal and gestational age). Currently, there is no way to determine these associations.

Current Status and the Future. Genome-scale methylation from DSB is a reality but requires a large proportion of individual DBS or amplification of DNA. More than $1 billion is currently being used to develop epigenomic technologies, including those related to the use of minute samples. It is predicted that within 3–5 years, all archival DBS will be amenable to genome-wide DNA methylation analysis.

Profound Deficit of IL10 at Birth in Children Who Develop ALL: New Research Using DBS

Joseph Wiemels, Ph.D., University of California, San Francisco, USA
Presented by Carol H. Kasten, M.D., NICHD, NIH, USA

DBS are a unique “common data element” collected at the same time (1–3 days postdelivery), in the same way (heel stick), and using the same instrument (filter paper). Only DBS storage conditions differ. They contain the child’s blood only and multiple other analytes that reflect prenatal conditions such as DNA, hemoglobin, albumin, cytokines, and nutritional status.

Several studies have found a correlation between childhood infectious exposures and subsequent acute leukemia. However, there is a paradox in the findings. Children with leukemia have lower prior exposure to infection. When they contract infection, children who get leukemia respond more strongly to the infection and go to the doctor more often. Researchers have hypothesized that children who get acute leukemia are born with differently tuned immune systems.

At the Brocher Foundation Symposium held on March 7–8, 2011, Dr. Wiemels presented preliminary results from the California Childhood Leukemia Study. Analyses were conducted on
DBS from the State of California Repository. If validated, the results will provide the first quantitative glimpse of what may be one of the antecedents of childhood leukemia, modifiers of conversion to leukemia may be more rapidly identified, and surveillance and ultimately intervention to prevent ALL may be possible.

The California Childhood Leukemia Study’s hypothesis was as follows: Children who grow up to get leukemia have a detectably different immune cytokine profile at birth. The study examined Guthrie cards from 116 leukemia (ALL) cases and 116 birth date-, gender-, and ethnicity-matched controls. Only five cytokines were detectable: IL4, IL6, IL10, IL12, and IL13. A univariate analysis found that IL4, IL6, IL10, and IL13 were significantly lower among childhood ALL cases than controls. Adjusting for the potential confounders, increased levels of IL4, IL6, IL10, and IL13 were all associated with a significantly decreased risk of childhood ALL. A multivariate analysis of the five detectable cytokines, using the same statistical model, showed that only IL10 remained statistically significant.

IL10 is produced by monocytes, regulatory T-cells, and B-1 cells. IL10 suppresses inflammatory Th1-type responses and enhances B-cell development and function. The critical role of IL10 in pregnancy is suppression of the immune system rejection of the fetus. Preliminary results from the California Childhood Leukemia Study revealed that children who get ALL are profoundly deficient in IL10 at birth. This congenital defect may make them more susceptible to strong reactions to infections and increased risk of leukemia.

The data suggest that children who develop leukemia may already have dysregulated immune function at birth. Future studies need to confirm this association and consider maternal and genetic factors in the development of childhood leukemia through their influence on the child’s immune development.

**Brocher Foundation Symposium Recommendations**

*Dr. Carol Kasten*

The Brocher Foundation symposium, titled “Ethical Issues in the Use of Neonatal Screening Bloodspots in International Population-Based Studies of Childhood Cancer,” was held on March 7–8, 2011, in Geneva, Switzerland. The symposium participants considered the secondary use of DBS and ancestry informative markers (AIMs) in the context of international epidemiological childhood cancer research. They recognized that DBS are unique and that differences in public engagement, ethical perspectives, and policies on DBS use and storage are major challenges in using DBS collaboratively. The symposium participants recommended the following:

- DBS should be stored for decades after initial screening has been performed.
- DBS collection and storage should be carried out according to local laws and stringent authorized procedures to ensure safety and security.
- Symposium participants will work as a group to prepare international guidelines on parental consent and the secondary use of de-identified DBS, using these guidance recommendations as templates.
- In international data pooling studies of childhood cancer using DBS for genetic and environmental analyses, the use of AIMs should be considered to control for confounding by genetic ancestry.
Ethical issues inherent to identification of ethnicity, such as discrimination based on one’s genetic and geographic ancestry, should also be considered.

The use of AIMs is appropriate in international data-pooling studies of childhood cancer; robust discussion of this opinion in scientific communities and the lay public must be promoted.

DBS must first and foremost serve their primary function as a screening tool for treatable congenital disorders, including infectious diseases in some countries.

Low and middle income countries should consider broadening their screening programs to include DBS for research as sustainability and resource allocation permits.

The consensus among symposium participants was as follows: There is an urgent need for internationally accepted ethics policies for large data-pooling studies on childhood cancer using DBS and AIMs.

Symposium participants emphasized the importance of engaging the public in developing internationally accepted guidelines to facilitate the use of DBS—and AIMs where appropriate—in data-pooling research.

Questions and Discussion

The following issues and topics were discussed:
- The “costs” of genotyping test results
- Sensitivity to ancestry based on genotyping
- Ethical and confidentiality issues of ancestral identification
- Recommendations of DBS storage conditions
- Other tissues for analysis (for example, placenta, cord blood, red blood cells, and monocytes)
- Collection and analysis of cord blood (for example, for metabolic disorders)
- Validation studies on the use of DBS for genetic analyses
- Comparison of methodologies.

Updates from the I4C Working Groups

Moderator: Manolis Kogevinas, M.D., Ph.D., CREAL, Spain

Update from the I4C Epigenetics Working Group

Hector Hernandez-Vargas, IARC, France

The objectives of the Epigenetics Working Group’s study of birth weight and childhood cancer are to define an epigenetic signature of the risk of childhood leukemia in blood obtained at birth and discover an epigenetic signature of cancer risk in blood obtained from high birth weight babies. The study’s hypothesis is as follows: Environmental exposure during early embryonic life is able to imprint an epigenetic signature that can be used as a biomarker of exposure and susceptibility to cancer.

There are several reasons to study epigenetics in childhood cancer:
- Translocations frequently target epigenetic mechanisms.
- DNA methylation markers are commonly deregulated in ALL.
- Hot spots for translocations are common in CpG-rich regions.
- Translocations frequently target epigenetic mechanisms.
  - Translocations frequently target chromatin modifiers.
  - Fusion proteins are involved in the recruitment of silencing complexes.
- DNA methylation marks are commonly deregulated in ALL.
- Hot spots for translocations are common in CpG-rich regions.

The approach for the study of birth weight and childhood cancer includes:
- 200 archived blood spots (stratified according to birth weight)
- DNA extraction, bisulfite modification, and whole genome amplification
- Epigenome analyses (using Infinium Human Methylation 450K)
- Bioinformatics to identify epigenetic signature of high birth weight
- Validation and replication.

This proof-of-principle approach will provide a starting point from which to explore the association of multiple environmental exposures to an epigenetic profile linked to childhood cancer. The identification of reversible epigenetic alterations associated with environmental cues may have a strong impact in understanding and preventing cancer.

The I4C Epigenetics Working Group currently consists of:
- Jia Chen, Sc.D., Mount Sinai School of Medicine, USA
- Jeff Craig, Ph.D., MCRI
- Terence Dwyer, MCRI
- Zdenko. Herceg, IARC
- Hector Hernandez-Vargas, IARC
- Rayjean J. Hung, Ph.D., M.S., University of Toronto, Canada.
- Yoshimi Inabi, MCRI
- Carol Kasten, NCS
- Sharon Savage, NCI
- Camilla Stoltenberg, National Institute of Public Health
- Gabriella Tikellis, MCRI
- Joseph Wiemels, University of California, San Francisco
- Nicholas Wong, MCRI

Update from the I4C Environmental Working Group on Birth Weight and Childhood Cancer

Dr. Ora Paltiel

The I4C provides a unique opportunity to examine childhood cancer associations using rich prospectively collected data, taking into account a large variety of covariates and modifiers of birth weight. The temporal and geographical diversity of cohorts participating in the I4C will allow the analysis of secular trends in the birth weight–cancer association as well as geographic and ethnic variations in this association.

The aim of the working group’s study on birth weight and childhood cancer is to investigate the
association between birth weight and other measures of fetal growth and childhood cancer, specifically leukemia, with specific attention to determinants of birth weight—such as maternal obesity, weight gain in pregnancy, and pregnancy complications—in a pooled analysis of childhood cancer cohorts.

The proposed study’s specific objectives are as follows:
- Examine the pattern of the association between birth weight and other measures of fetal growth and the risk of childhood AML, ALL, all leukemias, and other cancers.
- Examine birth weight both as a continuous and categorical variable.
- Examine these associations in specific age groups. The age groups are infant (up to age 1 year), early childhood (1–4 years), later childhood (5–9 years), and early adolescence (10–14 years)—controlling for and in strata of maternal prepregnancy body mass index (BMI) and weight gain during pregnancy.
- Determine whether the associations are consistent over time and across ethnicities and geographic groups using data from the various cohorts with inception times in different epochs.
- Investigate the hypothesis that diabetes in the mother, especially type 2 diabetes mellitus and gestational diabetes mellitus, increases the risk of cancer and that no such association is seen for paternal diabetes.

With regard to feasibility, basic data on birth weight, gender, birth order, and cancer incidence are available in all cohorts. In JPS, data on gestational age, prepregnancy BMI, and weight gain during pregnancy are only available for a subcohort. The participating cohorts are ALSPAC, TIHS, JPS, CPP, DNBS, and MoBA. From these cohorts, the total number of live births is 283,585, the total number of cancer cases is 528, and the total number of ALL cases is 134. Of the 283,585 births, 8.6 percent had a birth weight greater than 4,000 grams.

A case-cohort analysis of pooled data (n = 2,323) from ALSPAC, TIHS, CPP, and JPS has been conducted. The preliminary data show a higher incidence of all childhood cancers, any leukemia, and ALL for birth weight greater than 4,000 grams. The next steps in the analysis are to:
- Analyze birth weight as a continuous and categorical variable
- Add more covariates to analysis
- Stratify analysis by maternal prepregnancy BMI and weight
- Stratify by weight gain in pregnancy
- Use more sensitive indicators of fetal growth.

Update from the I4C Environmental Working Group on Parental Occupation in Farming and Childhood Cancer Risk in the I4C

Ann Olsson, Ph.D., IARC, France

Although case-control studies are prone to information and selection bias, the majority have shown an association between pesticide exposure and childhood cancer risk. These studies have shown stronger associations with indoor pesticide exposure. However, there have been internal inconsistencies (for example, lack of dose-response effect). The few cohort studies that have been conducted have shown no or weaker associations.
The Parental Occupation in Farming and Childhood Cancer Risk study will pool I4C data to investigate whether parental—both maternal and paternal—occupational exposures in different types of farming are associated with increased risk of childhood cancer in their offspring. Benefits to a pooled prospective cohorts study design include: data collection prior to diagnosis, data collection closer after cancer occurrence, and increased power to study “rare events.”

The objectives of the study are to:

- Investigate whether parental occupations in different types of farming are associated with increased risk of cancer in their offspring
- Evaluate the risk of childhood cancer associated with self-reported parental use of pesticides during pregnancy and compare type and frequency reported in case-control studies
- If possible, assess whether the risk of childhood cancer associated with parental exposure to occupation as a farmer and pesticides varies by exposure time-windows (preconception, prenatal, and postnatal).

The currently eligible I4C cohorts are ALSPAC, TIHS, DNBS, and MoBA. These cohorts have 231,600 births and more than 260 childhood cancer cases. The next steps are to determine the eligibility of other cohorts, transfer data to the I4C IDCC, harmonize disease and occupational data, analyze data, and draft manuscripts.

The study team currently consists of:

- Ann Olsson, IARC
- Joachim Schüz, Ph.D., IARC
- Kurt Straif, M.D., Ph.D., M.P.H., IARC
- Gabriella Tikellis, MCRI
- Martha Linet, NCI
- Martine Vrijheid, CREAL
- Principal investigators (PIs) of eligible cohorts.

The I4C Environmental Working Group Proposal: Agricultural Pesticide Exposures and Risk of Childhood Cancers

Mary Ward, Ph.D., NCI, NIH, USA

Most studies of childhood cancers (for example, leukemia, lymphoma, and brain cancer) have observed increased risk associated with parental occupational pesticide exposure. Meta-analyses of leukemia point to maternal occupational exposure and residential exposure during pregnancy as being most important. For occupational exposure, farm-related pesticides showed somewhat stronger risks than other occupational pesticides. Residential proximity to agricultural pesticides has been associated with elevated rates in some but not all studies. Only one case-control study evaluated the association of proximity of exposure and childhood cancer risk. However, no similar cohort studies have been conducted. Exposure studies have indicated that distances of up to 1,000 meters and periods up to 2 years may be important in predicting pesticide levels in homes.

The aims of the proposed Agricultural Pesticide Exposures and Risk of Childhood Cancers study are to (1) evaluate the risk of childhood cancers in relation to residential proximity to agriculture
during pregnancy and early life and (2) determine pesticide use around the birth residence and the association with childhood cancers. Parental occupational pesticide exposure will be considered.

The proposed study design is as follows:
- Case-cohort study, with 10 controls per case
- Cohorts with birth addresses for geocoding
- Land cover maps and crop-specific pesticide use data from each country for birth year
- Adjustment for parental occupational pesticide exposure.

Study design issues include:
- Geocoding of birth addresses (that is, centrally versus by each cohort)
- Need to create crop maps for some areas/years
- Sample size currently available
- Current cohorts and cancer cases.

The currently eligible cohorts and years of data collection are ALSPAC (1991–1992), TIHS (1987–1995), JPS (1964–1976), CPP (1960–1965), DNBS (1996–2002), and MoBA (1999–2007). From these cohorts, the total number of live births is 380,425, the total number of cancer cases is 228, and the total number of ALL cases is 102.

The study’s lead investigators are Dr. Ward and Leslie Stayner, Ph.D., M.S., University of Illinois at Chicago, USA. The co-investigators are:
- Dr. Vrijheid, CREAL
- Dr. Tikellis, MCRI
- Dr. Linet, NCI
- Dr. Straif, IARC
- Michael Dellarco, Dr.P.H., NICHD, NIH.

Questions and Discussion

The following issues and topics of the working groups’ reports were discussed:
- Time lapse after birth weight and childhood cancer data collection (1-year data are still valuable)
- Serial reporting of birth weight and childhood cancer findings (findings will be reported when there is reasonable statistical power)
- Children’s age at diagnosis (cohorts with data after 1 year of age can contribute to birth weight and childhood cancer data)
- Effects of low birth weight
- Use of parental occupation job modules
- Validity of exposure assessments
- Relationship of in-house dust collection data and self-reported exposure data
- Detail of information on type of farming and type of exposure
- Self-reported home and garden pesticide exposure
- Impact of differences in data collection periods (for example, 1960–1965 versus 1999–2007) and types of pesticides on exposure assessment
- Use of crop maps, pesticide registries, and cancer registries.

**Breakout Sessions**

The Epigenetic Working Group and Environmental Working Groups (pesticide and birth weight groups combined) convened separate breakout sessions to discuss issues related to their studies and make recommendations on future work.
Day 2

Cancer and Epigenomics
Manel Estellar, Ph.D., Bellvitge Medical Research Institute (IDIBELL), Spain

In an individual, each cell has the same genome, with the same genes and DNA sequence. However, cells in different tissues have different phenotypes. Although these cells have the same levels of DNA, RNA, and proteins, the expression and ability of the genome changes. Embryonic cells develop into tissues with different epigenomes. Starting from identical genetic sequences, changes in histone modifications and DNA methylation can produce organisms with different features and distinct susceptibility to sickness. An example is monozygotic twins. Monozygotic twins have the same DNA but different epigenetics. They can have differences in phenotype and gene penetrance due to changes in histone modifications, DNA methylation, and RNase. Environmental exposures, such as smoking, can change epigenomics and susceptibility to disease.

Cancer is an epigenetic disease characterized by the breakdown of the DNA methylation and histone modification patterns. The stability of the genome and correct gene expression are maintained in large measure thanks to a perfectly pre-established pattern of DNA methylation and histone modifications. Different parts of DNA can be altered by methylation. Genes have promoter regions that are rich in CpG (CpG “islands”) and affect expression of neighboring genes. In normal cells, these regions stay unmethylated, and the genes are expressed. In cancer cells, the CpG islands become hypermethylated, thereby silencing genes. Hypermethylated tumor genes are silenced by this epigenetic process. There are now markers based on methylation in cancer patients. For example, these markers have been used for the screening of prostate cancer.

The mechanisms underlying the disruption of the epigenetic landscape in transformed cells are unknown. It is possible that the enzymes that epigenetically modify DNA and histone are themselves targets of genetic disruption. Complications in genetics include mutations and histone modifications. In the nucleus of a cell, DNA fits into protein packets, which are different in different cells. The proteins can be modified by histone modifications and DNA methylation.

In the last few years, micro-RNAs have started a revolution in molecular biology and emerged as key players in the cancer process. For these reasons, it is extremely important to understand the physiological and disease-associated mechanisms underlying the regulation of these small, single-stranded RNAs. RNase modifications can affect many genes. In cancer cells, there are mutations in the machinery that produce micro-RNA, kinase, and receptors. Proteins involved in histone modifications are mutated in tumor cells. The mutated genes in tumor cells are epigenetic genes, such as histone modifiers, chromatin remodelers, and DNA methyltransferases. Drugs that target these enzymes can be used in cancer treatment. Five drugs are now approved for the treatment of two types of leukemia and two types of lymphoma.

New epigenomic techniques to unmask tumor suppressor genes in human cancer are being developed. These techniques can be used to look at the whole genome to find changes in DNA methylation and candidate genes. Every tumor type has its own DNA methylation profile. Micro-
array-based DNA studies looking at the pattern of gene expression have found candidate genes for breast, ovarian, and other tumor types. Matching DNA profiles can be used to determine primary tumor type. New large-scale epigenomic technologies might be useful in an attempt to define the complete DNA hypermethylome of tumor cells.

Questions and Discussion

The following issues and topics were discussed:

- Changes in DNA with aging
- Effect of aging on methylation profile
- Cause of epigenetic changes in tumor cells
- Methylation in DNA repair genes and modulation of mutation
- Specificity of DNA methylation profiles due to exposures to environmental toxins/carcinogens.

Presentations and Recommendations from the Breakout Sessions

Epigenetics Working Group. The working group discussed the following topics and issues:

- Other environmental exposures that could be related to epigenetics of cord blood
  - Maternal folate intake
  - Folate levels in maternal blood
  - Pesticide/herbicide exposure
  - Infections
- Linking epigenetic profile changes at birth with later profiles of cancer
  - Need for epigenetic profiles in cancer registries
  - Types of drugs used to treat cancer
  - Profiles after successful cancer treatment
- Possible genomic analyses.

Dr. Linet made the following comments:

- Measuring folate levels in maternal blood may not be productive and may be a waste of resources.
- The working groups could be more effective by communicating through the portal (for example, scheduling and announcing conference calls).
- The NCI is conducting a number of adult cancer consortia studies. In an effort to better manage consortia workload, submission of requests, proposals, and inquiries are now limited to certain periods only. Requests to I4C consortia could be consolidated similarly.

A participant commented that up to 90 percent of women take more than 1 milligram/day of folate. Studies have measured folate levels, and the findings are in the literature. An area of interest may be folate as a food additive. Studies so far have not shown an effect of folate on epigenetic processes. These types of studies are complex and challenging.

Environmental Working Groups. In a combined session, the working groups discussed the following topics and issues:

- The need for more detailed information on occupation in order to get better coding
Specific questions on job type and work environment
- Job at time of birth; job before birth
- Cohort survey on the type of occupation information available
  - Landsat satellite information
    - Need to provide details on the type of information obtained
    - Whether institutional review board approval is needed for address identification
    - Process of delinking geocoding data
    - Linking Landsat information to other environmental databases
  - Maternal BMI information
  - Lack of detailed information on paternal smoking
  - Cohorts’ data on maternal diabetes and children’s growth trajectories
  - Collection of biospecimens in the birth weight study
  - Publishing a paper on birth weight distribution among the I4C cohorts using existing data
  - Justifying requests for/use of biospecimens
  - Biospecimen analysis location and standardization.

New Working Group. A new clinical working group was proposed to address clinical aspects related to exposures (for example, infection) and cancer outcomes.

Development of New Cohorts
Moderator: Dr. Martha Linet

Europe

Concept for a German Environmental Health Birth Cohort: Design and Feasibility Issues
Börge Schmidt, University of Duisburg-Essen, Germany

The objectives for a German Environmental Health Birth Cohort are to study a broad range of environmental health problems with regard to socioeconomic aspects and migration, individual exposure assessment (for example, human biomonitoring), and detection of gene-environment interactions. Current German birth cohorts are medium-sized and focus on special aspects of child health. A large-scale birth cohort (100,000–200,000 participants) would serve as a platform for future research.

Progress to date includes reviewing existing birth cohort studies, developing a first idea for design issues and framework, generating possible research questions, determining design and main research questions, developing the exposure and outcome assessment, and elaborating the overall study concept.

The main research issues are the impact of the following:
- Environmental pollutants and noise on neurodevelopment and cognitive abilities
- Endocrine disruptors on reproductive development
- Environmental pollutants on pregnancy and birth outcomes
- Indoor/outdoor air pollution and inhalation allergens on asthma, allergies, and wheezing
- Environmental pollutants on obesity, insulin resistance, and diabetes.
The exposure measurements of interest are:
- Chemical exposures—organic/inorganic chemicals, and air pollutants
- Physical exposures—housing/neighborhood characteristics, and noise
- Biological exposures—allergens, bacteria, and infections
- Genetics—DNA and gene expression
- Psychosocial exposures—demographics, family support, and health behavior
- Biospecimens—blood, urine, breast milk, cord blood, placenta, meconium, and nails/hair
- Indoor/outdoor air pollution—allergens, particulate matter, and volatile organic compounds
- Noise—noise maps, models, and questionnaires.

The study will recruit early in pregnancy (first trimester) and enroll at different stages during pregnancy and even shortly after delivery. The recruitment strategy will use a multimodal approach via gynecologists, prenatal care providers, and maternity wards. Pharmacies and public institutions will be included. Incentives and public relations tools will be used to clearly state the benefits to participants.

Nine study locations are being considered. Selection of study locations will be based on socioeconomic, environmental, and demographic characteristics. Selection of study centers will be based on their recruitment experience. Study planners have developed a schedule for data collection, which will include medical examinations, home visits, phone interviews, and questionnaires from the first trimester to age 6 years. Follow-up is planned through age 18 years. A study timeline has been drafted.

Feasibility studies will explore effective strategies to recruit and retain participants, cooperation with local health care providers and community-based organizations, and improving response through incentives and public relations. The next steps are to announce the study locations, conduct feasibility studies in the near future, and create an expert group for ethics and privacy protection.

**European Cohorts: Overview and Lessons Learned**

*Dr. Martine Vrijheid*

Fetuses and infants are especially vulnerable to the effects of environmental contaminants. These effects may manifest themselves throughout the lifetime and even over generations. Pregnancy and birth cohort studies have played an important role in studying these effects. In Europe, there are many pregnancy and birth cohorts currently collecting a wealth of information on environmental exposure and child outcomes. However, data are often fragmented, and there is relatively little coordination to structure and consolidate scattered research. ENRIECO and CHICOS are two projects that are integrating research and coordinating the pooling and analysis of European birth cohort data.

The aim of ENRIECO is to coordinate European birth cohort research in the area of environmental exposures. Its objectives are to inventory birth cohort data; evaluate exposure, health, and exposure-response data; attempt to combine data from various cohorts; and make recommendations. Methods include inventory questionnaires to cohorts, expert working groups...
to evaluate cohort data, and workshops involving all cohorts. The participating birth cohorts have collected a variety of data on exposures, biospecimens, biomarkers, and birth outcomes. Data were collected by different methods, and not all data have been analyzed.

The aim of CHICOS is to develop a strategy for mother-child cohort research in Europe through the coordination of European cohorts. Its objectives are to inventory data from cohorts and registries, evaluate existing information, make recommendations for research action in key areas of policy concern, make recommendations for improved contribution of mother-child cohort research to policy at European level, and disseminate findings.

Most of the ENRIECO cohorts cannot or do not link exposure data to cancer. The following exposure-outcome data are currently being pooled:

- Persistent organic pollutants and birth outcomes
- Smoking, second-hand smoke, asthma, and allergies
- Alcohol consumption and birth outcomes
- Socioeconomic inequalities and birth outcomes
- Maternal occupation and birth outcomes
- Fish consumption and birth outcomes
- Maternal complications and asthma and allergies
- Obesity using BMI as predictor of body fat
- Other European Union projects: air pollution, water DBPs (disinfection byproducts), and biological agents.

Case studies are being conducted to evaluate combined data analyses using the original raw data from European birth cohorts on allergy/asthma to examine associations between indoor environmental exposures (for example, dampness/mold and secondhand tobacco smoke) and allergies, including asthma, allergic rhinitis, and eczema.

There are five steps in building analyses of pooled cohort data:

- Determining willingness to participate
- Assessing eligibility of cohorts
- Collecting individual participant data
- Harmonizing data
- Performing analyses (on individual cohorts and meta-analyses).

The successes for the ENRIECO data-pooling projects included (1) holding a kick-off meeting to establish personal contacts and interest in case studies; (2) assigning key responsibility for coordination and communication concentrated in one institution, with a single data collection and harmonization process; and (3) resolving complex data questions by telephone contacts with most cohorts in a fast and uncomplicated way.

The challenges included:

- Underestimation of time and effort for data management
- Heterogeneity of outcome and exposure data (compromises for outcome and exposure variables result in loss of information in harmonized datasets)
- Holding frequent meetings
Maintaining personal contacts.

Lessons learned are:

- There are more than 35 birth cohorts in Europe, with more than 350,000 mother-child pairs.
- Data and methods are fragmented.
- Combining information is possible and scientifically beneficial.
- Combining data from existing cohorts requires careful consideration of the aims, protocols, data, ethical issues, analyses, and management.
- Collecting, pooling, and analyzing data is time and labor intensive.

Asia

Japan Environment and Children’s Study (JECS)
Kei Mori, M.D., Ministry of the Environment, Japan

Background. From 1974 to 2004, there has been an increase in congenital anomalies such as Down syndrome, hydrocephaly, spina bifida, and hypospadias. Birth defect cases have doubled over a 25-year period. From 1980 to 2007, there has been an increase in immune system diseases. Childhood asthma has tripled over a 20-year period. There has also been an increase in psycho-neurodevelopment disorders such as children receiving medical treatment for mental and behavioral disorders. In 1997 and 2009, G8 Environmental Ministers called for policy changes and agreed to actions to improve children’s environmental health.

Pilot Study. A JECS pilot study will evaluate the following endpoints:

- Feasibility of conducting the intended research
- Limitations and problems observed in the pilot studies
- Ways of obtaining high recruitment and implementation rates in the follow-up
- Procedures that are less burdensome for health care professionals
- Data quality control system for biological samples and questionnaires.

The recruitment period was August 2008 to March 2010, with a follow-up from August 2008 to present. Four participating universities designed study methods, which will be compared in order to develop common and bias-free procedures for the JECS. A common questionnaire was developed and used by each university. The universities recruited 453 pregnant women. Biospecimens included maternal blood and urine, cord blood, umbilical cord, and breast milk.

The consent rate of pregnant women was 69 percent to 98 percent. When participants changed hospitals for delivery or had emergency deliveries, the chance of collecting cord blood was reduced. A better collection strategy is needed to increase the sampling rate for those cases. Motivations for participating in the study were “contributions to society” and “interest in the study.” Preferred methods for answering questionnaires were “filling out paper-based forms” and “a Web-based form or answering over a mobile phone.” About 60 percent of husbands agreed to donate their blood in the study.

JECS. The study’s core hypothesis is as follows: Chemical exposures during the fetal and infant stages adversely affect children’s health. The JECS will recruit nationwide 100,000 participants.
over a 3-year period, with 13 years of follow-up. The aims of the study are to:

- Identify environmental factors impacting children’s health
- Develop risk management systems that address vulnerabilities in children
- Ensure a sound environment where future generations are able to grow up in good health
- Establish infrastructures for children’s studies.

Questionnaires will be used to collect data on mothers, fathers, and children. The following biospecimens will be collected: cord blood; maternal blood, milk, urine, and hair; paternal blood; and child blood, urine, and hair.

**Chinese Children and Families Cohort Study (CFCS)**

*Peng Yin, M.D., Ph.D., Chinese Center for Disease Control and Prevention, China*

Since 1991, China and the United States have collaborated on folate studies such as the Community Intervention Program (CIP) and the Neural Tube Defects Prevention Project. Such studies provide unique opportunities to assess risk factors for pediatric and adult chronic diseases. The results of several folate studies suggest that periconception folate may reduce the risk of pediatric leukemia. There is some evidence that periconceptional folate supplements may be related to other rapidly increasing serious childhood diseases. Dietary studies suggest that early life folic acid may reduce the risk of cardiovascular diseases and certain cancers in adults.

The objectives and key questions of the CFCS focus on the role of periconceptional folate in (1) reducing the risks of pediatric leukemia or other pediatric cancers in offspring; (2) affecting the risk of pediatric asthma, autism, or other chronic diseases of children; and (3) having a protective effect on risk factors (for example, BMI and blood pressure) for cardiovascular disease or cancer in mothers.

The specific aims of the CFCS are:

- Re-establish contact with about 241,000 mothers who enrolled in the original CIP (1993–1996) to study the prevention of neural tube defects using folic acid supplementation
- Ascertain and verify all occurrences of pediatric cancers and deaths from any causes
- Examine growth, development, and developmental disabilities in this population of adolescents (ages 14–16 years).

The approach of a CFCS feasibility study involves three tasks: (1) follow up late effects of periconceptional folate in CIP mothers and offspring; (2) assess diet, nutrition, physical activity, ultraviolet radiation exposure, and associated biomarkers in CIP mothers and offspring; and (3) conduct a record linkage cohort and nested case-control study of postulated risk factors for pediatric cancer.

The objective of record linkage is to compare incidence of pediatric leukemia and other pediatric cancers in offspring of mothers who took periconceptional folic acid supplements with those who did not and assess prospective risk factors for pediatric leukemia and other pediatric cancers. The case-cohort study of the risk factors for pediatric cancer will seek hospital records for all pediatric cancers in hospitals serving the population among children born between 1994 and 1996 and match these back to the CIP population and prospectively assess pediatric
leukemia risks in exposed versus unexposed children.

Potential contributions of the CFCS to the I4C include:
- 200–236 pediatric leukemia cases in less than 15 years
- 600–800 total pediatric cancers cases in less than 15 years
- Birth weight and childhood cancer data
- Occupational exposure and childhood cancer data.

Questions and Discussion

The following issues and topics on Asian cohorts were discussed:
- The availability of stored biospecimens (for example, biospecimens are not stored in China due to issues of ownership)
- Questionnaire data collected in China
- The small number of Chinese hospitals that treat childhood cancer (seven or eight treat about 90 percent of all cases)
- Lack of data for the first 1–2 years of life
- Folate dosage and beneficial outcomes
- Lack of a national Japanese cancer registry (about 70 percent of local government have cancer registries and capture 80 percent to 90 percent of cases, with childhood leukemias combined).

Biospecimens: What Has Been Collected and What Is Available?

MoBA
*Dr. Tone Bjørge*

MoBA began in 1999 and stopped recruiting in 2008. The last baby was born in 2009. There are about 1,900 twins and 21 triplets in the study. Data were collected through questionnaires during pregnancy. Biospecimens include about 94,000 maternal blood samples from gestational week 17, maternal urine at gestational week 17, about 84,000 maternal blood samples at birth, about 19,000 cord blood samples, and about 68,000 paternal blood samples. Other biospecimens include plasma, DNA, and RNA. The biospecimens are controlled by steering committees for particular studies. There are a total of 106 childhood cancer cases in MoBA. The following lists the specific cancer sites and number of cases per site:
- Liver—2
- Ovary—1
- Testis—2
- Kidney—6
- Eye—6
- Brain and nervous system—36
- Other endocrine glands—8
- Connective tissue—2
- Lymphoma—2
- Leukemia—41.
DNBC

Data on biospecimen collection were presented in a poster.

ALSPAC

Data on biospecimen collection were presented in a poster.

The National Children’s Study
Dr. Carol Kasten

To date, the National Children’s Study has collected about 200,000 aliquotted biospecimens. There have been only 2,000 births. Maternal biospecimens include blood, RNA, urine, hair, saliva, vaginal swabs, breast milk, and placenta. Biospecimens from children include cord blood, cord sections, meconium, blood, urine, hair, and saliva. Nail clippings are no longer collected. Paternal biospecimens include blood, urine, hair, and saliva. There is no “real time” analysis of biospecimens. Data collection related to glucose metabolism includes (1) HbA1c at preconception and in the first trimester and (2) glucose, insulin, insulin-like growth factor (IGF), and IGF-BP3 in the third trimester. Some blood samples for glucose metabolism assays are collected during fasting.

Genetics, Genomics, and Epigenomics in the National Children’s Study
Dr. Carol Kasten

The National Children’s Study’s Vanguard Study—a pilot phase to assess methodologies, recruitment, logistics, and operations—has begun evaluating “legacy” data to determine the integrity and utility of aliquotted DNA, RNA, proteins, and tissues. Legacy data include family history and fetal ultrasounds. Formative research projects that are limited in scope and duration to augment the Vanguard Study are being conducted. Formative research projects will (1) address specific technical questions that can be evaluated in terms of feasibility, acceptability and cost and (2) explore new and potentially cost-effective approaches in areas such as genetics, genomics, and epigenomics.

One formative research project is investigating a dysmorphology screening tool. The project’s goal is to develop a protocol that can be used at the bed or crib by minimally trained data collectors. Digital photos of face, hands, and feet will be collected. The time to complete screening will be about 30 minutes. Analysis by nongeneticist physicians using dysmorphology screening tool software will identify birth defects.

Another formative research project involves fetal cells and DNA in the blood of pregnant women. Investigating fetal cells and maternal DNA offers a new window on human development. The goals of the project are to identify changes in fetal epigenetic and epigenomic expression; determine what cells, DNA, and RNA can currently be recovered; and ultimately link this information to environmental exposure data.
Current feasibility projects include:
- Whole genome sequencing of three National Children’s Study family trios
- Epigenome, methylation, ASM, and ASE
- Placental epigenome and histone modifications
- Premature birth genomics
- Pharmacogenomics and molecular ancestry.

Questions and Discussion

The following issues and topics were discussed:
- Versatility of aliquotted mixed red blood cells
- Telomere sequencing
- Compliance rate with biospecimen collection (about 80 percent rate for most biospecimens; about 40 percent for fathers’ hair samples)
- Formal application process for data and biospecimens requests from the National Children’s Study
- Applicability and standardization of dysmorphology tool to all ethnicities.

Genome-wide Association Studies (GWAS) versus Next Generation Sequencing (NGS)

*James McKay, Ph.D., IARC, France*

With regard to genetics, an individual’s heritable and nonheritable traits interact with the environment to produce a predicted phenotype. In terms of effects of genetic variants in cancer susceptibility, population frequency seems to impact on disease, the severity of the consequence on gene function, and the degree of functionality in gene variants. GWAS compare the frequency of a variant in cases with the variant’s frequency in controls. GWAS test all common genetic variations across the genome. So far, 21 types of cancer have been successfully identified with GWAS. These results are from four European studies with about 400 cases. The next steps are to use sufficiently large sample sizes and determine whether all SNPs are equal. Because many cancer loci are relevant to more than one type of cancer, new GWAS should start with known loci to decrease multiple testing burden. Large GWAS sample sizes (in the thousands) are needed because of the small relative risk and the large number of variants. GWAS only considers common genetic variants; rare variants are not assessed.

NGS uses massive parallel processing. This approach gives the ability to assay the entire genomic sequence of an individual. The costs for NGS are now affordable (about $3,000 for a single machine), and it has many applications other than DNA research. GWAS focus on common genetic variants, whereas NGS gives individual sequences, hence common information on rare variants. In addition, NGS data provide many short-sequence reads. Using variant filtering, 50–100 functional SNPs that are unique to each person can be identified in about 20,000 genes.

Questions and Discussion

The following issues and topics were discussed:
Effect size and heritability estimates
Insufficient number of I4C leukemia cases for GWAS
The association between viruses and Hodgkin’s lymphoma
Alternate strategies for determining genetic variance
Use of tumor tissue to inform germ lines.

Experiences in Pooling of Data: Opportunities for Harmonization of I4C Cohorts
Moderator: Dr. Martine Vrijheid

Experiences in Pooling of Data: Opportunities for Harmonization of I4C Cohorts
Dr. Manolis Kogevinas

Ad Hoc and Wider Collaborations. Most collaborations between mother-child cohorts are ad hoc initiatives focusing on a specific project with specific hypotheses and normally short duration (maximum of 5 years). For example, NewGeneris is a study of genotoxicity and biomarkers in European birth cohorts and biobanks. Results showed that micronuclei are biomarkers predictive of future cancer risk. Two initiatives—ENRIECO and CHICOS—are coordinating European birth cohorts and also conducting several case studies. ENRIECO is focusing on environmental health. CHICOS is focusing on all exposures and policy. ENRIECO is developing an inventory of European cohorts with environmental exposure data. These initiatives can be expanded worldwide by harmonizing data collection and analysis, sharing experience and expertise, coordinating with the I4C platform, and securing funding.

New Cohorts. New cohorts need to harmonize without restricting new ideas. They are necessary for I4C objectives by increasing exposure and protocol variation. New cohorts require extremely positive commitment for long-term funding and evaluation of environmental exposures. Development of new protocols and questionnaires should be coordinated with older cohorts, but innovation should be encouraged. Basic exposure and outcome variables and biological materials should be combinable. Questionnaires must be adapted to the cultures of the study populations. Because mother-child cohorts are complex projects, there is an absolute need for a multidisciplinary group to lead each cohort, including a very strong group in epidemiology.

Data Pooling versus Meta-Analyses. Some analyses can be done through meta-analysis (that is, without transfer of raw data). Depending on hypotheses and adjustment factors, individual cohorts may need to run hundreds of prespecified models that should then be meta-analyzed. Some analyses are impossible to do through meta-analysis (for example, gene-environment interactions or pathway analyses). Ethical issues must be considered for data transfer. There are two approaches to detect gene-environment interactions. A one-stage approach is a global genome-wide analysis. A two-stage approach to identify genes involved in gene-environment uses (1) screening tests to find SNPs most likely to be involved by testing for gene-environment associations and (2) case-control analyses. For association tests, logistic regressions may be performed when information on the exposure is available in both cases and controls. One of the practical issues for meta-analyses is that all studies need to use the same protocol. Data quality and comparability make pooling challenging.
Questions and Discussion

The following topic was discussed:
- Logistical issues of conducting meta-analyses and analyzing pooled data.

Consortia and Pooling of Data: Opportunities and Obstacles

*Ann Olsson, IARC, France*

IARC’s mission is to (1) coordinate and conduct research on the causes of human cancer and the mechanisms of carcinogenesis and (2) develop scientific strategies for cancer prevention and control. The agency is involved in both epidemiological and laboratory research and disseminates scientific information through publications, meetings, courses, and fellowships. Currently, nine collaborative research programs are coordinated by or have been initiated at IARC. Many include data pooling. SYNERGY, EPI-CT, and AGRICOH are three examples.

SYNERGY is a study of the joint effects of five occupational exposures and smoking in the development of lung cancer. EPI-CT was launched to (1) assess possible risks of leukemia and other cancers related to computerized tomography (CT) doses in the cohort of pediatric patients and (2) inform optimization strategies for CT use in pediatric health care. The aim of AGRICOH—an international consortium of agricultural cohort studies—is to study disease-exposure associations that individual cohorts do not have sufficient statistical power to study.

Data pooling offers opportunities but also has obstacles. It increases the power of a study and may allow studying effects in different subpopulations. Data pooling allows studying heterogeneity (for example, by geographical areas, over time, and precision of data) and often allows publishing results in higher ranked journals. Obstacles are presented by differences in study designs, inclusion criteria, questionnaires, protocol, countries, and times of data collection. In addition, cohort studies are often very different in size and are conducted independently with different purposes and outcomes. Because the studies are ongoing, data sets are dynamic, and pooled analyses must comply with national agendas. The challenges to data pooling are to create common variables for all studies by recoding, harmonizing (for example, different classifications), and compromising (for example, converting categories to average numbers). Cohort studies represent very large efforts that involve many stakeholders, and entities that approve access to data need to be identified. Requirements for optimal data pooling include questionnaires, a “code book,” a designated contact person, quality control, and addressing concerns of data owners.

Questions and Discussion

The following topics and issues were discussed:
- Use of proprietary software and database management
- Communication among PIs (for example, annual meetings, e-mails, and conference calls)
- Quality control of data analysis and programming errors
- Benefits of contributing studies’ involvement
- Providing protocols before beginning data analyses.
Opportunity for International Cooperation in the Next Generation of Birth Cohort Studies of Child Health and the Environment
Ruth Etzel, M.D., Ph.D., WHO, Switzerland

According to the WHO, about 24 percent of disease worldwide could be prevented by modifying the environment. To help prevent disease, there is a need to know about emerging environmental issues. New or “re-emerging” potential threats to children’s health and development include nanoparticles, ozone depletion, radiation, persistent organic pollutants, and endocrine disruption.

To address child health issues, the WHO has been promoting longitudinal child health cohort studies. Between 2003 and 2009, the WHO convened six consultations to promote a guide to undertaking a birth cohort study. This guide was developed by Professor Jean Golding of the University of Bristol in the United Kingdom.

Currently, very large birth cohort studies are being initiated. Similar to children’s “windows” of vulnerability to environmental insult, these large studies also have windows of vulnerability. During these windows, exposure of study investigators to positive and negative influences of other study investigators can profoundly shape the next generation of birth cohort studies. Positive influences include sharing of methods, data, and core protocols. The solid groundwork for the next generation of birth cohort studies has been laid by studies such the I4C, ENRIECO, CHICOS, the Consortium to Perform Human Biomonitoring on a European Scale (COPHES), and the Demonstration of a Study to Coordinate and Perform Human Biomonitoring on a European Scale (DEMOCOPHES).

The WHO’s interests are to promote coordination of the next generation of large birth cohort studies, develop core measures for collection across multiple large new birth cohorts, and provide a basis for further international collaborations. The WHO contends that childhood death and disability from environmental causes are preventable. However, much is still not understood about the effects of chemicals in the environment on child health and development. The WHO promotes translation the findings of birth cohort studies into public health policy to protect children of the next generation.

Questions and Discussion

The following topics and issues were discussed:
- Promoting birth cohort studies in developing countries
- Lack of resources and infrastructure for birth cohort studies in developing countries
- A WHO priority statement regarding child health and environment studies
- Advantages of large studies compared with small and medium-size studies
- The need to collect more environmental data.

Childhood Leukemia International Consortium (CLIC)
Catherine Metayer, M.D., Ph.D., University of California, Berkeley, USA

CLIC’s mission is to (1) develop and support collaborations among scientists involved in childhood leukemia research to accelerate knowledge on factors that influence the risk of
childhood leukemia through epidemiological studies and related research and (2) encourage free exchange of results (published or unpublished) and ideas in a collegial environment without fear of competition.

CLIC’s outstanding research questions are as follows:
- What are the critical windows of exposure?
- What are the risk factors of rarer childhood leukemia subtypes such as AML, acute progranulocytic leukemia (APL), and T-cell ALL?
- Are the risk factors specific to childhood leukemia cytogenetic and molecular subtypes of ALL and AML, or other tumor characteristics?
- What are the contributions of gene-environment interactions to childhood leukemia?

Currently, there are 22 CLIC studies in 13 countries. The studies have about 10,000 leukemia cases and 15,000 controls. About half of the studies have biospecimens. The study period is from the mid-1970s to 2011.

Ongoing pooled analyses include (1) fetal growth and the risk of childhood ALL; (2) vitamin/folate supplementation before and during pregnancy, MTHFR gene variants, and the risk of childhood ALL and AML; and (3) parental smoking, NQ01 gene, and the risk of childhood ALL.

Proposed pooled analyses include (1) markers of childhood infections and allergies and of ALL; (2) geographical distribution of AML cytogenetic type, including APL; and (3) indoor sources of benzene and other hydrocarbons (for example, parental smoking, home use of paints and solvents, and type of heating), NQ01 gene, and the risk of AML.

Proposals under development include (1) residential use of pesticides and the risk of childhood ALL and AML; (2) residential use of paints and solvents and the risk of childhood ALL; and (3) SES, environmental and genetic determinants, and childhood leukemia survival.

The challenges for CLIC include (1) study specificity regarding prevalence of exposure, design, and confounders (for example, SES and ethnicity); (2) meta-data analyses versus pooling of raw data; and (3) recognition of potential for selection bias in observational studies that require subject consent and attempt to assess the representativeness of participating families.

There are several opportunities of collaborations between CLIC and the I4C. For example, they could exchange ideas on genetic and epigenetic studies, exposure assessments in cohort studies that could inform case-control studies, and ethical issues. CLIC and the I4C could also exchange ideas on procedures such as standardizing data and compiling background information on exposure patterns by country and over time.

**Initiatives from the National Children’s Study**
*Michael Dellarco, Dr.P.H., NICHD, NIH, USA*

The National Children’s Study is an integrated system of activities to examine the relationships between environmental exposures and genetics on growth, development, and health. It is data-
driven, evidence-based, and community and participant informed. Environment is broadly defined to include factors such as air, water, soil, dust, noise, diet, social and cultural setting, access to health care, SES, and learning. The National Children’s Study will examine the multiple effects of environmental influences on the health and development of 100,000 children across the United States, following them from before birth until age 21, by providing high quality data to analyze scientific hypotheses. The overall goal of the National Children’s Study is to improve the health and well-being of children. There are two related phases of the National Children’s Study: the Vanguard Study and the Main Study. The Vanguard Study began prior to the Main Study and both studies will run in parallel.

Examples of exposure areas of interest include:
- Industrial chemicals and byproducts in the air, water, soil, and commercial products
- Natural products in the air, water, soil, and commercial products
- Pharmaceuticals used for therapy and in the environment
- Radiation exposure
- Proximity to manufacturing, transportation, and processing facilities
- Living with animals, insects, and plants
- Media and electronic device exposure; noise
- Access to routine and specialty health care
- Learning opportunities that are structured and unstructured
- Diet and exercise
- Family and social network dynamics in cultural and geographic context.

Examples of outcome areas of interest include:
- Interpersonal relationships and bonding
- Inflammatory processes including allergies, asthma, and infections
- Genetic and epigenetic status
- Epilepsy and other neurological disorders
- Cardiovascular screening and function
- Childhood cancer
- Multidisciplinary, multidimensional aspects of sensory input, learning, and behavior
- Precursors and early signs of chronic diseases such as obesity, asthma, hypertension, and diabetes.

The National Children’s Study Vanguard Study—the pilot phase for methods—is designed to evaluate feasibility (technical performance), acceptability (impact on participants, study personnel, and infrastructure), and cost (personnel, time, effort, and money) of recruitment, logistics, operations, visits, and visit assessments. The Vanguard Study will also evaluate recruitment strategies, recruitment parameters, retention, and cost. Assessments will include anthropomorphic measures, biospecimens, environmental measures, physical measures, questionnaires, sampling and analysis, sample collection procedures, and community environmental characterization. Using a decentralized approach, the National Children’s Study will implement new models for informatics and communication.

Formative research will (1) address specific technical questions that can be evaluated in terms of feasibility, acceptability, and cost; and (2) explore new and potentially cost-effective approaches
in areas such as genetics, cognition, and environment that have not been previously evaluated from an operational perspective. Formative research topic areas include:

- Real-time or near-time analysis of study samples, specimens, and measurements
- Study logistical analysis and improvements
- Biospecimen collection and processing
- Environmental sample collection and processing
- Physical measures
- Questionnaire development and validation
- Infrastructure development.

The National Children’s Study will contribute to children’s well-being by (1) identifying environmental factors that cause or contribute to health, development, and behavior problems; (2) understanding the biology and genetics of health, development, and behavior; (3) developing evidence-based information on which to base decisions about practice and policy regarding children’s physical and mental health; and (4) providing economic benefits and resources for future research.

Questions and Discussion

The following topics and issues were discussed:

- Public availability of the results of formative research projects
- Review and evaluation of new methodologies
- Public availability of the Main Study protocol (the Vanguard Study protocol will not be publicly available)
- Availability of the Vanguard Study data and specimens for ancillary studies
- Formal request process for the Vanguard Study data and specimens (data on exposure-outcome relationships will not be publicly available)
- Annual funding, reporting, and justification for continuation
- Availability of the Main Study questionnaires
- Need of the protocol for adjunct studies
- Commitment to data sharing.

Initiating the Testing of New Hypotheses in the I4C

_Discussion leaders: Drs. Dwyer and Linet_

The process of initiating the testing of new hypotheses generally begins by submitting a study proposal to the I4C Steering Committee. The hypothesis may be initiated by an individual associated with a birth cohort. A working group may be formed to collaborate on developing the proposal before submitting it. The proposal needs approval from the Steering Committee and the participating cohorts. Once approved, the working group requests the data or specific variables required for the study. The study investigators must agree to use the data for only the purpose described in their proposal and by only those investigators. The I4C International Data Coordinating Centre (IDCC) may perform data analyses for the study, with the investigators writing the paper. The I4C IDCC will provide input to ensure that all papers describe the variables consistently and that analyses on the same variables provide the same estimates.
Emerging cohorts are encouraged to develop hypotheses using existing I4C data. These cohorts can then develop hypotheses using their own data or integrating it with I4C data. Individuals with ideas for new hypotheses or areas for possible study can write short descriptions of them and post them on the I4C portal. Individuals from emerging cohorts are encouraged to get involved with I4C working groups.

**Funding Opportunities for Investigator-initiated Grants with Foreign Components at the NIH**

*Somdat Mahabir, Ph.D., M.P.H., NCI, NIH, USA*

The NIH invests more than $30 billion annually in medical research for the American people. The NIH supports external research through three mechanisms:

- Grants—The investigator decides the research design and approach.
- Contracts—The government decides the research scope to fill a specific need and establishes detailed requirements.
- Cooperative agreements—These are similar to grants but both awarding Institute/Center (I/C) and recipient have substantial involvement in design and conduct of the research.

The goal of these mechanisms is to provide an organized approach to obtain NIH funding. Grants can be used to fund foreign population-based studies such as the I4C, CLIC, ENRIECO, and CHICOS.

Grant applications require completed application forms and registration, which are mostly submitted electronically. These forms and required registrations are available on a number of Web sites. The NIH’s Research Portfolio Online Reporting Tools (RePORT) provides data on what NIH spends each year by I/C and by disease type. RePORT also provides information on scientific concepts, emerging trends, and success rates. It shows research that has already been funded and lists investigators in the funded research areas. This information can be used for planning and developing strategies for new grant applications. Applicants must use existing Funding Opportunity Announcements and mechanisms (for example, R01, R21, and R03), and identify dollar limits, submission dates, Awaiting Receipt of Application, and eligibility. Foreign applicants should decide whether to apply directly as a foreign institution or as a subcontractor with a U.S. institution. Foreign applicants should be aware of U.S. investigators who are conducting similar research and seek collaboration with U.S. researchers.

When considering a grant application, it is important to carefully read the *NIH Guide for Grants and Contracts*. The Guide lists searchable funding opportunities and describes the rules for writing and submitting applications. Before writing the application, investigators should (1) understand the mission of the funding agency and their guidelines, (2) check the NIH’s Center for Scientific Review (CSR) or the I/C itself, (3) review the review panelists’ interests and credentials, and (4) contact the person listed in the funding opportunity announcement. While writing the application, investigators should seek mentoring from funded investigators. Foreign applicants should justify why a foreign country should be involved and describe the benefit/relevance to the U.S. population. Applicants should test their ideas on colleagues, form a mini-panel to discuss the application, and seek collaborators. Applicants should know the group that is promoting the research and who the critical players are.
Grant applications are referred by the CSR, which assigns applications to specific areas in an I/C and assigns specific review groups. The reviewers may be some of the critical players. Investigators submitting grant applications can request specific reviewers and ask that the application not be reviewed by a particular study section.

The foreign research should have a public health impact on the U.S. population. Applicants should explain why the project presents special opportunities that either are not readily available in the United States or augment existing U.S. resources. Special opportunities may include unusual talent, resources, populations, or environmental conditions. Foreign researchers may be included when there is a specific need or expertise that is not readily available in a U.S. institution. There is generally no advantage if the lead institution is foreign. Both U.S and foreign PIs can be listed in the application.

There are five criteria in grant reviews: Significance, approach, innovation, investigators, and environment. The applications receive an impact score from 1 to 9, with 1 being exceptional/excellent and 9 being fair/poor. About half of grant applications are not reviewed. If an application is reviewed, it may not receive a good score and may not be accepted. However, some applications that do not score well are funded, and some applications that score well are not funded. If an application is not accepted, it can be resubmitted only once. When resubmitting, it is important to review each comment and address each one point by point. Resubmitted applications must be substantially revised.

**Meeting Wrap-Up, Next Steps, and Proposal for Future Meeting**

Dr. Dwyer thanked the workshop’s supporting organizations (the NCI, the NICHD, IARC, and MCRI), the WHO, the I4C Steering Committee, and CREAL for hosting the workshop. In particular, he thanked Drs. Linet and Tikellis. The next steps include developing new hypotheses, collecting new data, and accepting new cohorts into the I4C. Several locations were proposed for the future I4C workshop, including Berkeley, California; London or some other European location; East Asia; and Australia.

**Participants**

Therese W. Bakke, M.Sc.Pharm., Norwegian Institute of Public Health, Norway  
Tone Bjørge, M.D., Ph.D., Norwegian Institute of Public Health, Norway  
Magda Bosch de Basa Gómez, CREAL, Spain  
Patricia Buffler, Ph.D., M.P.H., University of California, Berkeley, USA  
Michael J. Dellarco, Dr.P.H., NICHD, NIH, USA  
Terry Dwyer, M.D., Ph.D., MCRI, Australia  
Manel Esteller, Ph.D., IDIBELL, Spain  
Ruth A. Etzel, M.D., Ph.D., WHO, Switzerland  
Hideki Hasunuma. M.E., Center for Environmental Information Science, Japan  
Maureen Hatch, Ph.D., NCI, NIH, USA  
Zdenko Herceg, Ph.D., IARC, France  
Hector Hernandez-Vargas, IARC, Spain