**Title**
Study of factors related to childhood leukemia and to central nervous system tumors in California



**Author**
Oksuzyan, Sona

**Publication Date**
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# *Study of factors related to childhood leukemia and to central nervous system tumors in California*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Epidemiology

by

Sona Oksuzyan, MD, MPH

2013

## Executive summary of the dissertation

*Study of factors related to childhood leukemia and*
*to central nervous system tumors in California*

by

Sona Oksuzyan
Doctor of Philosophy in Epidemiology
University of California, Los Angeles, 2013
Professor Leeka Kheifets, Chair

Childhood leukemia is the most common malignant disease in children, followed by childhood CNS/brain tumors. [1, 2] Both diseases have been recognized for a long time but little is known about the etiology of childhood leukemia and primary CNS/brain tumors.

The aims of this project were to evaluate: 1) the association of several perinatal factors with childhood leukemia as well as with CNS/brain tumors; 2) the relationships between childhood leukemia and child's and parental race, Hispanic ethnicity and socio-economic status (SES) at individual and census levels; 3) factors associated with residential mobility in childhood leukemia cases.

We conducted a large registry-based study in California using California birth and cancer registries to obtain information on childhood leukemia and CNS/brain tumor cases and controls. Information on case's diagnosis, histological subtypes, age at diagnosis, sex, and diagnosis address was obtained from California cancer registry; information on perinatal factors (birth weight, gestational age, birth order, parental age at birth, maternal complications during pregnancy, abnormal condition of a newborn), socio-demographic factors (date of birth, sex, race and ethnicity of parents and child, parental education, sources of payment for delivery, birth address) were obtained from the birth registry.

We linked California cancer and birth registries to obtain information on 5788 cases of childhood leukemia and 3308 cases of CNS/brain tumors and their 5788 and 3308 controls matched on age and sex (1:1). To address at least partially the problem with misclassification we

categorized birth weight, gestational age, parental age, child and parental race and Hispanic ethnicity, individual level proxies for and census-based SES in several ways. For all analyses we used conditional logistic regression, with adjustment for potential confounders.

Our results for childhood leukemia and perinatal factors indicate that high birth weight and LGA were associated with increased risk and SGA with decreased risk of total childhood leukemia and ALL, being first-born was associated with decreased risk of AML, and advanced paternal age was associated with increased risk of ALL. Our results for CNS/brain tumors and perinatal factors suggest that maternal genital herpes, blood and immunological disorders during pregnancy and newborn CNS abnormalities were associated with increased risk of CNS tumors. Maternal infections during pregnancy were associated with decreased risk of CNS tumors. Factors associated with childhood leukemia and CNS tumors varied by subtype, an indicator of different etiology for different subtypes.

We noted that Black children had decreased risk (ref. - White) and Hispanic children (ref. non-Hispanic) had increased risk of total childhood leukemia and ALL. Asian race was associated with increased risk of AML. These differences in the incidence of childhood leukemia indicate that some genetic and/or environmental/cultural (e.g., diet or lifestyle) factors are involved in etiology of childhood leukemia.

We found no evidence to support the suggestion that SES, as measured by variety of proxies is a determinant of childhood leukemia and of its both subtypes. It is likely that results of many previous studies that found an association between childhood leukemia and SES were largely influenced by selection or ecological bias.

The results of this study indicate that childhood leukemia cases in California are residentially very mobile, with 58% of them moving between birth and diagnosis. The residential mobility of childhood leukemia cases notably varied by time interval between birth and diagnosis, child's race/ethnicity, maternal age at birth, census-based SES, and the source of payment for delivery; it varied less by the distance to the nearest power line. Our results suggest that even if information on the residential mobility of subjects is unavailable, it might be possible and important to look at the distribution of factors that could be associated with residential mobility.

The dissertation of Sona Oksuzyan is approved.

Beate Ritz

Jorn Olsen

Abdelmonem Afifi

Catherine M. Crespi

Leeka Kheifets, Committee Chair

University of California

Los Angeles

2013

## Dedication

*This work is dedicated to my baby son Arno who was the biggest stimulus for me to finish the dissertation.*

*I want to thank my husband, my parents and friends without whom I would not be able to finish this work in the specified timeframe.*

## Table of Contents

**Acknowledgements**

I want to thank Dr. Leeka Kheifets for her professional, psychological, emotional, and financial support and for her patience throughout the years.

I also want to thank Dr. Crespi for her professional support, great advices and recommendations for analyses.

Special thanks to California Department of Public Health and California Cancer Registry for providing access to cancer and birth regirstries.

## Biographical sketch

| NAME | POSITION TITLE |
|---|---|
| Oksuzyan, Sona | PhD candidate |

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

| INSTITUTION AND LOCATION | DEGREE *(if applicable)* | YEAR(s) | FIELD OF STUDY |
|---|---|---|---|
| Medical State University of Armenia | M.D. | 1989-1995 | General Medicine |
| Medical State University of Armenia, residency | M.D. | 1995-1999 | Ob & Gyn |
| Emory University, Atlanta | M.P.H | 2003-2005 | Global Health |
| University of California, Los Angeles | PhD (c) | 2007-curr | Epidemiology |

### A. Positions and Honors.

**Positions and Employment**

| | |
|---|---|
| June-Sep 2012 | Intern at Amgen, Inc., Thousand Oaks, CA |
| 2008-2012 | Graduate Student Researcher, Department of Epidemiology, UCLA |
| 2008-2012 | Teacher Assistant for various courses in Public Health |
| 2005- 2006 | Consultant, American International Health Alliance, Washington D.C. |
| 1999-2003 | Obstetrician-gynecologist, Maternity Clinical Hospital #4, Yerevan, Armenia |
| 2002-2003 | Clinical Advisor, Intra Health International/Prime II project, Yerevan, Armenia |
| 2001-2002 | Monitoring of Distribution of Supplies Program, United Methodist Committee of Relief (UMCOR), Yerevan, Armenia |
| 1997-1998 | Consultant in Obstetrics and Gynecology, Instituto per la Cooperazione Universitaria, Yerevan, Armenia |

**Fellowships and trainings**

| | |
|---|---|
| Nov – Dec 2002 | Web-Learning course on Basic of Health Planning. American University of Armenia, Yerevan (honor certificate) |
| Apr – May 1999 | Advanced International Training Programme on "Reproductive Health Care", The University of Karlskrona, Sweden (certificate & diploma) |
| Sep – Oct 1998 | International Postgraduate Basic training Course on "Reproductive Health/Family planning", University Medical School of Debrecen, Hungary (certificate) |
| May - Nov 1997 | Fellowship in cervical cancer screening at the Institute of Obstetrics and Gynecology, The Catholic University of Rome (certificate) |

Aug 1997 Interregional Theoretical-Practical Course of The Micro-colpohysteroscopy, Vibo Valentia, Italy (certificate)

**B. Publications (in reverse chronological order).**

- Kheifets L, Crespi CM, Hooper C, **Oksuzyan S**, Cockburn M, Ly T, Mezei G. *Epidemiologic study of residential proximity to transmission lines and childhood cancer in California: Description of design, epidemiologic methods and study population.* Journal Of Exposure Science And Environmental Epidemiology: accepted for publication, 2013
- **Oksuzyan S**, Crespi CM, Cockburn M, Mezei G, Kheifets L. *Birth weight and other perinatal characteristics and childhood CNS tumors: a case-control study in California.* Cancer Epidemiol: in print.
- Vergara, X., Kheifets, L., Greenland, S., **Oksuzyan, S.**, Cho, Y. S., Mezei, G. *Occupational exposure to extremely low-frequency magnetic fields and neurodegenerative disease: a meta-analysis.* J Occup Environ Med, 55(2):135-46, Feb 2013
- **Oksuzyan S**, Crespi CM, Cockburn M, Mezei G, Kheifets L. *Birth weight and other perinatal characteristics and childhood leukemia in California.* Cancer Epidemiol: 36(6):e359-65, Dec 2012
- Kheifets L, Ahlbom A, Crespi CM, Feychting M, Johansen C, Monroe J, Murphy MFG, **Oksuzyan S**, Preston-Martin S, Roman E, Saito T, Savitz D, Schüz J, Simpson J, Swanson J, Tynes T, Verkasalo P, Mezei G. *A Pooled Analysis of Magnetic Fields and Childhood Brain Tumors.* Am J Epidemiol. 172(7):752-61, 2010 Oct 1
- Kheifets L, Ahlbom A, Crespi CM, Draper G, Hagihara J, Lowenthal RM, Mezei G, **Oksuzyan S**, Schüz J, Swanson J, Tittarelli A, Vinceti M, Wunsch Filho V. *Pooled Analysis of Recent Studies on Magnetic Fields and Childhood Leukemia.* Br J Cancer. 103(7):1128-35, 2010 Sep 28
- Kheifets L, **Oksuzyan S**. *Exposure Assessment And Other Challenges In Non-Ionising Radiation Studies Of Childhood Leukaemia.* Radiation Protection Dosimetry (Oct 2008), pp. 1–9
- **Oksuzyan S**. *Women's Health Curriculum: 9 modules for family physicians in Turkmenistan.* (English and Russian). 2006.
- **Oksuzyan S**. *The Special Gynecological Needs of Women Living with HIV/AIDS*. CommonHealth Winter 2004. Available at: http://www.aiha.com/resources/Doc/oksuzyan.pdf
- **Oksuzyan S**, MD, Rebecca Kohler, MPH, and Leah Levin, MHS. *Overcoming obstacles, increasing access and quality: A plan for family-friendly reproductive health care in Armenia*. Abstract on the 131st Annual Meeting of APHA (November 15-19, 2003)
- Engelbrecht S, Kiplinger N, L. Hurber L, **Oksuzyan S**. *9 learning modules for PHC providers in Armenia*. (English and Armenian). 2003.
- Giannice R, Margariti AP, **Oksuzyan S**. *Elements of cervico-vulva-vaginal colposcopy*. *Manual for Ob&Gyn in Armenia*. (Armenian). 1998.

# Chapter 1. Introduction

## Aims

The aims of this research were:

- to examine the association of childhood leukemia with perinatal factors, such as birth weight, gestational age, parity, birth order, history of miscarriages, maternal and paternal age, complications during pregnancy
- to examine the association of CNS/brain tumors with perinatal factors, such as birth weight, gestational age, parity, birth order, history of miscarriages, maternal and paternal age, complications during pregnancy
- to examine the association of childhood leukemia with individual- and community-level proxies for socio-economic status (SES)
- to examine the association of childhood leukemia with race and Hispanic ethnicity
- to examine factors associated with the residential mobility in childhood leukemia cases

## Background and significance

### Childhood leukemia

Childhood leukemia is the most common malignant disease in children. [1] The disease has been recognized for a long time but a little progress is made in understanding it; its determinants remain unknown.

Childhood leukemia affects lymphocytes, a type of white blood cells. Leukemic cells accumulate in the bone marrow, replace normal blood cells and spread to other organs including liver, spleen, lymph nodes, central nervous system, kidneys and gonads. [3]

Leukemia can be either fast growing (acute), or slower growing (chronic). Childhood leukemia in the vast majority of cases is acute. Chronic forms of childhood leukemia are very rare in all populations and particularly in children, accounting for ≤2–3% of all leukemia diagnoses. [3, 4] Acute leukemia has two major groups depending on cell types from which leukemia starts:

- acute lymphocytic leukemia (**ALL**, also called acute lymphoblastic leukemia) develops from lymphocytes in the bone marrow and accounts for approximately 80% of all childhood leukemia
- acute myelogenous leukemia (**AML**, also called acute myeloid leukemia) develops from other types of blood cells and accounts for less than 20% of all leukemia. [5, 6]

*Incidence and etiology*

Worldwide, estimated annual number of childhood cancers is approximately 200,000. Childhood leukemia is the most common type of cancer in children under age 15, accounting for about 30% of all childhood cancer diagnoses. [7-9] The most common form of childhood leukemia is acute lymphocytic leukemia with the annual incidence in different countries varying from 0.9 to 4.7 per 100,000 children. The incidence is around 4 per 100,000 children with a peak at 2-5 years of age in developed countries. AML incidence is less than one-quarter of ALL incidence and it doesn't have a noticeable peak at any age in childhood. [1, 10]

Leukemia accounts for the largest number of cases of childhood cancer in the United States. Each year approximately 3,250 children and adolescents younger than 20 years of age are diagnosed with leukemia, of which 2,400 are ALL. [11] Although due to advances in treatment the 5-year survival rate for ALL in children is currently more than 80% and for AML it is more than 50%, leukemia remains the primary cause of cancer mortality of children in the United States and accounts approximately for 25.5% of all cancer deaths in children. [12] According to California Cancer registry report, the incidence of childhood leukemia in California was between 4 and 5 with about 5.2 and 4.3 per 100,000 for males and females, respectively. Incidence was highest among the youngest age group (0-4 years), 8.23 and 6.80 cases per 100,000 for males and females, respectively, then it was declining with age until late adolescence. [13]

The etiology of childhood leukemia remains mainly unknown. There are some theories about genetic, viral and environmental factors that may be responsible for the development of this disease. Several genetic conditions associated with chromosomal instability, for example, Fanconi’s anemia, Bloom’s syndrome and ataxia-telangiectasia, are linked to an increased incidence of childhood leukemia. Down syndrome was found to be associated with 10 to 30-fold increase in childhood ALL and 100-fold and more increase in childhood AML. [8, 14, 15] These facts support a genetic component in the development of childhood leukemia. Approximately

80% of children with acute leukemia, particularly with ALL, have abnormal karyotypes. [16] According to Greaves hypothesis, hyperdiploidy or common the TEL-AML1 translocation originates in utero and seems to initiate the disease (the first hit). [17] Though, not all translocations will lead to conversion of the pre-leukemic clone to overt disease. Penetrance of such translocations is quite low. For the development of childhood leukemia postnatal genetic alterations, i.e. somatic mutations are needed (the second hit). [17]

Some researchers believe that a viral infection might be responsible for the second hit caused by abnormal immune responses to this infection. [18-22] In fact, many common viral and bacterial infections occur during early childhood which is also a peak age for childhood leukemia. Some viruses are detected in certain types of leukemia and lymphoma, such as human T-cell lymphotropic virus type 1 (HTLV-1) in adults with T-cell leukemia/lymphoma and EBV in children with Burkitt lymphoma [17, 23, 24], but no virus is yet detected for childhood leukemia.

Kinlen noticed that clusters of childhood leukemia, particularly ALL, occurred in children whose families had moved and mixed in a new setting. Kinlen proposed a "population mixing" hypothesis on the origins of leukemia and posited a specific unidentified viral infection as potentially causative in the leukemia "outbreaks" which occurred soon after mixing.[25] In one Canadian study Koushik et al. found that incidence of childhood leukemia was higher in rural areas where population had grown significantly, while it was lower in slowly growing urban areas. [26]

Greaves suggested that late exposure to a common infection could lead to an abnormal immune response and play an important role in the development of childhood leukemia. [20, 21] Studies that found that daycare attendance (as compared to home care) and increasing birth order had a protective effect against total childhood leukemia and ALL due to an early exposure to infections and support Greaves' hypothesis. [27-29]

Several maternal and perinatal characteristics were linked to childhood leukemia. The most consistent association was found between high birth weight (> 3,500 or 4,000g) and acute lymphoblastic leukemia [15, 30-32]; less consistent - with acute myeloid leukemia [15, 33]. Few researchers looked at birth weight at gestational age and found that large for gestational age (LGA) children had an increased risk of childhood leukemia [33, 34]. Other characteristics that are not consistently linked to childhood leukemia are gestational age, birth order, maternal

history of fetal loss prior the index child, maternal age > 35 at pregnancy, some pregnancy complications (hyperemesis, hydatiform mole, polyhydramnios, preeclampsia) [9, 32, 35, 36].

Some socio-demographic factors also deserve attention for their relationships with childhood leukemia since leukemia incidence varies substantially by age, sex, and race/ethnicity. Gender is one of the few factors that are consistently linked to childhood leukemia. In particular, male gender is found to be associated with about 30% higher incidence of childhood leukemia compared with female gender [11, 37-39].

Although for many other health outcomes low socio-economic status (SES) was associated with higher incidence, for childhood leukemia many studies found the association of low SES with lower incidence of childhood leukemia with 8 to 66% decrease in risk [40-42].

Ethnic differences in the incidence of childhood leukemia are of particular interest in California in the light of its diverse ethnic/racial population. In many studies White race was found to be associated with an increased risk of childhood leukemia [34, 37, 39, 43]. In California (1988-1999) Hispanics had the highest age-adjusted rates (5.62 cases per 100,000), followed by non-Hispanic whites (4.46 cases per 100,000). The lowest rates reported among non-Hispanic blacks (2.91 cases per 100,000) [13]. Interestingly, the incidence in several Latin American countries is the highest in the world: 5.65 cases per 100,000 in Costa Rica, 5.54 per 100,000 in Ecuador, 4.43 in Uruguay, 41.7 in Colombia, which may suggest genetic or environmental exposures that affect childhood leukemia risk in Hispanic population [44].

Many researchers believe that environmental factors may play a role in the development of childhood leukemia (as a second hit). Ionizing radiation at high doses is one of few exposures for which the causal relationship with childhood leukemia has been established. [9, 45, 46] The other most common environmental factor linked to childhood leukemia is exposure to extremely-low frequency magnetic fields (EMF) at relatively high exposure levels (above 0.3 μT) with about 2-fold increase. [47-51] The other factors that were considered as risk factors for childhood leukemia are some toxic chemicals (hydrocarbons, pesticides) and antineoplastic drugs. [8, 9, 52]

Many epidemiological studies examining the association between various environmental exposures (EMF, pesticides, traffic density and other ) or SES and childhood leukemia have used a residential address of subjects to evaluated risks associated with proximity of a child's residence to sources of EMF [50, 53-55], air pollution [56], traffic emissions [57], and

agricultural pesticides [58] and/or to assign community-based measures of SES. The majority of these studies used a single residential address of a child (birth, diagnosis, longest lived) and did not account for the residential mobility of subjects, which may result in misclassification of exposure and selection bias.

Leukemia sub-types, particularly ALL and AML, pose challenges for epidemiologic research because they may have different risk factors. For example, exposure to specific chemotherapy agents (e.g. alkylating agents) was found to be associated with an increased risk of childhood AML, in contrast to the rarity of treatment-related ALL [11]. Moreover, in few recent studies associations with factors such as benzene and pesticides suggested that childhood AML could share risk factors with adult AML [11]. ALL shows very distinctive age-distribution pattern, with an incidence peak at 2-5 years of age. The peak at 2-5 years of age is more apparent for White children compared to Black children. By contrast with ALL, the age-distribution pattern for AML shows highest rates in the first two years of life, thereafter the incidence is decreasing until 10 years of age followed by increasing rates thereafter. In contrast to ALL, the incidence of AML in Black children and White children is similar [11]. Table 1 below summarizes risk factors for AML and ALL at the state of current knowledge.

**Table 1.** Current knowledge on risk factors for AML and for ALL [11]

| | | AML | ALL |
|---|---|---|---|
| **Probable risk factors** | Sex | | ≈ 30% higher risk in boys then in girls [11, 37, 39] |
| | Age | | Peak in the incidence between the ages 2 and 5 [11, 37, 39] |
| | Race | The highest incidence rates are reported for Hispanic children [11, 37, 39] | ≈ 2-fold higher risk in white children compared to black children [11, 37, 39] |
| | SES | Inconsistent findings (39, 41, 43) | Positive association [11, 37, 39, 59] |
| | Ionizing radiation *in utero* and postnatal | Positive association (1.5-fold) [11, 37, 39] | Positive association (1.5-fold increase) [11, 37, 39] |
| | Chemotherapeutic agents | Positive association with alkylating agents or epipodophyllotoxins [11, 37, 39] | |
| | Down syndrome (DS) and other genetic conditions | Very strong positive (up to 500-fold increased risk for DS) [11, 37, 39, 60] | Strong positive association (up to 30-fold increased risk for DS) [11, 37, 39, 60] |
| **Factors for which evidence is suggestive but inconclusive** | Birth weight | Positive association; inconsistent findings [33, 36, 61] | Positive association (≈ 2-fold increased risk) [11, 14, 15, 31-34, 36-39, 61-63] |
| | Maternal history of fetal loss | Inconsistent findings [36] | Positive association; inconclusive findings [11, 35-37, 39] |
| | Maternal age (> 35) | | Positive association; inconclusive findings [11, 37, 39] |
| | Paternal age (> 40) | | Positive association; inconclusive findings (86) |
| | Birth order or parity | Inverse association or no association [15, 30, 33, 34, 36, 64] | Inverse association; inconclusive [11, 37, 39] |
| | Vitamin use | Inverse association, inconclusive [65, 66] | Inverse association [65-69] |
| **Factors for which evidence is inconsistent or limited** | Magnetic fields | Inconsistent findings [11, 37, 39, 49, 50, 54] | Positive association (≈ 2-fold increased risk) [11, 37, 39, 49, 50, 54] , inconclusive |
| | Hyperemesis | | Inconsistent findings or positive association [36] |
| | Polyhydramnios | Inconsistent findings; positive association [36] | |
| | Anemia | Inconsistent findings; positive association [36] | |
| | Hydatiform mole | | Inconsistent findings; positive association [36] |
| | Smoking prior to and during pregnancy | | Inconsistent findings; positive association [11, 37, 39] |
| | Parental occupational exposures | | Inconsistent findings or weak positive association [11, 37, 39] |
| | Postnatal infections | | Inconsistent findings; inverse association (101) |
| | Diet | | Inconsistent findings; positive association with meat consumption [70, 71] |

| | | | |
|---|---|---|---|
| | Air pollution/ traffic density | Inconsistent findings [56, 72, 73] | Inconsistent findings; weak positive association [56, 72, 73] |
| | Maternal alcohol consumption during pregnancy | Inconsistent findings; positive association (1.5-2 fold increased risk) [11, 37, 39] | |
| | Parental and child exposure to pesticides | Some positive association; inconsistent findings [11, 37, 39] | |
| | Parental exposure to benzene | Some positive association; inconsistent findings [11, 37, 39] | |
| | Maternal drugs abuse during pregnancy | One report suggested positive association | |
| | Radon | Some positive association [11, 37, 39] | |
| | Postnatal use of chloramphenicol | One study found substantial increased risks (10-fold) [74] | One study found substantial increased risk (10-fold) [74] |

**CNS and brain tumors**

Brain tumors are the second common malignancy in childhood after childhood leukemia and the most common of the solid tumors. Brain and CNS cancers make up 16.6-21% of all childhood cancers [75, 76]. Primary CNS tumors originate in the central nervous system and are classified as benign, with uncertain behavior or malignant. Malignant brain tumors in children rarely spread to other organs. Types of CNS tumors are named from the cells in which they originate. Astrocytomas/gliomas accounted for 52% of CNS malignancies, primitive neuroectodermal tumors (PNET) such as intracranial neuroblastoma, pineoblastoma, and medulloblastoma comprised 21%, other gliomas 15% and ependymomas about 9% [75].

*The incidence and etiology*

Worldwide incidence of CNS tumors varies from 1.7 to 4.1 per 100,000 with the lowest rates in Hong Kong and Costa Rica and the highest rates in Sweden [77]. About 2,820 new cases of childhood primary brain CNS tumors in children less than 15 years of age were diagnosed in the United States in 2007 [78]. In the US the incidence is around 3.2 per 100,000 [77].

The incidence rate of brain and CNS cancers in children has risen over the past three decades. From 1973 to 1994 the reported incidence increased by about 35% [79, 80]. Various factors have been suggested to account for this increase in incidence. The most probable factor is an improvement in diagnostic techniques. Despite improvements in health care and treatment the death rate has dropped slightly over this period. Brain tumors remain the leading cause of solid tumor cancer death in children [76].

Very little is known about the etiology of primary CNS and brain tumors. To date, there is no specific risk factor known to explain an occurrence of a substantial proportion of brain tumor. Ionizing radiation and some hereditary conditions have been found to be associated with increased risk for CNS cancer in children. For genetic conditions such as neurofibromatosis, tuberous sclerosis, nevoid basal cell syndrome, Turcot syndrome, Li- Fraumeni syndrome the association is very strong (e.g. 50-fold increase for neurofibromatosis and 70-fold increase for tuberous sclerosis). But together, these conditions account for less than 5% of all childhood brain tumors [2].

Sex is one of the factors that were consistently associated with childhood CNS and brain tumors. Males usually have higher incidence rate compared with females, mostly due to

pronounced differences in incidence rates for PNET and ependymomas. In the US in 1990-1995 boys had a 24% higher incidence rate compared to girls (30.0 vs. 24.2 per million respectively. White children have about 18% higher average CNS incidence rate compared with black children. The racial difference in rates for males are to some extent more pronounced compared to those for females [75, 79].

For all other factors the evidence is either suggestive/inconclusive or inconsistent/limited. Maternal meat consumption during pregnancy, history of brain tumor in parent or in sibling, family history of bone cancer, leukemia or lymphoma were reported positively associated in several studies [79]. Products containing N-nitroso compounds (beer, make-up, diuretics, antihistamines, rubber baby bottle and pacifiers), certain father's occupations and related exposures, pesticides, history of head injury, and family history of epilepsy/seizures and of mental retardation were found to be positively associated with brain tumors but the evidence is very inconsistent and based on a limited number of studies [75, 79]. While some early studies suggested an association with magnetic fields, latter studies and a pooled analysis did not confirm this association. [81]

Childhood CNS/brain tumors were included in the analysis on perinatal factors; but childhood leukemia remained the primary focus of the project.

**Knowledge gaps and contribution of this project**

Many studies on the association of childhood leukemia and perinatal factors such as birth weight, gestational age, birth order, parental age used single cut point for aforementioned factors and/or were interview- or questionnaire-based; therefore, these studies were subjects to misclassification and recall biases.

To our understanding, there was no previous research done using several cut points and moving windows analysis for birth weight and its association with childhood leukemia or CNS/tumors and their subtypes. Limited research was done on birth weight, birth order and parental age as risk factors for childhood CNS tumors. Practically, no research was done on relationships of maternal complications during pregnancy, newborn abnormal conditions and childhood CNS tumors, particularly for main subtypes.

Although many studies on childhood leukemia have looked at SES, but majority looked at it as confounder. Most of these studies used either individual level or community-based proxies for SES, while our study used both level proxies.

Limited number of studies specifically examined race and/or ethnicity in relation to childhood leukemia risk. [43, 82, 83] Most studies of childhood leukemia have considered race and/or ethnicity only as a covariate or a confounder in their analyses of childhood leukemia. The definition of race and ethnicity differed in all these studies. Even fewer studies have looked at relationships between race and ethnicity and the risk of major subtypes of childhood leukemia, acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML).

Most of research on abovementioned factors were limited in sample size, could be affected by selection bias and did not address problem of missing data.

Our data were based on population registries with almost complete registration of births and cancers in California and controls were randomly selected from birth registry. Since these registries are independent of each other and participation of subjects was not required for our data collection, selection bias was unlikely in this study. Misclassification of outcome was also unlikely in this study due to the completeness and high accuracy of the California Cancer Registry. Misclassification of birth weight, gestational age, birth order was possible but unlikely since these variables are usually recorded fairly accurately. In addition, we used different cut points for many variables to avoid dependence on an arbitrary cut point. We also adjusted for potential confounders that were available in registries. In order to address the problem of missing data, a major limitation in many studies, we multiply imputed dataset and repeated analyses using multiply imputed dataset and compared results with a complete-case analysis.

The major contribution of the current work is fourfold: First, we used a large dataset that allowed for major subtype analysis. Second, we reduced some sources of bias by using registry-based data and various cut points for variables of interest. Third, we repeated all analysis using multiply imputed dataset, thereby addressing an issue of missing data. Fourth, for several factors no previous research was done.

# Chapter 2

## Study design and methods

Two data sources were used for this study, the California Cancer Registry (CCR) and the California Birth Registry (CBR). California's population-based cancer registry was used to obtain information on all childhood leukemia and CNS/brain tumor cases diagnosed between 1988 and 2007 in children younger than 16 years who were born in California and who resided in California at the time of diagnosis. [84] The CCR provided information on child's age, sex, residence at the time of diagnosis and cancer diagnosis, subtypes and characteristics.

Controls, i.e. children who were 0-15 years old at the time of "pseudo-diagnosis" (i.e. at the time of diagnosis of a matched case), survived by age 15 without being diagnosed with any type of cancer, born in the same time period as cases in California, were selected randomly from the CBR. Controls were matched to cases (1 to 1) on the basis of date of birth and sex.

It is warranted to explain how subjects with Down syndrome were treated during sampling and analysis since Down syndrome is very strongly associated with childhood leukemia [11, 37, 39, 60, 62] and is considered one of the strongest potential confounders for the associations of childhood leukemia with perinatal and socio-demographic factors. Many researchers chose to exclude children with this syndrome from their analysis. The number of children with Down's syndrome included in our study was quite low (n=48). We have conducted analyses both with and without subjects with Down's syndrome; the results were similar. Thus, we decided to leave subjects with Down's syndrome in our sample.

Birth certificate data recently have become available statewide in an electronic format. Birth registry provided information for cases and controls on various factors, including the birth address, date of birth, sex, race and ethnicity; gestational age, birth weight, birth order; mother's and father's age, race, ethnicity and education; maternal birthplace, parity, terminations before and after and 20 weeks of gestation, the number of children born alive, the number of children ever born, the number of children living, maternal complications during delivery, newborn abnormal conditions, and the principal source of payment for medical services. [85] A detailed birth address was available in an electronic format only for years 1998 to 2007. Prior to 1997, only the zip code of the home address at the time of birth was available electronically. Year 1997 was a transition year in which only a portion of addresses were computerized. Since the peak of

childhood leukemia was between ages 2 and 5 we expected to have fewer cases born before 1983 (cancer registry is available from 1988). However, there were fewer cases born before 1986; therefore, it was worthy to spend most of efforts and money on extracting birth certificates for years 1986 and thereafter. For those years individual birth certificates were obtained; and maternal residential address and other variables were extracted from those manually.

All addresses were geocoded by our collaborators at the University of Southern California (USC) using their Geographic Information System (GIS) Laboratory's open-source geocoder, which uses parcel level data for Los Angeles County and street level data for the whole of California. Details on the geocoding practice, reference data, and procedures can be found in previous publications.[86] For the purposes of the study on the residential mobility of childhood leukemia cases, we included in the analysis only cases with more precise geocoded match (i.e. parcel and street segment matches) for both birth and diagnosis geocoded addresses, which consisted about 78% of all childhood leukemia cases.

Despite the anticipated high number of cases, the sample size was reduced due to missing data. Checks were made to compare data on complete and incomplete cases to decide whether it may be assumed that data were missing at random. With an assumption of data missing at random (MAR) multivariate imputation techniques was used to estimate missing values. [87, 88] Then analyses were repeated on imputed datasets and compared to a complete-case analyses.

Since this was a matched case-control study, the primary analysis was conditional logistic regression.[89] Unadjusted and adjusted for various covariates ORs were computed for each part of this project, but results were presented for adjusted analyses only. Various models including different subsets of covariates were implemented, including checks for potential influential observations. [89] The final models were chosen based on information on known or potential confounders and model fit statistics; the most parsimonious models with the lowest Akaike information criterion **(**AIC) values are presented. The analysis was conducted using statistical software SAS 9.2 and SAS 9.3.[90]

My role in the project was assisting in grant and proposal writing, obtaining approvals and approval renewals for the study from the University of California Los Angeles (UCLA) Institutional Review Board (IRB) and Committee of the Protection of Human Subjects (CPHS) of California, data cleaning, data management, analysis, and manuscript writing.

## Chapter 3

# Birth weight and other perinatal characteristics and childhood leukemia in California

Oksuzyan S[1], Crespi CM[2], Cockburn M[3], Mezei G[4], Kheifets L[5]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, soksuzyan@ucla.edu; [2]Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California, ccrespi@ucla.edu; [3]Department of Preventive Medicine, University of Southern California, Los Angeles, California, myles@med.usc.edu; [4]Environment Department, Electric Power Research Institute, Palo Alto, California, gmezei@epri.com; [5]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, kheifets@ucla.edu.

**Abstract**

*Aims:* We conducted a large registry-based study in California to investigate the association of perinatal factors and childhood leukemia with analysis of two major subtypes, acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML).

*Methods:* We linked California cancer and birth registries to obtain information on 5788 cases and 5788 controls matched on age and sex (1:1). We examined the association of birth weight, gestational age, birth and pregnancy order, parental ages, and specific conditions during pregnancy and risk of total leukemia, ALL and AML using conditional logistic regression, with adjustment for potential confounders.

*Results:* The odds ratio (OR) per 1000 gram increase in birth weight was 1.11 for both total leukemia and ALL. The OR were highest for babies weighing ≥4,500g with reference <2,500g: 1.59 (95% CI: 1.05-2.40) and 1.70 (95% CI: 1.08-2.68) for total leukemia and ALL, respectively.

For AML, increase in risk was also observed but the estimate was imprecise due to small numbers. Compared to average-for-gestational age (AGA), large-for-gestational age (LGA) babies were at slightly increased risk of total childhood leukemia (OR=1.10) and both ALL and AML (OR=1.07 and OR=1.13, respectively) but estimates were imprecise. Being small-for-gestational age (SGA) was associated with reduced risk of childhood leukemia (OR=0.81, 95% CI: 0.67-0.97) and ALL (OR=0.77, 95% CI: 0.63-0.94), but not AML. Being first-born was associated with decreased risk of AML only (OR=0.70; 95% CI: 0.53-0.93). Compared to children with paternal age <25 years, children with paternal age between 35 and 45 years were at increased risk of total childhood leukemia (OR=1.12; 95% CI: 1.04-1.40) and ALL (OR=1.23; 95% CI: 1.04-1.47). None of conditions during pregnancy examined or maternal age were associated with increased risk of childhood leukemia or its subtypes.

*Conclusions*: Our results suggest that high birth weight and LGA were associated with increased risk and SGA with decreased risk of total childhood leukemia and ALL, being first-born was associated with decreased risk of AML, and advanced paternal age was associated with increased risk of ALL. These findings suggest that associations of childhood leukemia and perinatal factors depend highly on subtype of leukemia.

**Key words:** childhood leukemia, birth weight, birth order, parental age, perinatal factors, pregnancy complications

## Introduction

Childhood leukemia is the most common malignant disease in children worldwide and in the United States.[1] The incidence of childhood leukemia in California is 5.2 and 4.3 per 100,000 for males and females, respectively. Incidence is highest among the youngest age group (0-4 years), with 8.2 and 6.8 cases per 100,000 for males and females, respectively, then incidence declines with age until late adolescence.[2]

The etiology of childhood leukemia remains largely unknown. Several perinatal characteristics have been linked to childhood leukemia but the relation of others to leukemia and particularly to its subtypes remains to be elucidated. Birth weight is one of a few perinatal factors that have been consistently reported to be related to childhood leukemia risk,[3] with over 30 studies examining the association. Most reported positive association between birth weight and ALL; less consistent associations have been reported for AML.[3-14] Few studies have taken into account gestational age in the analysis of birth weight. Of those that did, most showed that large-for-gestational age babies were at increased risk of childhood leukemia.[12, 15]

Birth order is another perinatal factor that has been examined with regard to childhood leukemia risk. Most studies have found that high birth order was associated with decreased risk of ALL compared with first-born babies (OR varied from 0.57 to 0.95).[10, 16] A few studies reported increased risk of ALL with increasing birth.[3, 13] For AML, studies have detected either positive association or no association with birth order.[3, 10-12, 16, 17]

The majority of studies that looked at the relationships of maternal and paternal age to childhood leukemia detected an increased risk for older paternal age. [5, 13, 17-19] Some of these also reported an increased risk for older maternal age.[18, 19]

Conflicting results have been reported for the association between childhood leukemia and several perinatal and reproductive factors such as history of fetal loss, preeclampsia, polyhydramnios, anemia and genital herpes.[7] The majority of studies found no association but a few reported positive associations. [11, 20]

In a large case-control study linking data from the California Cancer Registry (1988-2008) and Birth Registry (1986-2007), we examined the association of childhood leukemia with perinatal factors, including birth weight, birth order and history of pregnancy terminations; maternal and paternal age as well as the difference between them; and complications during

pregnancy. The large size of the sample allowed for detailed analysis by two major subtypes of leukemia, ALL and AML, which was not possible in most previous studies.

**Materials and methods**

The California Cancer Registry (CCR), a population-based statewide cancer registry, was used to obtain information on all childhood leukemia cases diagnosed between 1988 and 2008 in children younger than 16 years who were born in California and resided in California at the time of diagnosis. The CCR is recognized as one of the leading cancer registries in the world with almost complete registration (99%). It routinely records age, race/ethnicity, sex, and residence at the time of diagnosis as well as information on almost all cancers, cancer subtypes and characteristics.[21] Controls who had not been diagnosed in California with any type of cancer were randomly selected from the California Birth Registry and were matched to cases (1 to 1) on the basis of date of birth (±6 months) and sex.

Information on birth weight, maternal and paternal age, history of pregnancy terminations before and after 20 weeks of gestation, number of live births living, number of live births deceased and maternal complications during pregnancy, as well as child's date of birth, gender, father's education, and ethnicity were extracted from California birth records. Gestational age was calculated based on last menstrual period and date of birth. Birth order was inferred from the number of live births living and number of live births deceased; first pregnancy was inferred from the number of live births living, number of live births deceased, and number of terminations before and after 20 weeks of gestation.

Birth weight was evaluated as a categorical variable with reference <2500 g, a continuous variable with units of 1000 g, and as a dichotomous variable with cut points at $<$ or $\geq$ 3,500 g and $<$ or $\geq$ 4,000 g. Some authors suggested that birth weight for gestational age as an indicator for fetal growth could be a better predictor for childhood leukemia. [15] Birth weight-for-gestational age was constructed using U.S. national reference for fetal growth and was classified into three categories: small-for-gestational age (SGA), average-for-gestational age (AGA) and large-for-gestational age (LGA).[22] To evaluate dose-response relationships between birth weight and childhood leukemia we obtained odds ratios using a moving window of birth weight. These analyses used birth weight categories (windows) of 2500 to <3500 g, 3000 to <4000 g, 3500 to <4500 g, 4000-5000 g, and >4500 g with a reference category <2500 g.

Gestational age, maternal and paternal ages, and history of terminations were analyzed as dichotomous and as categorical variables. Birth and pregnancy orders, maternal complications during pregnancy were used as dichotomous variables (first vs. other, or yes vs. no).

The primary analysis method was conditional logistic regression using the one-to-one age and gender matched case-control pairs.[23] Analyses were conducted with and without subjects with Down's syndrome, which is known to be associated with increased risk of childhood leukemia. [10, 11, 17, 24] Results of those analyses were similar. Here we present analyses that include subjects with Down's syndrome (46 subjects: 42 cases and 4 controls). We confirmed the absence of unduly influential observations by fitting a variety of models with different subsets of covariates and examining the results for outliers and influence. Covariates in the models were chosen based on information about known or potential confounders and model fit statistics. Models with minimal Akaike information criterion **(**AIC) value and the lowest number of potential confounders were chosen as main models presented in this paper. Father's education and payment source for delivery were used in all models as proxies for socioeconomic status (SES), a known confounder. For payment source for delivery, governmental programs such as Medicare, Medi-Cal and others as well as 'No care' were coded as low SES; private insurance and other sources of payment were coded as middle-high SES. For father's education, ≤12 years of education was considered low SES, 13-17 years as middle SES, and ≥17 years as high SES.

Despite the large number of cases and controls, sample sizes for some analyses were reduced due to missing data that was attributable largely to differences in the information collected on birth certificates from year to year. No differences in patterns of missingness were detected between cases and controls. Missing data were multiply imputed using Monte Carlo Markov chain full-data imputation under a missing at random assumption [25, 26] implemented by the MI procedure in SAS 9.1.[27] The imputation model included all variables used in models (except abnormal fetal conditions and pregnancy complications) and auxiliary variables likely to be correlated with variables of interest (number of pregnancy visits, month prenatal care began, type of birth, planned place of birth, number of ever born children, number of children born alive, and number of children born alive now deceased). Analyses were repeated on the multiply imputed data using the MIANALYZE procedure.

Analyses were conducted using statistical software SAS 9.1.[27]

The study was approved by University of California, Los Angeles Office for the Protection of Research Subjects.

**Results**

A total of 6645 childhood leukemia cases were identified from the California cancer registry. Linkage to birth records was successful for 87.1% (5788/6645) of cases. Of the 5788 cases (55.8% males and 44.2% females) included in this analysis, 4721 were ALL cases (56.2% males and 43.8% females), 852 were AML cases (53.3% males and 46.7% females), and 215 were other childhood leukemia types. The mean age at diagnosis was 4.9 years with a range of 0 to 15.4 years. Table 1 shows additional characteristics of study subjects.

We assessed the association of childhood leukemia and variables of interest in unadjusted and adjusted conditional logistic regressions matched for child's age and sex. Results of these analyses were not materially different; therefore we present adjusted results only.

*Birth weight*

In Table 2 we show results for conditional logistic regression for the association of childhood leukemia and birth weight and birth weight-for-gestational age. We observed an increased risk of total childhood leukemia and ALL for high birth weight babies using increments by 1000g (OR=1.11, 95% CI: 1.01-1.21 and OR=1.11, 95% CI: 1.01-1.23, respectively). Increased risk of total leukemia and ALL was also detected in all categories of birth weight with a reference category of < 2500g; the highest increase was noted for the heaviest babies, weighing ≥4500g (OR=1.59, 95% CI: 1.05-2.40 and OR=1.70, 95% CI: 1.08-2.68, respectively). Increased risk using cutpoints of 3500g and 4000 g for birth weight and birth weight-for-gestational age was less pronounced.

To address concerns about exposure misclassification due to chosen cut points for birth weight and to further examine trend, we obtained odds ratios for total leukemia and subtypes using a moving window for birth weight categorized as 2500- <3500g, 3000- <4000g, 3500- <4500g, 4000- >5000g, and ≥4500g with a reference category <2500g, with adjustment for

potential confounders (see Figure 1). These showed elevated risk for ALL but not for AML, compared to the reference category.

Moving window analysis for total leukemia is available online (Figure 1a).

Models with interactions between birth weight and birth order, mother's age, child's race and source of payment for care were also considered. All of these factors are potentially associated with childhood leukemia and could act as effect modifiers. None of the models suggested interaction between the examined variables (results not presented).

*Gestational age*

Gestational age adjusted for birth weight, birth order, mother's age, father's education, child's race, and payment source for delivery when entered in the model as a continuous variable with 1-week and 2-week increments had an OR of 0.97 (95% CI: 0.95-1.00) and an OR of 0.95 (95% CI: 0.90-1.00), respectively. Gestational age entered into the model as a categorical variable gave similar results (not presented).

*Birth order*

We found that first-born babies were at decreased risk for AML (OR: 0.70, 95% CI 0.53-0.93), but not for total leukaemia or ALL (Table 3). Similar results were obtained for pregnancy order with OR=0.73 and 95% CI 0.55-0.96 for AML.

*Parental age*

For maternal age, slightly increased risk for total childhood leukaemia and both subtypes were observed for ages greater than 35 years old but none of the estimates was precise (Table 3). For paternal age, results were similar but estimates were more precise, with increased risk of total childhood leukaemia and ALL for 5-year increase in age (OR=1.01, 95% CI 1.00-1.02), for fathers aged 35 years and older compared to younger fathers (OR=1.13, 95% CI 1.01-1.26 and OR=1.12, 95% CI 0.99-1.27, respectively), and for fathers aged 35-45 years compared to <25 years old (OR=1.12, 95% CI 1.04-1.40 and OR=1.23, 95% CI 1.04-1.47, respectively). The

difference between paternal and maternal ages was not associated with childhood leukaemia nor its subtypes.

*Other prenatal factors*

Neither history of pregnancy termination before or after 20 weeks of gestation nor maternal complications/conditions during pregnancy (Table 3) were associated with increased risk of childhood leukaemia or its subtypes.

*Multiple imputations*

In analyses repeated using multiply imputed data, the association between birth weight and childhood leukemia was slightly weaker than the same association using complete case analyses, with narrower confidence intervals and a similar trend. The association between birth weight-for-gestational age and childhood leukemia was slightly stronger and more precise. No other important differences were observed between complete case and multiply imputed data analyses. Table 4 with results of the association between birth weight and childhood leukemia estimated using multiple imputation is available online.

**Discussion**

Consistent with other studies, we observed an increased risk of childhood leukemia for high birth weight babies. We observed an 11% increase in risk of total childhood leukemia and acute lymphocytic leukemia per 1-kg increase in birth weight. For acute myeloid leukemia, no such risk increase was observed. In analyses with birth weight as a categorical variable with a reference category <2500g, increased risk was present for all babies above 2500g. The highest increase in risk for all types of leukemia (1.59-fold) and ALL (1.70-fold) was observed for the heaviest babies (≥4500g). For AML, results were consistent with increased risk for the highest weight category but estimates were less precise. In general, statistical power to detect associations with AML was more limited due to smaller numbers.

Results of a moving window analysis suggested a slightly positive trend of increasing risk with increasing birth weight for total leukemia and ALL compared to <2500g birth weight; for AML no trend was observed.

When we examined birth weight-for-gestational age, results suggested that large for gestational age babies were at slightly increased risk of total childhood leukemia and both ALL and AML compared to average for gestational age babies, but the estimates were imprecise. Being small for gestational age was associated with reduced risk of total childhood leukemia and ALL, but not for AML. Our findings confirm results of other studies that looked at birth weight-for-gestational age [12].

There are two main theories explaining the association between high birth weight and childhood cancers, including childhood leukemia. The first theory is related to insulin-like growth factor 1 (IGF-1), a known procarcinogenic agent. IGF-1 level is associated with birth weight and could play a role in the development of childhood leukemia through the induction of pre-leukemic cell division.[16, 28] Another hypothesis suggests that, because there is an association between birth weight and bone marrow volume, i.e., the number of bone marrow cells, children with a higher birth weight have more cells at risk of malignant transformation, and thus are at a greater risk of leukemia.[16, 29-31] These two hypotheses are not mutually exclusive, e.g., pre-leukemic cells may secrete growth factors that increase birth weight.[28]

We observed decrease in risk of AML for first-born babies but no association with total leukemia and ALL. Some previous studies found an increased risk of total childhood leukemia and ALL with increasing birth order;[13, 32] a few reported an increased risk of childhood leukemia for first-born babies.[3, 10, 13, 16] Other studies found very weak or no association of birth order and childhood leukemia.[11, 18] For AML, most studies detected either increase in risk or no change in risk with increasing birth order.[3, 10-12, 16, 17] In developed countries, birth order is considered a proxy for exposure to infections in early childhood. [11] In recent years, use of day care facilities for first-born babies has increased due to growing numbers of working mothers. Day care use is associated with exposure to infections in early childhood.[11, 16, 33] These factors could explain the conflicting results of studies on birth order and childhood leukemia in recent years.

Our results for analysis of the association of maternal and paternal ages and childhood leukemia were very similar to findings of another study conducted in California.[22] We observed a small risk associated with older paternal age: for 5-year increase in paternal age, for father's aged ≥35 compared to younger fathers, and for father's aged 35-45 compared to father's ≤25 years of age. Small increase in risk was observed for older maternal age but estimates were imprecise.

We also considered a difference between paternal and maternal ages but we observed no associations with childhood leukemia.

Few studies looked at the relationships of childhood leukemia and perinatal factors such as pregnancy terminations and complications/conditions during pregnancy. Most found no association; only two studies found a small increase in risk of childhood leukemia in children of women with polyhydramnios and anemia [11] and genital herpes during index pregnancy.[3] In our study we found no association of pregnancy terminations or complications during pregnancy either with total leukemia risk or with ALL and AML. An unexpected finding was a decrease in risk with tobacco smoking for total leukemia and ALL, but these results did not stand for AML. Similar results were observed in Danish study by Westergaard. [10] These results could be related to the association of tobacco smoking and small-for-gestational age birth weight. [34] Mothers who smoke are at increased risk of having SGA babies [34] and SGA has a protective effect for childhood leukemia. Alternative explanations for these findings could include random error due to small numbers, misclassification issues, and probably very poor and inconsistent reporting of this variable on birth certificates. Therefore, our results on tobacco smoking and childhood leukemia should be interpreted with great caution.

The large size of the dataset allowed us to conduct separate analyses for two main subtypes of childhood leukemia, AML and ALL. The risk patterns were quite different for the two subtypes, which supports a theory that they have different etiology.

One of the major advantages of the current study was that our data were based on population registries with almost complete registration of births and cancers in California and controls were randomly selected from birth registry rather than by recruiting volunteers as in many case-control studies. Since these registries are independent of each other and participation of subjects was not required for our data collection, selection bias was unlikely in this study. Misclassification of outcome status was also unlikely in this study due to the completeness and high accuracy of the California Cancer Registry. Misclassification of birth weight was possible but unlikely. We think that birth weight is usually recorded fairly accurately. For other perinatal factors such as complications during pregnancy and abnormal fetal conditions, misclassification was possible due to many missing values and questionable accuracy of reporting.

We adjusted for potential confounders that were available in registries. There was no information available on such potential confounders as maternal and paternal occupation, diet,

and alcohol and drug abuse, maternal health conditions before pregnancy, fertility treatments and procedures and child's birth defects. Therefore, residual confounding was possible. Maternal occupation and alcohol and drug use are inversely associated with birth weight [35-37] but positively associated with childhood leukemia. [38-40] Therefore, confounding due to such variables would most likely pull estimates toward the null.

One of the limitations of the study was missing data. However, since information was missing mainly due to differences in the information collected on birth certificates from year to year rather than non-response, the potential for biases was probably small, and the impact was mainly on the precision of the estimates. There were no differences in the pattern of missingness between cases and controls. We repeated all analyses using multiply imputed dataset and obtained similar results but with narrower confidence intervals.

In summary, we found that high birth weight and LGA were associated with increased risk and SGA was associated with decreased risk of total childhood leukemia and ALL. For AML, increased risk was found for heaviest babies but estimates were imprecise. We also found that being first-born was associated with decreased risk of all leukemia types combined and AML, but not ALL. Increased risk of total leukemia and ALL was observed for advanced paternal age. These findings suggest that associations of childhood leukemia and perinatal factors depend highly on subtype of leukemia, which needs further evaluation.

**Acknowledgements**

This project was supported by a research contract from the Electric Power Research Institute (EPRI) to the UCLA and by UCLA Faculty Grants Program.

Crespi was also partially supported by National Institutes of Health grant P30 CA16042.

## References


1. Ma, S.K., T.S. Wan, and L.C. Chan, *Cytogenetics and molecular genetics of childhood leukemia.* Hematol Oncol, 1999. **17**(3): p. 91-105.
2. Campleman, S.L., Wright, W.E., *Childhood Cancer in California 1988 to 1999. Volume I: Birth to Age 14.*, in *Cancer Surveillance Section*, C.D.o.H. Services, Editor. 2004, California Cancer Registry: Sacramento, CA.
3. Milne, E., et al., *Fetal growth and acute childhood leukemia: looking beyond birth weight.* Am J Epidemiol, 2007. **166**(2): p. 151-9.
4. Adam, M., et al., *Childhood leukaemia and socioeconomic status: what is the evidence?* Radiat Prot Dosimetry, 2008. **132**(2): p. 246-54.
5. Kaye, S.A., et al., *Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia.* Cancer, 1991. **68**(6): p. 1351-5.
6. Robison, L.L. and J.P. Neglia, *Epidemiology of Down syndrome and childhood acute leukemia.* Prog Clin Biol Res, 1987. **246**: p. 19-32.
7. Ross, J.A., et al., *Evaluating the relationships among maternal reproductive history, birth characteristics, and infant leukemia: a report from the Children's Cancer Group.* Ann Epidemiol, 1997. **7**(3): p. 172-9.
8. Paltiel, O., et al., *Birth weight and other risk factors for acute leukemia in the Jerusalem Perinatal Study cohort.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(6): p. 1057-64.
9. Hjalgrim, L.L., et al., *Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies.* Am J Epidemiol, 2003. **158**(8): p. 724-35.
10. Westergaard, T., et al., *Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study.* J Natl Cancer Inst, 1997. **89**(13): p. 939-47.
11. Roman, E., et al., *Perinatal and reproductive factors: a report on haematological malignancies from the UKCCS.* Eur J Cancer, 2005. **41**(5): p. 749-59.
12. Schuz, J. and M.R. Forman, *Birthweight by gestational age and childhood cancer.* Cancer Causes Control, 2007. **18**(6): p. 655-63.
13. Ou, S.X., et al., *Birth characteristics, maternal reproductive history, hormone use during pregnancy, and risk of childhood acute lymphocytic leukemia by immunophenotype (United States).* Cancer Causes Control, 2002. **13**(1): p. 15-25.
14. Caughey, R.W. and K.B. Michels, *Birth weight and childhood leukemia: a meta-analysis and review of the current evidence.* Int J Cancer, 2009. **124**(11): p. 2658-70.
15. Sprehe, M.R., et al., *Comparison of birth weight corrected for gestational age and birth weight alone in prediction of development of childhood leukemia and central nervous system tumors.* Pediatr Blood Cancer, 2010. **54**(2): p. 242-9.
16. Hjalgrim, L.L., et al., *Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland.* J Natl Cancer Inst, 2004. **96**(20): p. 1549-56.
17. Cnattingius, S., et al., *Prenatal and neonatal risk factors for childhood myeloid leukemia.* Cancer Epidemiol Biomarkers Prev, 1995. **4**(5): p. 441-5.
18. Reynolds, P., J. Von Behren, and E.P. Elkin, *Birth characteristics and leukemia in young children.* Am J Epidemiol, 2002. **155**(7): p. 603-13.
19. Dockerty, J.D., et al., *Case-control study of parental age, parity and socioeconomic level in relation to childhood cancers.* Int J Epidemiol, 2001. **30**(6): p. 1428-37.

20. Dorak, M.T., et al., *Examination of gender effect in birth weight and miscarriage associations with childhood cancer (United Kingdom).* Cancer Causes Control, 2007. **18**(2): p. 219-28.
21. The California Cancer Registry. *Mission Statement & Purpose*. [cited March 2009]; Available from: http://www.ccrcal.org/abouttheccr.html.
22. Alexander, G.R., et al., *A United States national reference for fetal growth.* Obstet Gynecol, 1996. **87**(2): p. 163-8.
23. Hosmer, D.W., Jr., Lemeshow, S., *Applied Logistic Regression*. 2000, New York: Wiley.
24. Cnattingius, S., et al., *Prenatal and neonatal risk factors for childhood lymphatic leukemia.* J Natl Cancer Inst, 1995. **87**(12): p. 908-14.
25. Rubin, D.B. and N. Schenker, *Multiple imputation in health-care databases: an overview and some applications.* Stat Med, 1991. **10**(4): p. 585-98.
26. Little, R., Rubin, D., *Statistical Analysis with Missing Data*. 2nd Edition ed. 2002, New York: Wiley.
27. SAS Institute, I., *SAS Statistical Software*. 2006: Cary, NC.
28. Ross, J.A., et al., *Big babies and infant leukemia: a role for insulin-like growth factor-1?* Cancer Causes Control, 1996. **7**(5): p. 553-9.
29. Wilpshaar, J., et al., *Magnetic resonance imaging of fetal bone marrow for quantitative definition of the human fetal stem cell compartment.* Blood, 2002. **100**(2): p. 451-7.
30. Albanes, D. and M. Winick, *Are cell number and cell proliferation risk factors for cancer?* J Natl Cancer Inst, 1988. **80**(10): p. 772-4.
31. Albanes, D., *Energy balance, body size, and cancer.* Crit Rev Oncol Hematol, 1990. **10**(3): p. 283-303.
32. Altieri, A., et al., *Number of siblings and the risk of lymphoma, leukemia, and myeloma by histopathology.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(7): p. 1281-6.
33. Roman, E., et al., *Childhood acute lymphoblastic leukemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study.* Am J Epidemiol, 2007. **165**(5): p. 496-504.
34. Mitchell, E.A., et al., *Smoking, nicotine and tar and risk of small for gestational age babies.* Acta Paediatr, 2002. **91**(3): p. 323-8.
35. Bada, H.S., et al., *Low birth weight and preterm births: etiologic fraction attributable to prenatal drug exposure.* J Perinatol, 2005. **25**(10): p. 631-7.
36. Day, N.L., et al., *Prenatal exposure to alcohol: effect on infant growth and morphologic characteristics.* Pediatrics, 1989. **84**(3): p. 536-41.
37. Richardson, G.A., et al., *Growth of infants prenatally exposed to cocaine/crack: comparison of a prenatal care and a no prenatal care sample.* Pediatrics, 1999. **104**(2): p. e18.
38. Belson, M., B. Kingsley, and A. Holmes, *Risk factors for acute leukemia in children: a review.* Environ Health Perspect, 2007. **115**(1): p. 138-45.
39. Severson, R.K., et al., *Cigarette smoking and alcohol consumption by parents of children with acute myeloid leukemia: an analysis within morphological subgroups--a report from the Childrens Cancer Group.* Cancer Epidemiol Biomarkers Prev, 1993. **2**(5): p. 433-9.
40. van Duijn, C.M., et al., *Risk factors for childhood acute non-lymphocytic leukemia: an association with maternal alcohol consumption during pregnancy?* Cancer Epidemiol Biomarkers Prev, 1994. **3**(6): p. 457-60.

**Table 1.** Socio-demographic and perinatal characteristics of study subjects, California birth registry, 1986-2007.

| Variables | Total # | Cases (%) | Controls (%) | ALL # cases/controls [a] | AML # cases/controls [a] |
|---|---|---|---|---|---|
| **All** | 11576 | 5788 | 5788 | 4721 | 852 |
| **Child's age** | | | | | |
| < 1 year | 791 | (6.7) | (7.1) | 178/196 | 156/161 |
| 1-5 years | 7381 | (64.1) | (63.7) | 3196/3182 | 413/407 |
| 6-9 years | 2007 | (17.4) | (17.3) | 849/845 | 131/130 |
| 10-15 years | 1373 | (11.9) | (11.9) | 497/497 | 152/154 |
| Missing # | 24 | 12 | 12 | 1/1 | 0/0 |
| **Mother's age** | | | | | |
| < 25 years | 3952 | (32.6) | (35.7) | 1530/1689 | 296/301 |
| 25-35 years | 5998 | (52.3) | (51.4) | 2495/2430 | 415/435 |
| 35-45 years | 1603 | (14.9) | (12.8) | 687/594 | 136116 |
| ≥45 years | 21 | (0.2) | (0.1) | 9/7 | 5/0 |
| Missing # | 2 | 1 | 1 | 0/1 | 0/0 |
| **Father's age** | | | | | |
| < 25 years | 2354 | (20.7) | (22.4) | 925/993 | 175/179 |
| 25-35 years | 5808 | (53.1) | (53.2) | 2405/2373 | 405/407 |
| 35-45 years | 2443 | (23.5) | (21.2) | 1041/933 | 191/170 |
| ≥45 years | 323 | (2.8) | 169 (3.1) | 121/140 | 27/22 |
| Missing # | 648 | 286 | 362 | 228/281 | 46/66 |
| **Child's race** | | | | | |
| White | 8889 | (81.6) | (78.6) | 3777/3559 | 610/622 |
| Black | 780 | (5.2) | (8.9) | 198/399 | 76/68 |
| Asian | 1214 | (11.3) | (10.6) | 494/474 | 107/88 |
| Other | 209 | (1.9) | (1.9) | 81/82 | 21/18 |
| Missing # | 484 | 214 | 270 | 171/207 | 38/56 |
| **Father's education** | | | | | |
| ≤12 years | 6352 | (77.9) | (77.6) | 2600/2582 | 484/461 |
| 13-17 years | 1307 | (15.8) | (16.2) | 554/562 | 68/76 |
| ≥17 years | 509 | (6.3) | (6.2) | 213/208 | 3536 |
| Missing [Not collected] # [b] | 3408 [3014] | 1691 [1509] | 1717 [1505] | 1354/1369 | 265/279 |

| | | | | | |
|---|---|---|---|---|---|
| **Source for payment for delivery** | | | | | |
| Governmental programs | 4457 | (44.0) | (46.4) | 1741/1861 | 346350 |
| Other insurance | 5397 | (56.0) | (53.6) | 2296/2166 | 378/367 |
| Missing [Not collected] # [b] | 1722 [1695] | 853 [845] | 869 [850] | 684/694 | 128/135 |
| **Birth weight** | | | | | |
| <2500 g | 620 | (5.8) | (4.9) | 219/274 | 52/53 |
| 2500-3000 g | 1654 | (14.8) | (13.8) | 634/695 | 138/127 |
| 3000-3500 g | 4377 | (38.2) | (37.4) | 1783/1782 | 313/355 |
| 3500-4000 g | 1654 | (30.1) | (30.7) | 1449/1435 | 252/235 |
| 4000-4500 g | 1170 | (9.2) | (11.0) | 527/441 | 83/71 |
| ≥4500 g | 238 | (1.9) | (2.2) | 109/94 | 14/11 |
| Missing # | 2 | 1 | 1 | 1/1 | 0/0 |
| **Gestational age** | | | | | |
| <28 weeks | 47 | (0.3) | (0.5) | 13/23 | 4/6 |
| 28-37 weeks | 1103 | (10.8) | (9.4) | 453/418 | 116/81 |
| 37-42 weeks | 8842 | (80.6) | (81.4) | 3629/3602 | 619/657 |
| >42 weeks | 927 | (8.3) | (8.7) | 376/385 | 62/60 |
| Missing/implausible values # | 657 | 311 | 346 | 250/293 | 51/48 |
| **Birth weight for gestational age** | | | | | |
| SGA | 862 | (7.0) | (8.8) | 295/386 | 70/71 |
| AGA | 8593 | (78.4) | (79.0) | 3516/3492 | 627/646 |
| LGA | 1464 | (14.6) | (12.2) | 660/550 | 104/87 |
| Missing/implausible values # | 657 | 311 | 346 | 250/293 | 51/48 |
| **Birth order** | | | | | |
| First | 4556 | (38.7) | (40.7) | 1840/1896 | 309/356 |
| Other | 7008 | (61.3) | (59.3) | 2878/2819 | 496 |
| Missing # | 12 | 6 | 6 | 3/6 | 1/0 |
| **History of terminations before 20 weeks** | | | | | |
| No | 9685 | (83.4) | (84.1) | 3923/3964 | 724/720 |
| Yes | 1880 | (16.6) | (16.0) | 795/749 | 128/132 |
| Missing # | 11 | 3 | 8 | 3/8 | 0/0 |
| **History of terminations after 20 weeks** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| No | 11392 | (98.5) | (98.6) | 4652/4643 | 832/844 |
| Yes | 170 | (1.5) | (1.4) | 66/69 | 20/8 |
| Missing # | 14 | 4 | 10 | 3/9 | 0/0 |
| **Maternal conditions during index pregnancy** [c] | | | | | |
| Preeclampsia/Eclampsia | 182 | (1.6) | (1.5) | 78/69 | 13/15 |
| Anemia | 82 | (0.7) | (0.7) | 32/35 | 5/5 |
| Genital herpes | 129 | (1.1) | (1.2) | 54/51 | 5/15 |
| Chronic diseases | 305 | (2.7) | (2.6) | 133/128 | 16/16 |
| Polyhydramnios | 51 | (0.5) | (0.4) | 19/17 | 10/3 |
| Blood and immune disorders | 105 | (0.8) | (1.0) | 38/46 | 9/9 |
| Tobacco use | 162 | (1.1) | (1.7) | 49/78 | 14/16 |
| Missing # | 10 | 5 | 5 | 3/5 | 2/0 |

[a] Number of cases and controls for ALL and AML do not add up for the total number of cases and controls for childhood leukemia because there were few other subtypes in the dataset.

[b] Patterns of missingness varied by variable and by year due to differences in data collection for different years.

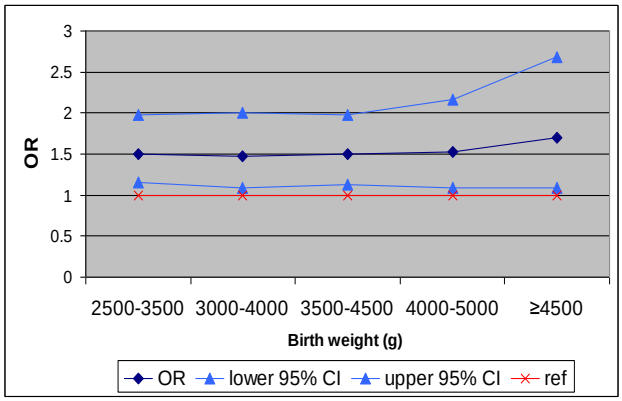
[c] Conditions are not mutually exclusive and thus frequencies exceed the totals. Information was not routinely collected for many of these conditions prior 1990.

**Table 2**. Conditional odds ratios (OR) and 95% confidence intervals (CI) for childhood leukemia associated with birth weight and birth weight-for-gestational age matched on age and sex and adjusted for gestational age, birth order, mother's age, father's education, child's race, and payment source for delivery. California birth registry, 1986-2007.

| | **All types** (3334 cases/3334 controls) | | | **ALL** (2744 cases/2744 controls) | | | **AML** (462 cases/462 controls) | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% Confidence Intervals** | | **OR** | **95% Confidence Intervals** | | **OR** | **95% Confidence Intervals** | |
| **Birth weight (g)** | | | | | | | | | |
| Reference <2500 | 1 | - | - | 1 | - | - | 1 | - | - |
| 2500-3500 | 1.39 | 1.09 | 1.77 | 1.51 | 1.15 | 1.98 | 1.03 | 0.54 | 1.98 |
| 3500-4500 | 1.47 | 1.14 | 1.88 | 1.50 | 1.13 | 1.98 | 1.26 | 0.65 | 2.44 |
| ≥4500 | 1.59 | 1.05 | 2.40 | 1.70 | 1.08 | 2.68 | 1.45 | 0.45 | 4.70 |
| **Birth weight** (≥3,500g vs <3,500g) | 1.08 | 0.98 | 1.20 | 1.03 | 0.92 | 1.15 | 1.23 | 0.93 | 1.63 |
| **Birth weight** (≥4,000g vs <4,000g) | 1.05 | 0.90 | 1.22 | 1.04 | 0.89 | 1.23 | 1.03 | 0.66 | 1.61 |
| **Birth weight, 1000g increase** | 1.11 | 1.01 | 1.21 | 1.11 | 1.01 | 1.23 | 1.01 | 0. 80 | 1.28 |
| **Birth weight- for-gestational age** | | | | | | | | | |
| SGA | 0.81 | 0.67 | 0.97 | 0.77 | 0.63 | 0.94 | 1.15 | 0.71 | 1.88 |
| AGA | 1 | - | - | 1 | - | - | 1 | - | - |
| LGA | 1.10 | 0.95 | 1.27 | 1.07 | 0.91 | 1.25 | 1.13 | 0.75 | 1.68 |

**Figure 1**. Odds ratios (OR) and 95% confidence intervals (CI) for acute lymphocytic leukemia and acute myeloid leukemia at moving windows of birth weight matched on age and sex and adjusted for child's race, gestational age, mother's age, birth order, father's education and source of payment for delivery. Reference level: < 2500g.

a) Acute lymphocytic leukemia




b) Acute myeloid leukemia

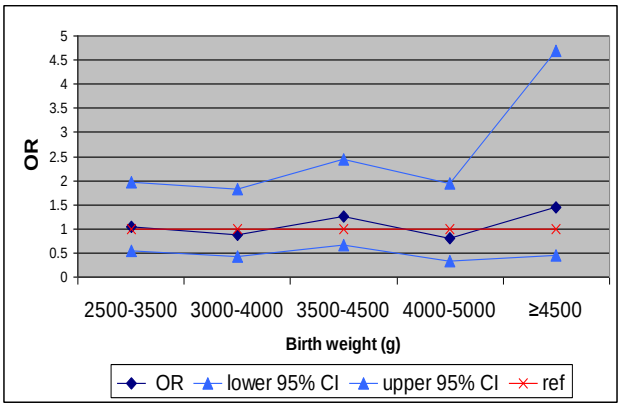


**Table 3**. Conditional odds ratios (OR) and 95% confidence intervals (CI) for childhood leukemia associated with several perinatal factors matched on age and sex and adjusted for confounders. California birth registry, 1986-2007.

| | All leukemia types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variables** | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Birth order** (1st vs Other) [a] | 0.93 | 0.84 | 1.03 | 0.97 | 0.87 | 1.08 | 0.70 | 0.53 | 0.93 |
| **Pregnancy order** (1st vs Other) [a] | 0.96 | 0.87 | 1.07 | 1.00 | 0.89 | 1.12 | 0.73 | 0.55 | 0.96 |
| **History of terminations before week 20** [d] | 0.95 | 0.83 | 1.09 | 0.97 | 0.83 | 1.12 | 0.94 | 0.65 | 1.37 |
| **History of termination after week 20** [d] | 1.13 | 0.74 | 1.73 | 1.02 | 0.65 | 1.62 | 2.81 | 0.74 | 10.66 |
| **Mother's age** [b] | | | | | | | | | |
| By 5-year increase | 1.01 | 0.99 | 1.02 | 1.00 | 0.98 | 1.02 | 1.02 | 0.98 | 1.06 |
| ≥35 years vs < 35 years | 1.14 | 0.88 | 1.49 | 1.07 | 0.80 | 1.45 | 1.52 | 0.81 | 2.87 |
| ≥40 years vs < 40 years | 1.10 | 0.82 | 1.74 | 0.92 | 0.52 | 1.64 | 2.18 | 0.77 | 6.14 |
| 25-35 years vs < 25 years | 1.00 | 0.85 | 1.32 | 1.13 | 0.88 | 1.44 | 0.84 | 0.48 | 1.48 |
| 35-45 years vs < 25 years | 1.13 | 0.87 | 1.61 | 1.19 | 0.83 | 1.68 | 1.31 | 0.64 | 2.70 |
| ≥45 years vs < 25 years | 1.52 | 0.33 | 6.93 | 0.76 | 0.12 | 4.71 | - | - | - |
| **Father's age** [c] | | | | | | | | | |
| By 5-year increase | 1.01 | 1.00 | 1.02 | 1.01 | 1.00 | 1.02 | 1.00 | 0.98 | 1.02 |
| ≥35 years vs < 35 years | 1.13 | 1.01 | 1.26 | 1.12 | 0.99 | 1.27 | 1.11 | 0.90 | 1.27 |
| ≥40 years vs < 40 years | 1.11 | 0.95 | 1.30 | 1.10 | 0.92 | 1.31 | 1.02 | 0.74 | 1.61 |
| 25-35 years vs < 25 years | 1.03 | 0.91 | 1.17 | 1.07 | 0.93 | 1.23 | 0.80 | 0.57 | 1.12 |
| 35-45 years vs < 25 years | 1.12 | 1.04 | 1.40 | 1.23 | 1.04 | 1.47 | 0.90 | 0.60 | 1.35 |
| ≥45 years vs < 25 years | 0.90 | 0.68 | 1.20 | 0.89 | 0.65 | 1.22 | 0.92 | 0.44 | 1.94 |
| **Difference between paternal and maternal age** [c] | | | | | | | | | |
| 5-10 years vs < 5 years | 0.95 | 0.81 | 1.11 | 0.97 | 0.82 | 1.15 | 0.95 | 0.58 | 1.54 |
| ≥10 years vs < 5 years | 0.84 | 0.65 | 1.08 | 0.87 | 0.65 | 1.15 | 0.78 | 0.41 | 1.46 |
| **Conditions during index pregnancy** [d] | | | | | | | | | |
| Eclampsia | 1.08 | 0.75 | 1.57 | 1.19 | 0.78 | 1.80 | 0.84 | 0.33 | 2.15 |
| Anemia | 1.09 | 0.64 | 1.87 | 1.13 | 0.62 | 2.05 | 1.16 | 0.22 | 6.17 |

| | All leukemia types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variables** | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| Blood disorders | 0.76 | 0.48 | 1.23 | 0.84 | 0.50 | 1.42 | 0.61 | 0.17 | 2.18 |
| Chronic conditions | 1.07 | 0.80 | 1.43 | 0.99 | 0.72 | 1.36 | 1.24 | 0.48 | 3.21 |
| Tobacco use | 0.59 | 0.38 | 0.92 | 0.56 | 0.33 | 0.93 | 0.84 | 0.31 | 2.33 |
| Polyhydramnios | 1.10 | 0.59 | 2.06 | 0.95 | 0.46 | 1.94 | 2.32 | 0.44 | 12.28 |
| Genital herpes | 0.81 | 0.43 | 1.54 | 0.87 | 0.43 | 1.78 | 0.46 | 0.09 | 2.44 |

[a] Adjusted for child's race, mother's age, source of payment for delivery
[b] Adjusted for child's race, birth order, father's education, terminations after 20 weeks of gestation, source of payment for delivery
[c] Adjusted for child's race, birth order, father's education, and source of payment for delivery
[d] Adjusted for child's race, birth weight, gestational age, birth order, mother's age, source of payment for delivery

**Figure 1a**. Odds ratios (OR) and 95% confidence intervals (CI) for childhood leukemia at moving windows of birth weight, matched on age and sex and adjusted for child's race, gestational age, mother's age, birth order, father's education and source of payment for delivery. Reference level: < 2500g.

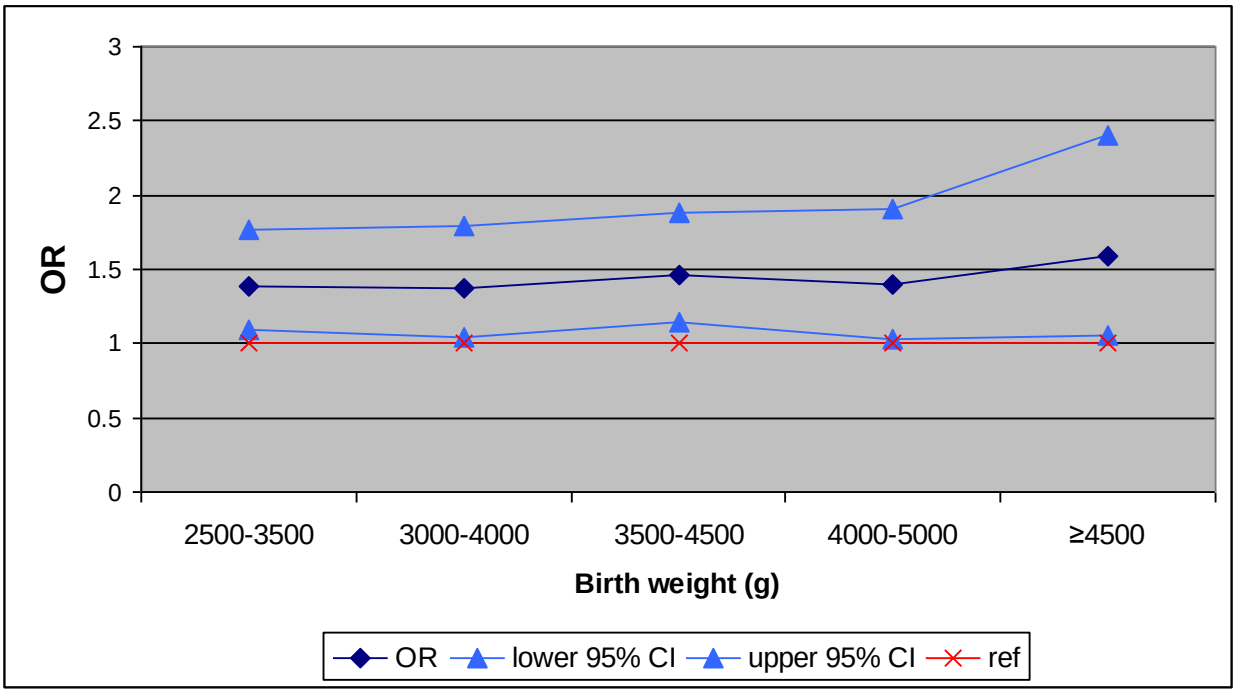


**Table 4.** Conditional odds ratios for childhood leukemia associated with birth weight and birth weight for gestational age matched on age and sex and adjusted for gestational age, birth order, mother's age, father's education, child's race, and payment source for delivery based on multiple imputation of missing data. California birth registry, 1986-2007.

| | **All leukemia types** | | | **ALL** | | | **AML** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Birth weight (g)** | | | | | | | | | |
| Reference <2500 | 1 | - | - | 1 | - | - | 1 | - | - |
| 2500-3500 | 1.25 | 1.04 | 1.50 | 1.27 | 1.04 | 1.56 | 1.22 | 0.76 | 1.94 |
| 3500-4500 | 1.36 | 1.13 | 1.64 | 1.35 | 1.09 | 1.67 | 1.46 | 0.90 | 2.34 |
| >4500 | 1.45 | 1.06 | 1.98 | 1.48 | 1.05 | 2.08 | 1.79 | 0.71 | 4.52 |
| **Birth weight** (≥3,500g vs <3,500g) | 1.11 | 1.03 | 1.20 | 1.08 | 0.99 | 1.18 | 1.22 | 1.00 | 1.50 |
| **Birth weight** (≥4,000g vs <4,000g) | 1.19 | 1.06 | 1.34 | 1.19 | 1.05 | 1.34 | 1.20 | 0.87 | 1.66 |
| **Birth weight,** 1000g increase | 1.16 | 1.09 | 1.24 | 1.17 | 1.08 | 1.25 | 1.12 | 0.95 | 1.34 |
| **Birth weight-for-gestational age** | | | | | | | | | |
| SGA vs AGA | 0.82 | 0.71 | 0.94 | 0.78 | 0.66 | 0.92 | 0.93 | 0.63 | 1.37 |
| LGA vs AGA | 1.19 | 1.06 | 1.32 | 1.18 | 1.04 | 1.33 | 1.23 | 0.90 | 1.68 |

## Chapter 4

# Birth weight and other perinatal factors and childhood CNS tumors: a case-control study in California

Oksuzyan S[1], Crespi CM[2], Cockburn M[3], Mezei G[4], Kheifets L[5]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, soksuzyan@ucla.edu; [2]Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California, ccrespi@ucla.edu; [3]Department of Preventive Medicine, University of Southern California, Los Angeles, California, myles@med.usc.edu; [4]Environment Department, Electric Power Research Institute, Palo Alto, California, gmezei@epri.com; [5]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, kheifets@ucla.edu.

**Abstract**

*Aims:* We conducted a large registry-based study in California to investigate the association of perinatal factors and childhood CNS tumors, with analysis by tumor subtype.

*Methods:* We linked California cancer and birth registries to obtain information on 3308 cases and 3308 controls matched on age and sex. We examined the association of birth weight, gestational age, birth order, parental ages, maternal conditions during pregnancy, newborn abnormalities and the risk of childhood CNS tumors using conditional logistic regression, with adjustment for potential confounders.

*Results:* The odds ratio (OR) per 1000 gram increase in birth weight was 1.11 (95% CI: 0.99-1.24) for total childhood CNS tumors, 1.17 (95% CI: 0.97-1.42) for astrocytoma and 1.28 (95% CI: 0.90-1.83) for medulloblastoma. Compared to average-for-gestational age, large-for-gestational age infants were at increased risk of glioma (OR=1.86, 95% CI: 0.99-3.48), while

small-for-gestational age infants were at increased risk of ependimoma (OR=2.64, 95% CI: 1.10-6.30). Increased risk of childhood CNS tumors was observed for 5-year increase in maternal and paternal ages (OR=1.06, 95% CI: 1.00-1.12 and 1.05, 95% CI: 1.00-1.10 respectively). Increased risk of astrocytoma was detected for 5-year increase in paternal age (OR=1.08; 95% CI:1.00-1.16) and increased risk of glioma for maternal age ≥35 years old (OR=1.87; 95% CI:1.00-3.52). Maternal genital herpes during pregnancy was associated with a pronounced increase in risk of total CNS tumors (OR=2.74; 95% CI:1.16-6.51). Other (non-sexually transmitted) infections during pregnancy were associated with decreased risk of total CNS tumors (OR=0.28, 95% CI: 0.09-0.85). Maternal blood/immune disorders during pregnancy were linked to increased risk of CNS tumors (OR=2.28, 95% CI: 1.08-4.83) and medulloblastoma (OR=7.13, 95% CI: 0.82-61.03). Newborn CNS abnormalities were also associated with high risk of childhood CNS tumors (OR=4.08, 95% CI: 1.13-14.76).

*Conclusions*: Our results suggest that maternal genital herpes, blood and immunological disorders during pregnancy and newborn CNS abnormalities were associated with increased risk of CNS tumors. Maternal infections during pregnancy were associated with decreased risk of CNS tumors. Advanced maternal and paternal ages may be associated with a slightly increased risk of CNS tumors. Factors associated with CNS tumor subtypes varied by subtype, an indicator of different etiology for different subtypes.

**Key words:** childhood CNS tumors, childhood brain tumors, birth weight, birth order, parental age, perinatal factors, pregnancy complications, newborn abnormalities

## Introduction

Brain and central nervous system (CNS) cancers are the second most common malignancy in childhood after childhood leukemia and the most common of the solid tumors, making up 16.6-21% of all childhood cancers in the U.S.[1-3]

Very little is known about the etiology of primary CNS and brain tumors. Several maternal and perinatal characteristics are suspected risk factors for CNS tumors in children but evidence for most is limited or inconsistent. Some consistent findings have been reported for exposure to infections during pregnancy and perinatally, with the majority of studies finding an increased risk of CNS tumors for children whose mothers were exposed to infections during pregnancy. [4-6] The polyomavirus family has been suspected of playing a role in the development of CNS and brain cancers since their DNA has been found in brain tumor tissues, but whether the virus causes these tumors remains unknown. [7-9]

Findings on the association of birth weight and childhood CNS tumors are inconsistent. Two studies reported an association between CNS tumors and low birth weight. [10, 11] Other investigators found an increased risk of certain subtypes of CNS tumors in children with high birth weight but the associations were dependent on reference group and most of these associations were imprecise. Some studies detected increased risk of astrocytoma for high birth weight (>4000 g) compared with children with birth weight of 2500–3999 g, but no risk for other subtypes. [12, 13] A meta-analysis of 8 studies reported slightly increased risk of astrocytoma and medulloblastoma for high birth weight children. [14] Most other studies found no association with birth weight or birth weight-for-gestational age. [15, 16]

The majority of studies that have examined birth order and risk of childhood brain tumors have found no association. [11-13, 17, 18] However, two studies have reported a moderate increase in risk of childhood brain tumors among children who were first born, [19] and several studies found an increase associated with being second or higher born.[6, 20]

Maternal history of miscarriages was found to be protective against childhood CNS tumors in one study. [13] In another study maternal history of one miscarriage was associated with slightly increased risk of malignant CNS tumors.[21] Most other studies did not find any association. [5, 10, 11, 18]

A very few studies have found that the risk of brain tumors increases with advanced maternal age [22]; the majority of studies did not observe an association. [5, 10-13, 18, 23] For paternal age, findings are also inconsistent. [23, 24]

The majority of studies on parental smoking before and during pregnancy found no or limited association with CNS tumors. [5, 18, 25]

Most previous studies on perinatal risk factors for CNS tumors have had limited numbers of cases. Very few studies have looked at such risk factors as newborn abnormalities and selected maternal complications/conditions during index pregnancy.

In a large case-control study linking data from the California cancer (1988-2008) and birth (1986-2007) registries, we examined the association of childhood CNS tumors with perinatal factors, including birth weight, birth order and history of pregnancy terminations; maternal and paternal age; newborn abnormalities; and complications/conditions during index pregnancy. The large size of the sample allowed for detailed analysis by subtypes of tumors, which was not possible in most previous research.

**Materials and methods**

The California Cancer Registry (CCR), a statewide population-based cancer registry, was used to obtain information on diagnosis and subtype of all CNS and brain tumor cases diagnosed between 1988 and 2008 in children younger than 16 years who were born in California and resided in California at the time of diagnosis. The CCR is recognized as one of the leading cancer registries in the world with almost complete registration (99%). The cancer registry routinely recorded age, race/ethnicity, sex, and residence at the time of diagnosis as well as information on cancer types and characteristics.[26] Control subjects were randomly selected from the California Birth Registry and matched to cases (1 to 1) on the basis of date of birth (±6 months) and sex; controls who had been diagnosed with any type of cancer in California by the age of diagnosis of a case were excluded.

Information on birth weight, birth order, maternal and paternal age, maternal history of pregnancy terminations, complications/conditions during index pregnancy, and newborn abnormalities as well as on child's date of birth, gender, maternal and paternal education, and ethnicity were extracted from California birth records.

Birth weight was evaluated as a categorical variable with 500 g and 1000 g increments and reference <2500 g, as a categorical variable with normal birth weight (2500-4000 g) as reference, as a continuous variable with units of 500 and 1000 g, and as a dichotomous variable with cut points at 3500 g and 4000 g. Categories of birth weight for gestational age, small-for-gestational-age (SGA), average-for-gestational-age (AGA), large-for-gestational-age (LGA), were constructed using U.S. national reference for fetal growth by Alexander (1996).[27]

To evaluate dose-response relationships between birth weight and CNS tumors, we obtained odds ratios using a moving window of birth weight with a width of 1000 g. This analysis used overlapping birth weight categories (windows) of 2500 to <3500 g, 3000 to <4000 g, 3500 to <4500 g, 4000 to 5000 g and ≥4500 g with reference category <2500 g, and was adjusted for potential confounders.

Gestational age, maternal and paternal ages were analyzed as categorical variables. Gestational age was calculated based on last menstrual period and date of birth. Birth order ($1^{st}$ versus other), history of maternal pregnancy terminations, maternal conditions/complications during index pregnancy and newborn abnormalities were analyzed as dichotomous variables (yes or no).

The primary analysis method was conditional logistic regression using the one-to-one age and gender matched case-control pairs.[28] We evaluated the association of CNS tumors and variables of interest both unadjusted and adjusted for potential confounders. Results of these analyses were not materially different; hence we present adjusted results only. For adjusted analyses, covariates were chosen based on information about known or potential confounders and model fit statistics; models with the lowest Akaike information criterion **(**AIC) value and lowest number of potential confounders are presented. Father's education and payment source for delivery were used as proxies for socioeconomic status (SES), a potential confounder. For payment source for delivery, governmental programs such as Medicare and Medi-Cal and 'No care' were coded as low SES; private insurance and other sources of payment were coded as middle/high SES. For father's education, ≤12 years of education was considered low SES, 13-17 years as middle SES, and ≥17 years as high SES. We confirmed the absence of unduly influential observations by fitting a variety of models with different subsets of covariates and examining the results for outliers and influence.

Despite the large number of cases and controls, sample sizes for some analyses were reduced due to missing data that was attributable largely to differences in the information collected on birth certificates from year to year. No differences in patterns of missingness were detected between cases and controls. Missing data were multiply imputed using Monte Carlo Markov chain full-data imputation under a missing at random assumption [29, 30] implemented by the MI procedure in SAS 9.2, [31] and analyses were repeated on the multiply imputed data using the MIANALYZE procedure. The imputation model included all variables used in models (except newborn abnormalities and complications/conditions during index pregnancy) and auxiliary variables likely to be correlated with variables of interest (number of pregnancy visits, month prenatal care began, type of birth, planned place of birth, number of ever born children, number of children born alive, and number of children born alive now deceased).

Analyses were conducted using SAS 9.2 statistical software. [31]

The study was approved by University of California, Los Angeles Office for the Protection of Research Subjects and the California Office for the Protection of Human Subjects.

**Results**

A total of 3858 childhood CNS and brain tumor cases that met our inclusion criteria were identified from the California cancer registry. Linkage to birth records was successful for 3308 cases (85.7%), and this constituted the analytic sample. There were 1375 astrocytoma, 434 glioma, 279 ependymoma, 311 primitive neuroectodermal embriogenic tumor (PNET), 394 medulloblastoma and 323 other CNS tumor cases; 192 cases were missing information on type of tumor. Of the 3308 subjects, 1769 were male (53.5%) and 1539 were female (46.5%). The mean age at diagnosis was 5.6 years with a range of 0 to 15.6 years. Table 1 shows other characteristics of study subjects.

Socio-demographic and perinatal characteristics of study subjects by CNS tumor subtype are available as online supplemental material.

In Table 2 we show results of conditional logistic regression analyses for associations of childhood CNS tumors and various perinatal factors.

*Birth weight*

In the analysis of birth weight as a continuous variable, risk increased with increasing weight, with adjusted ORs for total CNS per 500g and 1000 g increase. In categorical analyses, including the one with normal birth weight as reference, adjusted ORs also exceeded unity for higher birth weight babies, but estimates were imprecise. (Table 1)

To address concerns about birth weight misclassification due to chosen cut points and to further examine trend, we obtained odds ratios using a moving window for birth weight categorized as 2500-<3500 g, 3000-<4000 g, 3500-<4500 g, 4000 to 5000 g, ≥4500 g with a reference category <2500 g, with adjustment for potential confounders (Figure 1). No clear trend was detected for the association of birth weight and childhood CNS tumors. Analyses using weight-for-gestational-age categories similarly yielded elevated point estimates but were also consistent with no increased risk.

*Gestational age*

Analysis of gestational age adjusted for birth weight, birth order, mother's age, father's education, child's race, and payment source for delivery, entered in the model as a dichotomous variable or 3-level categorical variable, did not yield evidence of increased risk associated with more advanced gestational age (Table 2).

*Birth order*

No association was detected between birth order and childhood CNS tumors.

*Parental age*

A small increased risk of CNS tumors was observed for 5-year increase in maternal and paternal ages (Table 2). Similar result was found in the categorical analysis but with wider confidence intervals, with highest risk among mothers who were more than 35 year-old.

*Other prenatal factors*

Maternal history of pregnancy terminations, either before or after 20 weeks of gestation, was not associated with risk of CNS tumors in children.

Among maternal complications/conditions during index pregnancy, genital herpes and blood and immune disorders were associated with increased risk and non-sexually transmitted infections were associated with decreased risk (Table 2). Among newborn abnormalities, only CNS abnormalities were associated with higher risk of childhood CNS tumors.

Models with interactions between birth weight and birth order, mother's age, child's race, and source of payment for care were also considered. None of the interaction models showed any important findings (results not presented).

*CNS tumor subtypes*

The large sample size allowed us to perform analyses by CNS tumor subtype, although numbers for some subtypes were small. In Table 3, we present results of adjusted conditional logistic regression analysis for associations of various perinatal factors with subtypes of childhood CNS tumors.

High birth weight was associated with slightly increased risk of astrocytoma and glioma, although confidence intervals were wide. Birth weight-for-gestational age was positively associated with the risk of glioma and negatively associated with the risk of ependimoma. (Table 3)

Advanced maternal age was associated with increased risk of glioma and advanced paternal age was associated with small increase in risk of astrocytoma (Table 3).

There was little evidence of any associations of maternal complications during pregnancy or newborn abnormal conditions with risk of any type of childhood CNS tumors examined. There was a numerically high but imprecisely estimated risk of medulloblastoma associated with maternal blood and immune disorders during pregnancy and numerically low but imprecisely estimated risk of astrocytoma and medulloblastoma with maternal non-sexually transmitted infections (Table 3).

After performing complete case analyses, analyses were repeated using multiply imputed data. Estimates from these two types of analyses were very similar, but more precise for

multiply imputed data. No other important differences were observed for the associations of perinatal factors and childhood CNS tumors between complete case and multiply imputed data analysis (results not presented).

**Discussion**

In our analysis we observed a slightly increased risk for total CNS tumors for high birth weight infants but no clear dose response. These findings are in line with other studies that reported a slightly increased risk of CNS tumors for high birth weight babies, but findings have been inconsistent. Meta-analysis by Harder et al. (2008) found slightly increased risk of astrocytoma and medulloblastoma for high birth weight children.[14] Some studies detected increased risk of astrocytomas only. [12, 13] In analyses by subtype, we observed numerically higher point estimates of risk for astrocytoma and medulloblastoma as well as for other subtypes.

Some authors have suggested that birth weight alone is not a good predictor for the development of childhood cancers and that birth weight for gestational age as an indicator for fetal growth could be a better predictor for brain tumors. When we examined birth weight for gestational age, the results indicated no increase in risk of total CNS tumors for either LGA or SGA babies compared to AGA. In analyses by subtype, a 1.9-fold increased risk of glioma was observed for LGA babies. Interestingly, for ependymoma, we found a 2.6-fold increased risk for SGA babies. Such results for subtype analysis could be due to chance or to different risk factors for different subtypes.

Like the majority of studies, we did not find any association between birth order and either total CNS tumors or subtypes.

Our results revealed a weak positive association of advanced maternal and paternal ages and CNS tumors: 6% increased risk for 5-year increase in maternal age and 5% for 5-year increase in paternal age. In subtype analysis we observed increased risk of astrocytoma for 5-year increase in paternal age and of glioma for paternal age greater than 35 years old (< 25 years old as reference). Although many studies did not find associations with parental age [12, 19, 32, 33], some studies showed results similar to ours. Hemminki et al. [23] found an increased risk of CNS tumors and advanced paternal age; Macmahon et al (1962) found an increased risk with advanced maternal age. [34] Reproductive parental age may affect the risk of childhood cancers in several ways. Frequencies of paternal and maternal germ cell mutations increase with age and

thus may increase the chance of developing cancer in a child. [35] Other mechanisms could include differential expression of genes in cell cycle control and changes in DNA damage response and repair pathways in oocytes of older mothers. [36] Aging may also change physiological parameters, such as estrogens levels, which may also increase risk of cancer. [35] Older age, particularly of the mother, appears to be a risk factor across many of the common childhood cancer types. [36]

There were several maternal complications or conditions during pregnancy that we found potentially associated with CNS tumors and some subtypes. Genital herpes and blood and immunological disorders during index pregnancy were associated with marked increased risk of total CNS tumors (OR=2.74 and OR=2.28, respectively). Increase in risk associated with blood and immune disorders was probably due to a large (about 7-fold) but imprecise increase in risk of medulloblastoma. Several other studies found an increased risk of CNS tumors in children whose mothers had viral infections during pregnancy [4], but there were no studies that investigated genital herpes infection specifically. Viral material has been found in almost all types of CNS tumors. [7, 9] Although the presence of viral substance does not prove a causal relationship with tumor formation, it does suggest the hypothesis that viruses may be involved. [7-9, 37] Interestingly, we observed a decreased risk of CNS tumors in children whose mothers had other non-sexually transmitted infections during pregnancy. It was not possible to identify type of infection during pregnancy; thus our results could be due to bacterial or various combinations of infections.

Newborn CNS abnormalities were associated with 4-fold increased risk of CNS tumors. Other newborn abnormalities such as Down's syndrome, musculoskeletal, digestive system, face and neck abnormalities were not associated with the risk of CNS tumors. We did not find any other studies that looked at the association of CNS/brain tumors and newborn abnormalities except that of Partap et al. (2011), which found an increased risk associated with newborn birth defects (non-specific). [38]

We did not find any association between CNS tumors or their subtypes and other perinatal factors such as gestational age, history of pregnancy terminations, eclampsia or preeclampsia, anemia, chronic conditions, and maternal tobacco smoking. We did not find studies that looked at similar factors in association with CNS tumors, except studies on maternal tobacco smoking that also found no association. [5, 18, 25]

Our study had strengths and limitations. A major strength was that our data were from population registries with almost complete registration of births and cancers in California and controls were randomly selected from the birth registry rather than recruiting volunteers as in many case-control studies. Since these registries are independent of each other and participation of subjects was not required for our data collection, selection and information biases are unlikely in this study. This is not the case for studies based on questionnaires and interviews, which are subject to serious participation and recall bias issues. Misclassification of outcome status was also unlikely in this study due to the completeness and high accuracy of the California Cancer Registry. Misclassification of birth weight was possible but unlikely since birth weight is usually recorded very accurately.[39] There are very few validation studies on quality of reporting for some variables on California birth certificates. Validation studies on accuracy of reporting of health insurance and birth weight has shown high concordance between recorded and interview data, but for other variables, such as gestational age, accuracy was lower, especially after 1992. [39, 40] Bases on afore-mentioned information we assumed that it is possible that some perinatal factors, particularly complications during pregnancy and newborn abnormalities, are subject to some misreporting. Information was not routinely collected for many maternal and newborn conditions prior to 1990 and in addition, the very small numbers of such subjects leads us to suspect underreporting. Uncertain accuracy of reporting together with underreporting could have led to misclassification in those variables, which most likely would have biased results towards the null.

Another advantage of this study was that the large size of the dataset allowed us to carry out analyses for subtypes of CNS tumors.

We adjusted for potential confounders that were available in registries such as ethnicity, parental education, and sources of payment for delivery as proxies for socioeconomic status. Adjustment for the majority of these variables did not make a difference in estimates of risk for birth weight and other perinatal factors. There was little or no information available on such potential confounders as maternal and paternal occupation, diet, alcohol and drug use, maternal health conditions before pregnancy and child's abnormal conditions. Therefore, residual confounding was possible. It was unlikely that maternal occupation and alcohol and drug use could considerably bias the association of birth weight and brain tumors since these factors are inversely associated with birth weight [41-43] but positively associated with childhood CNS

tumors. [13, 24, 44-48] Therefore, even if we considered bias due to these variables possible, it would most likely pull the estimate toward the null.

A limitation of the study was missing data. However, since information was missing mainly due to differences in the information collected on birth certificates from year to year rather than non-response, the potential for biases was probably small, and the impact was mainly on the precision of the estimates. For example, information on parental education and several maternal and newborn conditions was available only for more recent years. There was no difference in the pattern of missingness between cases and controls. We reanalyzed the data using multiple imputation and obtained similar results.

In summary, we found that maternal genital herpes, blood and immunological disorders during pregnancy and newborn CNS abnormalities were associated with increased risk of CNS tumors. Maternal infections during pregnancy were associated with decreased risk. Advanced maternal and paternal ages were also associated with slightly increased risk of CNS tumors. Factors associated with subtypes of CNS tumors varied by subtype, which could be expected due to different etiology for different subtypes.

**Acknowledgements**

This project was supported by a research contract from the Electric Power Research Institute (EPRI) to the UCLA and by UCLA Faculty Grants Program.

Crespi was also partially supported by National Institutes of Health grant P30 CA16042.

## References


1. Gurney, G., Smith, MA, Bunin, GR,, *CNS and Miscellaneous Intracranial and intraspinal neoplasms* in *SEER Pediatric Monograph*, N.C. Institute, Editor 1999, SEER: Bethesda. p. 51-64.
2. National Cancer Institute, *A snapshot of brain and central nervous system cancers*, in *Snapshots*2008, National Cancer Institute, Office of Science Planning and Assessment
3. Ries LAG, S.M., Gurney JG, Linet M, Tamra T, Young JL, Bunin GR., *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995.*, in *SEER Pediatric Monograph*, N.C. Institute, Editor 1999, SEER: Bethesda, MD. p. 51-64.
4. Dickinson, H.O., T.A. Nyari, and L. Parker, *Childhood solid tumours in relation to infections in the community in Cumbria during pregnancy and around the time of birth.* Br J Cancer, 2002. **87**(7): p. 746-50.
5. Fear, N.T., et al., *Malignant neoplasms of the brain during childhood: the role of prenatal and neonatal factors (United Kingdom).* Cancer Causes Control, 2001. **12**(5): p. 443-9.
6. Shaw, A.K., P. Li, and C. Infante-Rivard, *Early infection and risk of childhood brain tumors (Canada).* Cancer Causes Control, 2006. **17**(10): p. 1267-74.
7. Baldwin, R.T. and S. Preston-Martin, *Epidemiology of brain tumors in childhood--a review.* Toxicol Appl Pharmacol, 2004. **199**(2): p. 118-31.
8. Vilchez, R.A. and J.S. Butel, *SV40 in human brain cancers and non-Hodgkin's lymphoma.* Oncogene, 2003. **22**(33): p. 5164-72.
9. Khalili, K., et al., *Human neurotropic polyomavirus, JCV, and its role in carcinogenesis.* Oncogene, 2003. **22**(33): p. 5181-91.
10. Schuz, J., et al., *Risk factors for pediatric tumors of the central nervous system: results from a German population-based case-control study.* Med Pediatr Oncol, 2001. **36**(2): p. 274-82.
11. McCredie, M., et al., *SEARCH international case-control study of childhood brain tumours: role of index pregnancy and birth, and mother's reproductive history.* Paediatr Perinat Epidemiol, 1999. **13**(3): p. 325-41.
12. Von Behren, J. and P. Reynolds, *Birth characteristics and brain cancers in young children.* Int J Epidemiol, 2003. **32**(2): p. 248-56.
13. Kuijten, R.R., et al., *Gestational and familial risk factors for childhood astrocytoma: results of a case-control study.* Cancer Res, 1990. **50**(9): p. 2608-12.
14. Harder, T., A. Plagemann, and A. Harder, *Birth weight and subsequent risk of childhood primary brain tumors: a meta-analysis.* Am J Epidemiol, 2008. **168**(4): p. 366-73.
15. Sprehe, M.R., et al., *Comparison of birth weight corrected for gestational age and birth weight alone in prediction of development of childhood leukemia and central nervous system tumors.* Pediatr Blood Cancer, 2010. **54**(2): p. 242-9.
16. Smith, A., et al., *Birth weight, sex and childhood cancer: A report from the United Kingdom Childhood Cancer Study.* Cancer Epidemiol, 2009. **33**(5): p. 363-7.
17. Heuch, J.M., et al., *Risk of primary childhood brain tumors related to birth characteristics: a Norwegian prospective study.* Int J Cancer, 1998. **77**(4): p. 498-503.
18. Linet, M.S., et al., *Maternal and perinatal risk factors for childhood brain tumors (Sweden).* Cancer Causes Control, 1996. **7**(4): p. 437-48.

19. Emerson, J.C., et al., *Childhood brain tumor risk in relation to birth characteristics.* J Clin Epidemiol, 1991. **44**(11): p. 1159-66.
20. Pavlovic, M.V., et al., *Risk factors from brain tumors in children and adolescents: a case-control study in Belgrade, Serbia.* Eur J Neurol, 2005. **12**(7): p. 508-13.
21. Mallol-Mesnard, N., et al., *Birth characteristics and childhood malignant central nervous sytem tumors: the ESCALE study (French Society for Childhood Cancer).* Cancer Detect Prev, 2008. **32**(1): p. 79-86.
22. Macmahon, B. and V.A. Newill, *Birth characteristics of children dying of malignant neoplasms.* J Natl Cancer Inst, 1962. **28**: p. 231-44.
23. Hemminki, K., P. Kyyronen, and P. Vaittinen, *Parental age as a risk factor of childhood leukemia and brain cancer in offspring.* Epidemiology, 1999. **10**(3): p. 271-5.
24. Cordier, S., et al., *Parental exposure to polycyclic aromatic hydrocarbons and the risk of childhood brain tumors: The SEARCH International Childhood Brain Tumor Study.* Am J Epidemiol, 2004. **159**(12): p. 1109-16.
25. Norman, M.A., et al., *Prenatal exposure to tobacco smoke and childhood brain tumors: results from the United States West Coast childhood brain tumor study.* Cancer Epidemiol Biomarkers Prev, 1996. **5**(2): p. 127-33.
26. The California Cancer Registry. *Mission Statement & Purpose*. March 2009]; Available from: http://www.ccrcal.org/aboutttheccr.html.
27. Alexander, G.R., et al., *A United States national reference for fetal growth.* Obstet Gynecol, 1996. **87**(2): p. 163-8.
28. Hosmer, D.W., Jr. , Lemeshow, S., *Applied Logistic Regression*. 2000, New York: Wiley.
29. Rubin, D.B. and N. Schenker, *Multiple imputation in health-care databases: an overview and some applications.* Stat Med, 1991. **10**(4): p. 585-98.
30. Little, R., Rubin, D., *Statistical Analysis with Missing Data*. 2nd Edition ed. 2002, New York: Wiley.
31. SAS Institute, I., *SAS Statistical Software*, 2006: Cary, NC.
32. Samuelsen, S.O., et al., *Head circumference at birth and risk of brain cancer in childhood: a population-based study.* Lancet Oncol, 2006. **7**(1): p. 39-42.
33. McKinney, P.A., et al., *Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study.* Br J Cancer, 1999. **80**(11): p. 1844-51.
34. Macmahon, B., *Prenatal x-ray exposure and childhood cancer.* J Natl Cancer Inst, 1962. **28**: p. 1173-91.
35. Yip, B.H., Y. Pawitan, and K. Czene, *Parental age and risk of childhood cancers: a population-based cohort study from Sweden.* Int J Epidemiol, 2006. **35**(6): p. 1495-503.
36. Johnson, K.J., et al., *Parental age and risk of childhood cancer: a pooled analysis.* Epidemiology, 2009. **20**(4): p. 475-83.
37. Bondy, M. and B.L. Ligon, *Epidemiology and etiology of intracranial meningiomas: a review.* J Neurooncol, 1996. **29**(3): p. 197-205.
38. Partap, S., et al., *Birth Anomalies and Obstetric History as Risks for Childhood Tumors of the Central Nervous System.* Pediatrics, 2011.
39. Remy, L., Oliva, G, *Impact of data quality on birth related health indicators*, in *Family Health Outcomes Project*2006, University of California San Francisco: San Francisco.
40. Braveman, P., et al., *Validity of insurance information on California birth certificates.* Am J Public Health, 1998. **88**(5): p. 813-6.

41. Bada, H.S., et al., *Low birth weight and preterm births: etiologic fraction attributable to prenatal drug exposure.* J Perinatol, 2005. **25**(10): p. 631-7.
42. Day, N.L., et al., *Prenatal exposure to alcohol: effect on infant growth and morphologic characteristics.* Pediatrics, 1989. **84**(3): p. 536-41.
43. Richardson, G.A., et al., *Growth of infants prenatally exposed to cocaine/crack: comparison of a prenatal care and a no prenatal care sample.* Pediatrics, 1999. **104**(2): p. e18.
44. Cordier, S., et al., *Incidence and risk factors for childhood brain tumors in the Ile de France.* Int J Cancer, 1994. **59**(6): p. 776-82.
45. Cordier, S., et al., *Parental occupations and childhood brain tumors: results of an international case-control study.* Cancer Causes Control, 2001. **12**(9): p. 865-74.
46. McKean-Cowdin, R., et al., *Parental occupation and childhood brain tumors: astroglial and primitive neuroectodermal tumors.* J Occup Environ Med, 1998. **40**(4): p. 332-40.
47. Milne, E., et al., *Parental alcohol consumption and risk of childhood acute lymphoblastic leukemia and brain tumors.* Cancer Causes Control, 2013. **24**(2): p. 391-402.
48. Preston-Martin, S., et al., *N-Nitroso compounds and childhood brain tumors: a case-control study.* Cancer Res, 1982. **42**(12): p. 5240-5.

**Table 1.** Socio-demographic and perinatal characteristics of childhood brain and central nervous system tumor cases and controls, California birth registry, 1986-2007.

| Variables | Total N | Cases N (%) | Controls N (%) |
|---|---|---|---|
| | 6616 | 3308 | 3308 |
| **Child's age at diagnosis** | | | |
| < 1 year | 670 | 335 (10.1) | 335 (10.1) |
| 1-5 years | 3231 | 1609 (48.6) | 1622 (49.0) |
| 6-9 years | 1528 | 768 (23.2) | 760 (23.0) |
| 10-15 years | 1187 | 596 (18.0) | 591 (17.9) |
| **Mother's age at birth** | | | |
| < 25 years | 2261 | 1054 (31.9) | 1207 (36.5) |
| 25-35 years | 3359 | 1711 (51.7) | 1648 (49.8) |
| ≥35 years | 995 | 543 (16.4) | 452 (13.7) |
| Missing | 1 | 0 | 1 |
| **Father's age at birth** | | | |
| < 25 years | 1369 | 635 (20.4) | 734(23.6) |
| 25-35 years | 3251 | 1647 (52.9) | 1604 (51.7) |
| ≥35 years | 1598 | 831 (26.7) | 767 (24.7) |
| Missing | 398 | 195 | 203 |
| **Child's race** | | | |
| White | 5000 | 2564 (81.0) | 2436 (77.8) |
| Black | 555 | 257 (8.2) | 298 (9.5) |
| Asian | 640 | 295 (9.3) | 345 (11.0) |
| Other | 101 | 48 (1.5) | 53 (1.7) |
| Missing | 320 | 144 | 176 |
| **Hispanic origin of child** | | | |
| Non-Hispanic | 3432 | 1846 (57.1) | 1586 (49.2) |
| Hispanic | 3021 | 1385 (42.9) | 1636 (50.8) |
| Missing | 163 | 77 | 86 |
| **Father's education** | | | |
| ≤12 years | 3521 | 1720 (75.0) | 1801 (79.8) |
| 13-17 years | 739 | 416 (18.1) | 323 (14.3) |

| | | | |
|---|---|---|---|
| ≥17 years | 291 | 157 (6.9) | 134 (5.9) |
| Missing/Not collected [1] | 2065 | 1015 | 1050 |
| **Source of payment for delivery** | | | |
| Governmental programs | 2422 | 1075 (38.1) | 1347 (47.8) |
| Other insurance | 3220 | 1748 (61.9) | 1472 (52.2) |
| Missing/Not collected [1] | 974 | 485 | 489 |
| **Birth weight** | | | |
| <2500 g | 408 | 192 (5.8) | 216 (6.5) |
| 2500-3500 g | 3406 | 1655 (50.0) | 1751 (52.9) |
| 3500-4500 g | 2679 | 1399 (42.3) | 1280 (38.7) |
| ≥4500 g | 123 | 62 (1.9) | 61 (1.8) |
| **Gestational age at birth** | | | |
| <37 weeks | 645 | 319 (10.3) | 326 (10.5) |
| 37-42 weeks | 5005 | 2499 (80.8) | 2506 (80.8) |
| >42 weeks | 542 | 274 (8.9) | 268 (8.7) |
| Missing/implausible values | 424 | 216 | 208 |
| **Birth order** | | | |
| First | 2606 | 1331 (40.3) | 1275 (38.6) |
| Other | 4003 | 1973 (59.7) | 2030 (61.4) |
| Missing | 7 | 4 | 3 |
| **Maternal history of terminations before 20 weeks** | | | |
| No | 5508 | 2741 (82.9) | 2767 (83.8) |
| Yes | 1102 | 564 (17.1) | 538 (16.3) |
| Missing | 6 | 3 | 3 |
| **Maternal history of terminations after 20 weeks** | | | |
| No | 6512 | 3252 (98.5) | 3260 (98.6) |
| Yes | 96 | 51 (1.5) | 45 (1.4) |
| Missing | 8 | 5 | 3 |
| **Complications/conditions during index pregnancy**[2] | | | |
| None | 3173 | 1561 (47.4) | 1612 (48.9) |

| | | | |
|---|---|---|---|
| Preeclampsia/eclampsia | 93 | 45 (1.4) | 48 (1.5) |
| Anemia | 38 | 19 (0.6) | 19 (0.6) |
| Genital herpes | 75 | 50 (1.5) | 25 (0.8) |
| Infections (non-sexually transmitted)[3] | 33 | 11 (0.3) | 22 (0.7) |
| Chronic diseases[3] | 145 | 84 (2.6) | 61 (1.8) |
| Blood and immune disorder[3] | 51 | 32 (1.0) | 19 (0.6) |
| Tobacco use | 122 | 62 (1.9) | 60 (1.8) |
| **Newborn abnormal conditions** [2] | | | |
| CNS | 18 | 15 (0.5) | 3 (0.1) |
| Eye, ear, neck, face | 9 | 3 (0.1) | 6 (0.2) |
| Musculoskeletal | 11 | 7 (0.2) | 4 (0.1) |
| Down's syndrome | 12 | 8 (0.2) | 4 (0.1) |

[1] Patterns of missingness varied by year due to differences in data collection for different years.

[2] Information was not routinely collected for many of these conditions for years prior to 1990; therefore there is substantial amount of missing data for these conditions.

[3] Infections during pregnancy include pyelonephritis, hepatitis B and rubella.
Chronic diseases include renal, cardiac, lung diseases, hypertension and diabetes.
Blood and immune disorders include Rh sensitization, hemaglobinopathy, uterine bleeding before labor.

**Figure 1**. Conditional odds ratios (OR) and 95% confidence intervals (CI) for childhood CNS tumors at moving windows of birth weight, matched on child's age and sex and adjusted for child's race, gestational age, mother's age, birth order, father's education and source of payment for delivery. Reference level: < 2500 g.

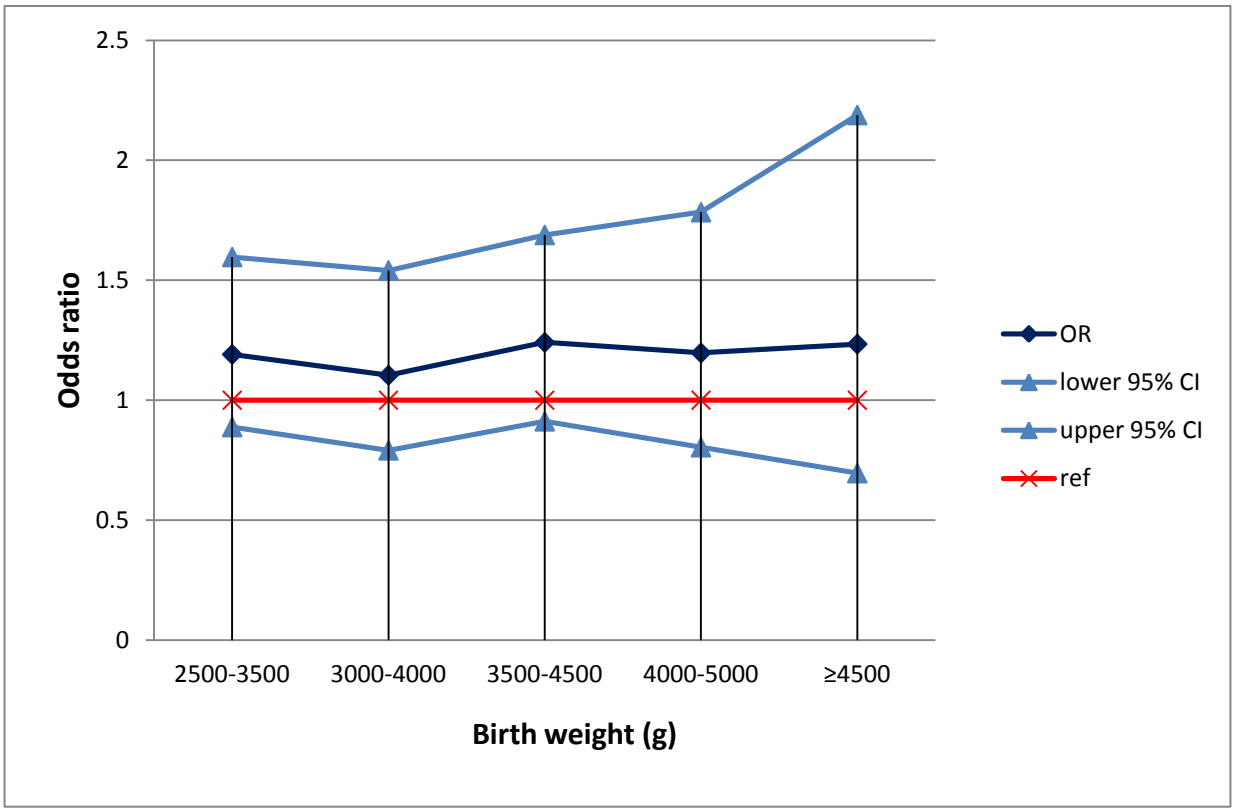


**Table 2.** Adjusted conditional odds ratios (OR) and 95% confidence intervals for childhood CNS tumors and various risk factors, matched on age and sex. California birth registry, 1986-2007.

| Variables | Adjusted OR | Lower 95% confidence interval | Upper 95% confidence interval |
|---|---|---|---|
| **Birth weight** [1] | | | |
| Continuous, per 500 g increase | 1.06 | 0.99 | 1.12 |
| Continuous, per 1000 g increase | 1.11 | 0.99 | 1.24 |
| <2500 g | 1 | - | - |
| 2500-3500 g | 1.19 | 0.89 | 1.60 |
| 3500-4500 g | 1.24 | 0.91 | 1.69 |
| ≥4500 g | 1.24 | 0.70 | 2.19 |
| <3500 g | 1 | - | - |
| ≥3500 g | 1.06 | 0.92 | 1.22 |
| <4000 g | 1 | - | - |
| ≥4000 g | 1.12 | 0.91 | 1.39 |
| <2500 g | 0.84 | 0.63 | 1.12 |
| 2500-4000g | 1 | - | - |
| ≥4000 g | 1.12 | 0.91 | 1.38 |
| SGA | 0.96 | 0.75 | 1.23 |
| AGA | 1 | - | - |
| LGA | 1.09 | 0.89 | 1.27 |
| **Gestational age** [2] | | | |
| <37 weeks | 1 | - | - |
| 37-42 weeks | 0.96 | 0.74 | 1.23 |
| ≥42 weeks | 1.08 | 0.76 | 1.52 |
| **Birth order** [3] | | | |
| First | 1 | - | - |
| Other | 0.92 | 0.79 | 1.06 |
| **Mother's age** [4] | | | |
| by 5-year increase | 1.06 | 1.00 | 1.12 |

| | | | |
|---|---|---|---|
| < 25 years | 1 | - | - |
| 25-35 years vs < 25 years | 1.12 | 0.96 | 1.30 |
| ≥35 years vs < 25 years | 1.21 | 0.98 | 1.49 |
| **Father's age** [4] | | | |
| by 5-year increase | 1.05 | 1.00 | 1.10 |
| < 25 years | 1 | - | - |
| 25-35 years vs < 25 years | 1.18 | 0.99 | 1.41 |
| ≥35 years vs < 25 years | 1.16 | 0.94 | 1.43 |
| **Maternal history of terminations before 20 weeks** [1] | | | |
| No | 1 | - | - |
| Yes | 1.13 | 0.94 | 1.35 |
| **Maternal history of terminations after 20 weeks** [1] | | | |
| No | 1 | - | - |
| Yes | 1.28 | 0.71 | 2.32 |
| **Complications/conditions during index pregnancy** [1] | | | |
| Preeclampsia/Eclampsia | 0.78 | 0.42 | 1.41 |
| Anemia | 1.04 | 0.48 | 2.29 |
| Genital herpes | 2.79 | 1.17 | 6.62 |
| Infections (non-sexually transmitted) | 0.28 | 0.09 | 0.85 |
| Chronic diseases | 1.31 | 0.84 | 2.03 |
| Blood and immune disorders | 2.32 | 1.09 | 4.91 |
| Tobacco use | 1.86 | 0.73 | 1.91 |
| **Newborn abnormal conditions** [1] | | | |
| CNS | 4.30 | 1.19 | 15.53 |
| Eye, ear, neck, face | 0.42 | 0.08 | 2.30 |
| Musculoskeletal | 1.12 | 0.30 | 4.24 |
| Down's syndrome | 2.06 | 0.50 | 8.47 |

[1] Adjusted for child's race, gestational age, birth order, mother's age, father's education and source of payment for delivery

[2] Adjusted for child's race, birth weight, birth order, mother's age, father's education, source of payment for delivery

[3] Adjusted for child's race, gestational age, father's education, mother's age, source of payment for delivery

[4] Adjusted for child's race, birth order, father's education, and source of payment for delivery

**Table 3.** Adjusted conditional odds ratios (OR) and 95% confidence intervals (CI) for childhood CNS tumors by subtype and several perinatal factors, matched on child's age and sex. California birth registry, 1986-2007.

| | Astrocytoma | | | Glioma | | | Ependymoma | | | PNET | | | Medulloblastoma | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Birth weight** [1] | | | | | | | | | | | | | | | |
| by 500 g increase | 1.07 | 0.97 | 1.19 | 1.10 | 0.92 | 1.32 | 0.95 | 0.77 | 1.17 | 1.11 | 0.91 | 1.36 | 1.12 | 0.92 | 1.35 |
| by 1000 g increase | 1.17 | 0.97 | 1.42 | 1.23 | 0.88 | 1.72 | 0.94 | 0.65 | 1.37 | 1.23 | 0.83 | 1. 82 | 1.28 | 0.90 | 1.83 |
| <2500 g | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 2500-3500 g | 1.12 | 0.68 | 1.83 | 1.93 | 0.86 | 4.35 | 1.58 | 0.47 | 5.30 | 0.98 | 0.38 | 2.54 | 0.83 | 0.30 | 2.34 |
| 3500-4500 g | 1.26 | 0.75 | 2.11 | 1.54 | 0.865 | 3.68 | 0.91 | 0.26 | 3.22 | 1.11 | 0.41 | 2.98 | 1.12 | 0.39 | 3.21 |
| ≥4500 g | 0.99 | 0.40 | 2.48 | 1.94 | 0.30 | 12.71 | n/e* | n/e | n/e | 1.82 | 0.35 | 9.48 | 0.52 | 0.10 | 2.74 |
| ≥3,500 g vs <3,500 g | 1.11 | 0.88 | 1.40 | 0.85 | 0.53 | 1.37 | 0.66 | 0.39 | 1.10 | 1.20 | 0.74 | 1.95 | 1.26 | 0.81 | 1.96 |
| ≥4,000 g vs <4,000 g | 1.09 | 0.79 | 1.51 | 1.80 | 0.93 | 3.48 | 1.58 | 0.75 | 3.31 | 1.01 | 0.44 | 2.35 | 1.06 | 0.58 | 1.93 |
| <2500 g | 0.87 | 0.53 | 1.41 | 0.55 | 0.25 | 1.23 | 0.78 | 0.24 | 2.57 | 0.96 | 0.37 | 2.45 | 1.08 | 0.39 | 2.95 |
| 2500-4000g | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| ≥4000 g | 1.09 | 0.79 | 1.50 | 1.80 | 0.93 | 3.50 | 1.55 | 0.74 | 3.27 | 1.01 | 0.44 | 2.35 | 1.06 | 0.58 | 1.94 |
| SGA | 0.75 | 0.50 | 1.13 | 0.95 | 0.44 | 2.06 | 2.64 | 1.10 | 6.30 | 0.63 | 0.29 | 1.39 | 1.91 | 0.73 | 5.03 |
| AGA | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| LGA | 0.96 | 0.70 | 1.33 | 1.86 | 0.99 | 3.48 | 1.41 | 0.66 | 3.04 | 0.93 | 0.45 | 1.91 | 1.22 | 0.68 | 2.19 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Maternal age** [4] | | | | | | | | | | | | | | | |
| Mother's age (by 5-year increase) | 1.07 | 0.98 | 1.17 | 1.12 | 0.95 | 1.33 | 1.06 | 0.88 | 1.28 | 0.89 | 0.73 | 1.09 | 1.13 | 0.94 | 1.36 |
| < 25 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 25-35 years vs < 25 years | 1.21 | 0.95 | 1.54 | 1.16 | 0.72 | 1.87 | 1.15 | 0.71 | 1.87 | 0.80 | 0.49 | 1.31 | 1.14 | 0.71 | 1.83 |
| ≥35 years vs < 25 years | 1.09 | 0.79 | 1.52 | 1.87 | 1.00 | 3.52 | 0.92 | 0.45 | 1.87 | 0.86 | 0.40 | 1.84 | 1.60 | 0.79 | 3.21 |
| **Paternal age** [4] | | | | | | | | | | | | | | | |
| Father's age (by 5-year increase) | 1.08 | 1.00 | 1.16 | 1.10 | 0.97 | 1.26 | 1.11 | 0.95 | 1.31 | 0.89 | 0.76 | 1.04 | 1.08 | 0.93 | 1.25 |
| < 25 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 25-35 years vs < 25 years | 1.24 | 0.94 | 1.64 | 1.45 | 0.83 | 2.53 | 1.06 | 0.57 | 1.99 | 1.11 | 0.63 | 1.94 | 0.89 | 0.53 | 1.50 |
| ≥35 years vs < 25 years | 1.33 | 0.96 | 1.84 | 1.63 | 0.88 | 3.02 | 1.54 | 0.71 | 3.34 | 0.49 | 0.24 | 1.03 | 0.99 | 0.51 | 1.93 |
| **Maternal complications/conditions during index pregnancy** [3] | | | | | | | | | | | | | | | |
| Eclampsia | 0.55 | 0.21 | 1.40 | 0.82 | 0.11 | 6.27 | * | * | * | * | * | * | 0.95 | 0.06 | 15.93 |
| Anemia | 2.09 | 0.52 | 8.41 | * | * | * | * | * | * | 3.86 | 0.36 | 41.67 | 0.74 | 0.10 | 5.32 |
| Genital herpes | 2.25 | 0.44 | 11.43 | 3.41 | 0.37 | 31.79 | * | * | * | 0.78 | 0.05 | 13.18 | 1.51 | 0.26 | 8.64 |
| Infections (non- sexually transmitted) | 0.38 | 0.07 | 2.01 | * | * | * | * | * | * | ** | ** | ** | 0.59 | 0.05 | 6.99 |
| Chronic conditions | 1.94 | 0.91 | 4.12 | 1.70 | 0.45 | 6.46 | 3.08 | 0.32 | 29.39 | 0.98 | 0.20 | 4.91 | 0.41 | 0.10 | 1.61 |
| Blood disorders | 1.14 | 0.37 | 3.51 | 1.73 | 0.24 | 12.69 | ** | ** | ** | * | * | * | 8.28 | 0.96 | 71.50 |
| Tobacco use | 1.79 | 0.79 | 4.02 | 2.10 | 0.49 | 9.02 | * | * | * | 1.03 | 0.28 | 3.79 | 0.84 | 0.20 | 3.43 |
| **Newborn abnormal conditions** [3] | | | | | | | | | | | | | | | |
| CNS | 0.72 | 0.04 | 13.15 | * | * | * | ** | ** | ** | * | * | * | * | * | * |
| Eye, ear, neck, face | * | * | * | ** | ** | ** | ** | ** | ** | * | * | * | * | * | * |
| Musculoskeletal | 0.25 | 0.02 | 2.73 | ** | ** | ** | ** | ** | ** | * | * | * | * | * | * |
| Down's syndrome | * | * | * | ** | ** | ** | ** | ** | ** | 1.75 | 0.14 | 22.37 | * | * | * |

[1] Adjusted for child's race, gestational age, birth order, mother's age, father's education and source of payment for delivery
[2] Adjusted for child's race, gestational age, father's education, mother's age, source of payment for delivery

[3] Adjusted for child's race, birth weight, gestational age, birth order, mother's age, father's education, source of payment for delivery
[4] Adjusted for child's race, birth order, father's education, and source of payment for delivery
* Not estimable due to small cell counts
** No observations

# Chapter 5

## Race/ethnicity and childhood leukemia: a case-control study in California

Oksuzyan S[1], Cockburn M[2], Crespi CM[3], Mezei G[4], Vergara X[5], Kheifets L[6]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, soksuzyan@ucla.edu; [2]Department of Preventive Medicine, University of Southern California, Los Angeles, California, myles@med.usc.edu; [3]Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California, ccrespi@ucla.edu; [4]Environment Department, Electric Power Research Institute, Palo Alto, California, gmezei@epri.com; [5]Environment Department, Electric Power Research Institute, Palo Alto, California, xvergara@epri.com; [6]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, kheifets@ucla.edu.

**Key words:** childhood leukemia, child's race, parental race, child's Hispanic ethnicity, parental Hispanic ethnicity

**Corresponding author:** Sona Oksuzyan, MD, MPH, PhD (c)
Mailing address: 17072 Signature Drive, Granada Hills, CA 91344
Phone: 818-731-5854
Email: soksuzyan@ucla.edu

**Abstract**

*Aims:* We conducted a large registry-based study in California to investigate the association between race/ethnicity and childhood leukemia, focusing on two subtypes: acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML).

*Methods:* We obtained information on 5788 cases and 5788 controls by linking California cancer and birth registries. We evaluated relative risk of childhood leukemia by race and ethnicity of the child and their parents using conditional logistic regression, with adjustment for potential confounders.

*Results:* Compared to Whites, Black children had lower risk of ALL (odds ratio [OR], 0.49, 95% CI: 0.37-0.65). Children of Black/Black and Black/Asian parents were also at decreased risk of ALL (OR=0.40, 95% CI: 0.30-0.52 and OR=0.23, 95% CI: 0.07-0.80, respectively). Hispanic ethnicity was associated with increased risk of ALL (OR=1.36, 95% CI: 1.22-1.52). Children with one Hispanic parent were at higher risk of ALL (OR=1.19, 95% CI: 1.00-1.42); for children with two Hispanic parents, the risk increased further (OR=1.42, 95% CI: 1.26-1.60). The highest risk of ALL was observed for Hispanic White children compared to non-Hispanic Whites (OR=1.28, 95% CI: 1.13-1.46). The lowest risk of ALL was observed for non-Hispanic Blacks (OR=0.48, 95% CI: 0.37-0.62). Asian and Black races were associated with increased risk of AML compared to Whites but estimates were imprecise. Associations for total childhood leukemia were similar to ALL.

*Conclusions*: Our results confirm that there are ethnic and racial differences in the incidence of childhood leukemia. These differences indicate that some genetic and/or environmental/cultural factors are involved in etiology of childhood leukemia.

## Introduction

Leukemia accounts for the largest number of cases of childhood cancer in the United States and worldwide.(1) Risk factors for this disease are mostly unknown.

A limited number of studies have specifically examined race and/or ethnicity in relation to childhood leukemia risk. (2-4) Most studies of childhood leukemia have considered race and/or Hispanic ethnicity only as a covariate or a confounder in their analyses of childhood leukemia. One interview-based study found that the proportion of Whites among controls was higher than among cases of childhood leukemia, but this may have been due to selection bias. (5) Another study did not find any association between child's race and childhood leukemia, but may have been underpowered due to small sample size. (6) A vast majority of studies have found that Black race was associated with decreased risk of childhood leukemia (2, 3, 7-12) compared to Whites. The definition of race and ethnicity differed in all these studies.

Fewer studies have looked at relationships between race and ethnicity and risk of major subtypes of childhood leukemia, acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML). Most of these studies detected similar results for total childhood leukemia and ALL, which comprises the majority of childhood leukemia cases (2, 3, 8, 13-15). Two studies found no association with any subtypes (6, 15). For the association of AML and race, results varied: two studies found that White race was associated with increased risk of AML, one study detected higher risk for Asian children (8) and other did not find any association of AML with race.(14) Such varied findings, particularly for subtypes, could be explained by differences in etiology of ALL and AML, underlying socio-economic status, differences in age distribution, and/or differences in classification of race and ethnicity.(16)

Most studies on Hispanic origin of the child have report an increased risk of total childhood leukemia and ALL for Hispanic ethnicity (2, 13, 17, 18); however, some studies found no association between Hispanic origin of the child and the risk of childhood leukemia. (6, 19). In California, Hispanics have been reported to have the highest age-adjusted incidence rates for childhood leukemia (5.62 cases per 100,000), followed by non-Hispanic whites (4.46 cases per 100,000); the lowest incidence rates were among non-Hispanic blacks (2.91 cases per 100,000). (18)

The aims of this large-scale study were to examine the relationships between race/ethnicity of the child, mother and father and childhood leukemia and its subtypes in California. California is particularly suitable for studying racial and ethnic differences in the incidence of childhood leukemia due to its diverse racial/ethnic distribution allowing examination of mixed races. In addition to being one of the most diverse states in the U.S. (20), California has cancer and birth registries that have almost complete (99%) registration. (21, 22) Registry-based studies have an important advantage over interview- or questionnaire-based studies as they have lower potential for selection bias.(23)

Most previous record-based studies used a single definition of child's race and ethnicity. Race/ethnicity of child available from a registry may have sizeable amount of missing data. In our study we explored several definitions of child's race and ethnicity and their combination available from birth registry. We reconstructed child's race and ethnicity from mother's and father's race and ethnicity, hence reducing missing data and increasing sample size.

**Materials and methods**

Eligible cases included in this analysis were childhood leukemia cases diagnosed between 1988 and 2008 in children younger than 16 years who were born in California and who resided in California at the time of diagnosis. Information about cases was extracted from the population-based California Cancer Registry, which routinely records age, race/ethnicity, sex, and residence at the time of diagnosis as well as information on cancer types and characteristics. Cancer registry data were linked to the California Birth Registry, which was used to select controls for 1986-2007 birth years. Controls were selected randomly and matched to cases (1 to 1) on date of birth (±6 months) and sex. The California Birth Registry also provided information on socio-demographic (race/ethnicity of child and parents, parental education and age) and perinatal factors (birth weight, birth order) for cases and controls.

Measures

Race and ethnicity of a child was available in both cancer and birth registries, but in cancer registry it was available for cases only; maternal and paternal race was not available to us from the Cancer registry. Therefore, we used birth records to extract information on child's and parental race and ethnicity. Although racial origin of the child was available in California birth records, more than 30% of values were missing. Significantly fewer subjects were missing information on mother's and father's race, less than 1% and about 4%, respectively (see Table 1). We therefore created classifications of child's race based on race of birth parents as recorded in the child's birth certificate. Racial groups available from child's birth records for mother and father were as follows: White, Black, American Indian, Asian-unspecified and Asian-specified (not included in the standard list), Chinese, Japanese, Korean, Vietnamese, Cambodian, Thai, Laotian, Indian, Filipino, Guamanian, Samoan, Eskimo, Aleut, Pacific Islander, Hawaiian, and

Other. These groups were combined into four main racial categories as follows: White (White), Black (Black), Asian (Asian-unspecified and Asian-specified, Chinese, Japanese, Korean, Vietnamese, Cambodian, Thai, Laotian, Indian, & Filipino) and Other (all others). These four categories were used in analysis for maternal and paternal race.

We used several alternative approaches to create classifications of child's race based on afore-mentioned four categories of mother's and father's races. For comparison in some analyses we used the original child's race available from birth records categorized into four main racial groups. The only difference between our classification and the Surveillance, Epidemiology and End Results (SEER) recommendations was that we classified Pacific Islanders and American Indians into Other racial category. By the first classification of reconstructed child's race with four categories, a child was considered White if both parents were White; Black if either parent was Black; Asian if both parents were Asian or if one of the parents was White and another was Asian; Other if both parents were of Other race or mixed Other/White or Other/Asian race. By the second reconstructed classification with five categories a child was considered the race if both parents were that particular race: White if both parents were White, Black if both parents were Black, Asian if both parents were Asian, Other if both parents were from Other racial groups, and Mixed if parents were from two or more different racial groups (e.g. White/Black, White/Asian, Asian/Other and etc.). We compared results using these two classifications of reconstructed child's race. We found very similar results; in addition, the distribution of race using the first classification was very close to that of the original child's race available from birth certificates. Therefore, we present results based only on the first classification of reconstructed child's race with four categories (White, Black, Asian and Other).

Similar to a system used by Chow et al. (19), we also constructed a 10-category variable for child's race consisting of the following combinations of parental races: White/White, Black/Black, Asian/Asian, Other/Other, White/Black, White/Asian, White/Other, Black/Asian, Black/Other, Asian/Other. The distribution of this combined parental race variable is presented in the Table 1.

Parental Hispanic ethnic origin was considered as a dichotomous variable separate from racial categorization: Hispanic (Mexican, Puerto Rican, Cuban, Central/South American, Other Spanish/Hispanic) and non-Hispanic. Hispanic origin of child was not recorded and was derived from parental Hispanic ethnicity available in birth registry. We classified a child as Hispanic if either parent was Hispanic, regardless of race. Analyses were also conducted that classified children of Hispanic ethnicity as having one or both parents Hispanic.

For some analyses, we combined child's race and Hispanic ethnicity to create one variable with 8 categories: Non-Hispanic White, Hispanic White, Non-Hispanic Black, Hispanic Black, Non-Hispanic Asian, Hispanic Asian, Non-Hispanic Other and Hispanic Other.

We used parental education as a proxy for socio-economic status (SES) at the individual level. Maternal educational attainment (in years) was only collected in certain years and missing in more than 50% of subjects; therefore, we used father's educational attainment as the SES proxy. It was categorized into four levels: <12 years (less than high school), 12 years (high school), 13-16 years (some college, college), and 17 years and more (graduate school).

We used a measure of community-based SES derived from U.S. Census data using principal components analysis based on seven indicator variables at a census block group level:

education index developed by Liu et al. (24) that weights the proportion of people with a given level of education by a number of years need to attain that level, proportion with a blue-collar job, proportion older than 16 without a job, median household income, proportion below 200% of the federal poverty level, median rent, median home value. Available to us were component scores categorized into quintiles from principal component analysis for SES index. (25) We adjusted models for census-based SES and compared the results with models adjusted for individual level SES (i.e. father's education).

Statistical Analysis

The primary analysis method was conditional logistic regression utilizing the matched case-control pairs. (26) We conducted both unadjusted analyses and analyses adjusted for birth order, birth weight, mother's age and father's education. Since these were not materially different, we present results of adjusted analysis. Several models using different subsets of covariates were fitted and checked for potentially influential observations. For adjusted analysis, the models for race and ethnicity were chosen based on information on known or potential confounders and model fit statistics; the most parsimonious models with the lowest Akaike information criterion **(**AIC) values are presented. Models for child, maternal and paternal race and ethnicity were fit separately to avoid close correlations.

Since risk factors for the two main subtypes of childhood leukemia (ALL and AML) could be different, analyses were also conducted for these subtypes separately using the same models.

Despite the large number of cases and controls, sample sizes for some analyses were reduced due to missing data. Due to differences in data collection by year the pattern of

missingness varied by year but no differences in patterns of missingness were detected between cases and controls. Under a missing at random assumption, multivariate imputation techniques were used to impute missing values for father's and mother's education, payment source for delivery and for census-based SES by the MI procedure in SAS. (27-29) The imputation models included all variables used in that models. Analyses were repeated using the multiply imputed data using the MIANALYZE procedure.(27)

Analyses were conducted using SAS 9.3. (27)

The study was approved by University of California, Los Angeles Office for the Protection of Research Subjects.

**Results**

A total of 6645 childhood leukemia cases were identified from the California Cancer Registry. Linkage to birth records was successful for 87.1% (5788/6645) of cases. Of the 5788 cases (55.8% males and 44.2% females) included in this analysis, 4721 were ALL cases (56.2% males and 43.8% females), 852 were AML cases (53.3% males and 46.7% females), and 215 were other childhood leukemia types. The median age at diagnosis was 3.8 years with a range of 0 to 15.4 years with the peak for ALL between 2 and 5 years of age and between 0 and 2 years of age for AML. Table 2 shows other characteristics of study subjects by case/control status and by the two main subtypes.

Results of conditional logistic regression analyses assessing the association of childhood leukemia and racial origin of parents and of child (both original and reconstructed) presented in the Table 3 indicate that Black race of mother, of father and of child were each associated with a

decreased risk of childhood leukemia compared to White race. Similar findings were observed for ALL, but not for AML. For AML, increased risk was observed for Black children, Asian children, children of Asian fathers but estimates were not very precise.

Adjusted analysis of a child's race as defined by father's and mother's races combined showed that Black/Black and Black/Asian race children were at lower risk of total childhood leukemia and ALL compared to White/White children. Black/Black and Asian/Asian children had about 49% and 42%, respectively, increased risk of AML but estimates were imprecise (Table 3).

We repeated the analyses adjusting for census-based SES instead of father's education as the SES proxy. Results were very similar with two exceptions: the association of Black/Asian race was less precise and the association of Asian race with AML was more precise. Children of Asian fathers had higher risk of AML compared to children of White fathers (OR=1.75, 95% CI: 1.13-2.72). Asian children had higher risk of AML in all analyses (OR=1.64, 95% CI: 1.09-2.45 for Asian vs. Whites and OR=1.67, 95% CI: 1.03-2.69 for Asian/Asian vs. White/White children).

Hispanic ethnicity was first considered independent of race. Analysis revealed that Hispanic origin of mother, father and child was associated with increased risk of total childhood leukemia and ALL, but not AML. Since estimates and 95% confidence intervals for mother's and father's Hispanic ethnicity were almost identical to child's Hispanic ethnicity, in Table 4 we presented results for child's Hispanic ethnicity only.

We observed a gradient in the risk of total childhood leukemia and ALL when we compared Hispanic children with both parents Hispanic, one parent Hispanic and non-Hispanic children, with the highest risk observed for children with both parents Hispanic (positive linear trend p-value < 0.0001).

Models with interactions between child's race and birth order, mother's age, father's education (or census-based SES), and Hispanic ethnicity of child were also considered. Interactions were detected between child's race and Hispanic ethnicity (p-value=0.02). For further investigation of this association, we combined child's race and Hispanic ethnicity. Results of regression with combined race/ethnicity variable and childhood leukemia are presented in Table 5.

The highest OR for total childhood leukemia and for ALL was observed in Hispanic White and the lowest in non-Hispanic Black children compared to non-Hispanic Whites. non-Hispanic Asians were at slightly increase risk of total leukemia, ALL and AML, but with imprecise estimates.

After performing complete case analyses, analyses were repeated using multiply imputed data. Results were very similar to complete case analysis. Some of these results are presented in Appendix Table 6 online.

**Discussion**

We examined the relationships between childhood leukemia and parental and child's race and found that compared to White race, being Black was highly protective for the development of childhood leukemia and in particular acute lymphoblastic leukemia (ALL). This association was observed for both paternal and maternal race and for child's race, regardless classification used. Although there were several studies that observed similar results (2, 3, 7-10, 13, 19, 30), our study had larger sample size, we used several classification for child race and we looked at maternal and paternal race too.

Some researchers have suggested underlying socio-economic status as a possible explanation of these associations (16, 31) whereas others suggest the associations may be explained low birth weight among Blacks.(32-35) Many studies have shown that high birth weight was associated with increase in risk of total childhood leukemia and ALL (36, 37); and consequently, low birth weight could have a protective effect on incidence of childhood leukemia and ALL. However, since we controlled for SES using two types of proxies and for birth weight in our study, these factors are unlikely to account for the observations.

We observed an association between Asian race and AML when we adjusted for census-based SES. Elevated risk of AML for Asian children was less pronounced and less precise after adjustment for father's education. The association of AML with Asian race was also observed by Reynolds et al (2002), unadjusted, and the association became imprecise when adjusted for father's education; they did not control for community-level SES.(8) We cannot offer any specific explanation of this findings, but as noted by many researchers, AML may have different risk factors than ALL, including genetic factors.

We observed that Hispanic ethnicity of father, mother and child was associated with a 29% increase in risk of total childhood leukemia and with a more than 37% increase in risk for ALL. We observed a trend in the risk of total childhood leukemia and ALL when we compared Hispanic children with both parents Hispanic, one parent Hispanic and non-Hispanic children, with the highest risk observed for children with both parents Hispanic. We detected an interaction between child's race and ethnicity. Analysis with combined race and Hispanic ethnicity showed that non-Hispanic Black children had the lowest risk of childhood leukemia compared to White non-Hispanics. The highest risk was observed in White Hispanic children. Our findings are in line with results of other studies showing increased risk of childhood leukemia for children of Hispanic origin. (18) The incidence in several Latin American countries is the highest in the world: 5.65 cases per 100,000 in Costa Rica, 5.54 per 100,000 in Ecuador, 4.43 in Uruguay. (38)

Some researchers have suggested that nutrition and diet could contribute to racial and ethnic differences in cancer incidence.(39) Emerging evidence suggests that genetic risk factors may also explain the markedly different risk of childhood leukemia in Hispanics and Blacks. A recent study by Xu (2013) detected new susceptibility variants at 10p12.31-12.2 of the BMI1-PIP4K2A gene. This polymorphism is common in Hispanic ethnic groups and rare in Black populations; it could at least partially explain observed associations.(40)

A major strength of our study was that data were obtained from nearly complete population registries of births and cancers in California and controls were randomly selected from the birth registry rather than recruiting volunteers as in many case-control studies. Interview- and questionnaire-based case-control studies are prone to selection bias. Meta-analysis by Slusky et al. (2012) has shown that interview-, questionnaire-based case-control

studies of childhood cancer could suffer from overrepresentation of Whites and under-representation of other races in participating controls. (23) This is not a case for registry-based non-contact studies. Since birth and cancer registries were independent of each other and participation of subjects was not required for our data collection, selection bias due to participation was unlikely in this study.

A potential bias could arise if controls of a particular race or ethnicity selectively moved out of state before their "pseudo-diagnosis" date and became cases in other states or countries. For example, if Hispanic children moved out of state and became leukemia cases outside of California, we would underestimate that association. These scenarios are very unlikely since probability of controls moving out of California and subsequently developing childhood leukemia, a rare cancer, is quite low. We tested this assumption by randomly selecting different number of Hispanic control and treating them as "cases" in a regression model. This analysis showed that we would get the same OR and 95% CIs if hypothetical twenty three Hispanic controls moved out of state and became leukemia cases there. Only if more than these hypothetical twenty three Hispanic controls moved out of state and became cases out-of-state we would get ORs and 95% CIs higher than the original one. The probability of having twenty three and more of our controls of Hispanic ethnicity moving out of California and after that getting childhood leukemia is extremely low. In addition, the literature supports that cases of childhood leukemia are more mobile than controls and not vice versa.(5, 41, 42)

We matched on age and gender, which are known confounders, and adjusted for other potential confounders available in registries, although we cannot completely exclude residual confounding due to other unknown or unmeasured factors.

Another advantage of this study was that the large size of the dataset allowed us to carry out analyses for two main subtypes of childhood leukemia, ALL and AML. The risk pattern for ALL was very similar to total childhood leukemia. The risk patterns observed for AML were different, and included higher risk for AML for Asian children. This highlights the importance of conducting disaggregated analyses by subtype, as different subtypes are likely to exhibit different patterns of risk factors, which are disguised when subtypes are aggregated.

A study limitation is potential misclassification on variables of interest and covariates. Misclassification of the outcome was unlikely due to the completeness and high accuracy of the California Cancer Registry. Race and ethnicity on birth certificates could be reported by parents, abstracted from a medical chart, or recorded by hospital staff based on their own observations. (19) Nonetheless, in validation studies in California where birth certificate data were compared with structured post-partum interviews, the sensitivity of birth records to correctly identify most racial and ethnic groups was greater than 94% with the exception of Native Americans. (43) Some misclassification of race and ethnicity is still possible due to categorization of these variables. We attempted to address potential misclassification of race and ethnicity by examining sensitivity to re-classification, which, reassuringly, did not alter our results. Even if misclassification was present, we believe that it was not different for cases and controls. Non-differential misclassification would pull the estimates toward the null in the case of binary categorization. (44) Father's education was not consistently recorded on birth certificates for earlier years and we could not estimate its accuracy, but we believe that misclassification of father's education was not differential for cases and controls.

Another weakness of the study was missing data. To at least partially address the issue of missing values, instead of only using records containing child's race available from birth registry

that was missing for about a third of subjects, we constructed a classification of child's race using more complete information about mother's and father's races. For a majority of factors, missingness did not vary considerably by race and ethnicity. Black and Other races of children had slightly higher missing data on father's education than Whites and Asians. However, since information was missing mainly due to differences between years in the collected information on birth certificates rather than non-response, and did not differ between cases and controls, the potential for biases was probably small, and the impact was mainly on the precision of the estimates. We re-analyzed the data using multiple imputations and, as expected, obtained very similar results with slightly narrower confidence intervals.

In summary, we found that children of Black race were at lower risk of childhood leukemia and ALL, but not AML. Hispanic ethnicity was associated with high risk of total childhood leukemia and ALL. The highest risk of childhood leukemia was observed for Hispanic White children and the lowest risk for non-Hispanic Black children. Asian race was associated with increased risk of AML. Ethnic and racial differences in incidence of childhood leukemia indicate that some genetic and environmental/cultural factors may be involved in etiology of childhood leukemia.

**Acknowledgements**

This project was supported by a research contract from the Electric Power Research Institute to UCLA and by UCLA Faculty Grants Program. Crespi was also partially supported by National Institutes of Health grant P30 CA16042. The authors want to thank California Department of Public Health for providing support and access to birth regirstry.

**References:**


1. Ma SK, Wan TS, Chan LC. Cytogenetics and molecular genetics of childhood leukemia. Hematol Oncol. 1999 Sep;17(3):91-105. PubMed PMID: 10641030.
2. Matasar MJ, Ritchie EK, Consedine N, Magai C, Neugut AI. Incidence rates of the major leukemia subtypes among US Hispanics, Blacks, and non-Hispanic Whites. Leukemia & lymphoma. 2006 Nov;47(11):2365-70. PubMed PMID: 17107911.
3. Okcu MF, Goodman KJ, Carozza SE, Weiss NS, Burau KD, Bleyer WA, et al. Birth weight, ethnicity, and occurrence of cancer in children: a population-based, incident case-control study in the State of Texas, USA. Cancer Causes Control. 2002 Sep;13(7):595-602. PubMed PMID: 12296506.
4. Yamamoto JF, Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. Cancer Causes Control. 2008 May;19(4):379-90. PubMed PMID: 18064533.
5. McBride ML, Gallagher RP, Theriault G, Armstrong BG, Tamaro S, Spinelli JJ, et al. Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. Am J Epidemiol. 1999 May 1;149(9):831-42. PubMed PMID: 10221320.
6. Reynolds P, Elkin E, Scalf R, Von Behren J, Neutra RR. A case-control pilot study of traffic exposures and early childhood leukemia using a geographic information system. Bioelectromagnetics. 2001;Suppl 5:S58-68. PubMed PMID: 11170118.
7. Macdougall LG, Jankowitz P, Cohn R, Bernstein R. Acute childhood leukemia in Johannesburg. Ethnic differences in incidence, cell type, and survival. The American journal of pediatric hematology/oncology. 1986 Spring;8(1):43-51. PubMed PMID: 3459378.
8. Reynolds P, Von Behren J, Elkin EP. Birth characteristics and leukemia in young children. Am J Epidemiol. 2002 Apr 1;155(7):603-13. PubMed PMID: 11914187.
9. Ross JA, Davies SM, Potter JD, Robison LL. Epidemiology of childhood leukemia, with a focus on infants. Epidemiologic reviews. 1994;16(2):243-72. PubMed PMID: 7713179.
10. Sandler DP, Ross JA. Epidemiology of acute leukemia in children and adults. Seminars in oncology. 1997 Feb;24(1):3-16. PubMed PMID: 9045302.
11. Gurney JG, Severson RK, Davis S, Robison LL. Incidence of cancer in children in the United States. Sex-, race-, and 1-year age-specific rates by histologic type. Cancer. 1995 Apr 15;75(8):2186-95. PubMed PMID: 7697611.
12. Linabery AM, Ross JA. Trends in childhood cancer incidence in the U.S. (1992-2004). Cancer. 2008 Jan 15;112(2):416-32. PubMed PMID: 18074355.
13. Dores GM, Devesa SS, Curtis RE, Linet MS, Morton LM. Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. Blood. 2012 Jan 5;119(1):34-43. PubMed PMID: 22086414. Pubmed Central PMCID: 3251235.
14. Milne E, Laurvick CL, Blair E, Bower C, de Klerk N. Fetal growth and acute childhood leukemia: looking beyond birth weight. Am J Epidemiol. 2007 Jul 15;166(2):151-9. PubMed PMID: 17443021.
15. Podvin D, Kuehn CM, Mueller BA, Williams M. Maternal and birth characteristics in relation to childhood leukaemia. Paediatr Perinat Epidemiol. 2006 Jul;20(4):312-22. PubMed PMID: 16879503.
16. Stiller CA, Parkin DM. Geographic and ethnic variations in the incidence of childhood cancer. British medical bulletin. 1996 Oct;52(4):682-703. PubMed PMID: 9039726.

17. McNeil DE, Cote TR, Clegg L, Mauer A. SEER update of incidence and trends in pediatric malignancies: acute lymphoblastic leukemia. Med Pediatr Oncol. 2002 Dec;39(6):554-7; discussion 2-3. PubMed PMID: 12376977.
18. Campleman SL, Wright, W.E.,. Childhood Cancer in California 1988 to 1999. Volume I: Birth to Age 14. . Sacramento, CA: California Cancer Registry, 2004 July 2004. Report No.
19. Chow EJ, Puumala SE, Mueller BA, Carozza SE, Fox EE, Horel S, et al. Childhood cancer in relation to parental race and ethnicity: a 5-state pooled analysis. Cancer. 2010 Jun 15;116(12):3045-53. PubMed PMID: 20564410. Pubmed Central PMCID: 2903004.
20. Humes K, Jones, NA, Ramirez, RR. Overview of Race and Hispanic Origin: 2010. U.S. Census Bureau, 2011 March 2011. Report No.: Contract No.: C2010BR-02.
21. The California Cancer Registry. Mission Statement & Purpose Sacramento, CA [March 2009]. Available from: http://www.ccrcal.org/aboutttheccr.html.
22. Schoendorf KC, Branum AM. The use of United States vital statistics in perinatal and obstetric research. American journal of obstetrics and gynecology. 2006 Apr;194(4):911-5. PubMed PMID: 16580275.
23. Slusky DA, Mezei G, Metayer C, Selvin S, Von Behren J, Buffler PA. Comparison of racial differences in childhood cancer risk in case-control studies and population-based cancer registries. Cancer epidemiology. 2012 Feb;36(1):36-44. PubMed PMID: 22018954.
24. Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). Cancer Causes Control. 1998 Aug;9(4):369-80. PubMed PMID: 9794168.
25. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control. 2001 Oct;12(8):703-11. PubMed PMID: 11562110.
26. Hosmer DW, Jr. , Lemeshow, S. Applied Logistic Regression. New York: Wiley; 2000.
27. SAS Institute I. SAS Statistical Software. 9.1 ed. Cary, NC2006.
28. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991 Apr;10(4):585-98. PubMed PMID: 2057657.
29. Little R, Rubin, D. Statistical Analysis with Missing Data. 2nd Edition ed. New York: Wiley; 2002 2002.
30. Zahm SH, Devesa SS. Childhood cancer: overview of incidence trends and environmental carcinogens. Environ Health Perspect. 1995 Sep;103 Suppl 6:177-84. PubMed PMID: 8549470. Pubmed Central PMCID: 1518893.
31. Draper G. The Geographical Epidemiology of Childhood Leukaemia and Non-Hodgkin Lymphomas in Great Britain, 1966-83. Surveys OoPCa, editor. London: H.M. Stationery Office; 1991. 131 p.
32. David RJ, Collins JW, Jr. Differing birth weight among infants of U.S.-born blacks, African-born blacks, and U.S.-born whites. The New England journal of medicine. 1997 Oct 23;337(17):1209-14. PubMed PMID: 9337381.
33. Fang J, Madhavan S, Alderman MH. Low birth weight: race and maternal nativity--impact of community income. Pediatrics. 1999 Jan;103(1):E5. PubMed PMID: 9917485.
34. McGrady GA, Sung JF, Rowley DL, Hogue CJ. Preterm delivery and low birth weight among first-born infants of black and white college graduates. Am J Epidemiol. 1992 Aug 1;136(3):266-76. PubMed PMID: 1415148.

35. Shiono PH, Klebanoff MA, Graubard BI, Berendes HW, Rhoads GG. Birth weight among women of different ethnic groups. JAMA : the journal of the American Medical Association. 1986 Jan 3;255(1):48-52. PubMed PMID: 3940304.
36. Caughey RW, Michels KB. Birth weight and childhood leukemia: a meta-analysis and review of the current evidence. Int J Cancer. 2009 Jun 1;124(11):2658-70. PubMed PMID: 19173295.
37. Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H, et al. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. Am J Epidemiol. 2003 Oct 15;158(8):724-35. PubMed PMID: 14561661.
38. Howard SC, Metzger ML, Wilimas JA, Quintana Y, Pui CH, Robison LL, et al. Childhood cancer epidemiology in low-income countries. Cancer. 2008 Feb 1;112(3):461-72. PubMed PMID: 18072274.
39. Ross JA, Kasum CM, Davies SM, Jacobs DR, Folsom AR, Potter JD. Diet and risk of leukemia in the Iowa Women's Health Study. Cancer Epidemiol Biomarkers Prev. 2002 Aug;11(8):777-81. PubMed PMID: 12163333.
40. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, et al. Novel Susceptibility Variants at 10p12.31-12.2 for Childhood Acute Lymphoblastic Leukemia in Ethnically Diverse Populations. J Natl Cancer Inst. 2013 May 15;105(10):733-42. PubMed PMID: 23512250.
41. Green LM, Miller AB, Villeneuve PJ, Agnew DA, Greenberg ML, Li J, et al. A case-control study of childhood leukemia in southern Ontario, Canada, and exposure to magnetic fields in residences. Int J Cancer. 1999 Jul 19;82(2):161-70. PubMed PMID: 10389746.
42. Michaelis J, Schuz J, Meinert R, Menger M, Grigat JP, Kaatsch P, et al. Childhood leukemia and electromagnetic fields: results of a population-based case-control study in Germany. Cancer Causes Control. 1997 Mar;8(2):167-74. PubMed PMID: 9134240.
43. Baumeister L, Marchi K, Pearl M, Williams R, Braveman P. The validity of information on "race" and "Hispanic ethnicity" in California birth certificate data. Health services research. 2000 Oct;35(4):869-83. PubMed PMID: 11055453. Pubmed Central PMCID: 1089157.
44. Rothman KJ, Greenland, S., Lash, T.L.,. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.

**Table 1.** Race of child based on combination of parental races, California birth registry, 1986-2007.

| Mother's race | Father's race | | | | |
|---|---|---|---|---|---|
| | **White** | **Black** | **Asian** | **Other** | **Total** |
| **White** | 8889 | 147 | 114 | 72 | 9222 |
| **Black** | 41 | 556 | 5 | 3 | 605 |
| **Asian** | 217 | 19 | 852 | 4 | 1092 |
| **Other** | 70 | 9 | 3 | 65 | 147 |
| **Total** | 9217 | 731 | 974 | 144 | 11066 |

**Table 2.** Characteristics of study subjects, California birth registry, 1986-2007.

| Variables | Total | Cases (%) | Controls (%) | ALL # cases/controls [a] | AML # cases/controls [a] |
|---|---|---|---|---|---|
| **All** | 11576 | 5788 | 5788 | 4721 | 852 |
| **Mother's race** | | | | | |
| White | 9534 | 4859 (84.7) | 4675 (81.6) | 4020/3815 | 665/686 |
| Black | 646 | 230 (4.0) | 416 (7.3) | 154/342 | 62/56 |
| Asian | 712 | 366 (6.4) | 346 (6.0) | 277/283 | 71/54 |
| Other | 575 | 282 (4.9) | 293 (5.1) | 238/241 | 35/38 |
| Missing | 109 | 51 | 58 | 32/40 | 19/18 |
| **Father's race** | | | | | |
| White | 9220 | 4706 (84.6) | 4514 (81.9) | 3910/3697 | 630/649 |
| Black | 732 | 274 ( 4.9) | 458 (8.3) | 186/375 | 73/62 |
| Asian | 617 | 327 (5.9) | 290 (5.3) | 239/231 | 69/51 |
| Other | 501 | 254 (4.6) | 247(1.1) | 208/208 | 36/28 |
| Missing | 506 | 227 | 279 | 178/210 | 44/62 |
| **Child's race** (original) | | | | | |
| White | 6257 | 3184 (81.9) | 3073 (78.8) | 2632/2506 | 451/451 |
| Black | 577 | 222 (5.7) | 355 (9.1) | 146/283 | 62/54 |
| Asian | 807 | 411 (10.6) | 396 (10.2) | 317/318 | 72/65 |
| Other | 150 | 73 (1.9) | 77 (2.0) | 59/61 | 12/10 |
| Missing | 3785 | 1898 | 1887 | 1567/1553 | 270/272 |
| **Child's race** (reconstructed from parental race) | | | | | |
| White | 8889 | 4550 (81.9) | 4339 (78.8) | 3777/3559 | 610/622 |
| Black | 780 | 290 (5.2) | 490 (8.9) | 198/399 | 76/68 |
| Asian | 709 | 338 (6.8) | 331 (6.0) | 283/265 | 75/56 |
| Other | 688 | 341 (6.1) | 347 (6.3) | 283/286 | 47/44 |
| Missing | 510 | 229 | 281 | 180/212 | 44/62 |
| **Hispanic origin of mother** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Hispanic | 5417 | 2858 (49.6) | 2559 (44.5) | 2380/2082 | 387/381 |
| Non-Hispanic | 6089 | 2899 (50.4) | 3190 (55.5) | 2317/2607 | 458/464 |
| Missing | 70 | 31 | 39 | 24/32 | 7/7 |
| **Hispanic origin of father** | | | | | |
| Hispanic | 5189 | 2737 (49.1) | 2452 (44.3) | 2281/2021 | 368/341 |
| Non-Hispanic | 5920 | 2842 (50.9) | 3087 (55.7) | 2269/2501 | 451/459 |
| Missing | 467 | 209 | 258 | 171/199 | 33/52 |
| **Hispanic origin of child** | | | | | |
| Both parents Hispanics | 4685 | 2481 (43.7) | 2204 (39.1) | 2074/1809 | 331/314 |
| One parent Hispanic | 1236 | 633 (11.1) | 603 (10.7) | 513/485 | 93/94 |
| Both parents non-Hispanic | 5396 | 2569 (45.2) | 2827 (50.2) | 2047/2304 | 411/417 |
| Missing | 259 | 105 | 154 | 87/123 | 17/27 |

[a] Number of cases and controls for ALL and AML do not add up for the total number of cases and controls for childhood leukemia because there were few other subtypes in the dataset.

**Table 3**. Conditional odds ratios (95% CIs) for childhood leukemia and race of child (N=7982), mother (N=8096) and father (N=7984) matched on child's age and sex and adjusted for birth order, birth weight, mother's age, and father's education. California birth registry, 1986-2007.

| | All types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Mother's race** | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.53 | 0.43 | 0.66 | 0.43 | 0.33 | 0.56 | 1.24 | 0.74 | 2.08 |
| Asian | 0.98 | 0.84 | 1.14 | 0.95 | 0.80 | 1.13 | 1.12 | 0.74 | 1.71 |
| Other | 0.70 | 0.47 | 1.05 | 0.70 | 0.45 | 1.09 | 0.86 | 0.28 | 2.60 |
| **Father's race** | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.55 | 0.45 | 0.67 | 0.45 | 0.36 | 0.57 | 1.31 | 0.81 | 2.13 |
| Asian | 1.01 | 0.86 | 1.19 | 0.94 | 0.78 | 1.13 | 1.41 | 0.91 | 2.18 |
| Other | 0.92 | 0.61 | 1.39 | 1.01 | 0.64 | 1.59 | 0.96 | 0.31 | 2.96 |
| **Child's race** (original) | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.61 | 0.48 | 0.78 | 0.49 | 0.37 | 0.65 | 1.50 | 0.86 | 2.61 |
| Asian | 0.90 | 0.74 | 1.10 | 0.86 | 0.69 | 1.07 | 1.20 | 0.70 | 2.04 |
| Other | 0.82 | 0.53 | 1.27 | 0.87 | 0.54 | 1.40 | 0.92 | 0.27 | 3.17 |
| **Child's race** (reconstructed from parental races) | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.54 | 0.44 | 0.66 | 0.45 | 0.36 | 0.56 | 1.19 | 0.75 | 1.90 |
| Asian | 0.99 | 0.85 | 1.15 | 0.94 | 0.80 | 1.12 | 1.26 | 0.85 | 1.89 |
| Other | 0.89 | 0.64 | 1.25 | 0.94 | 0.65 | 1.36 | 1.07 | 0.42 | 2.75 |
| **Child's race** (combined parental race) | | | | | | | | | |
| White/White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black/Black | 0.50 | 0.40 | 0.63 | 0.40 | 0.30 | 0.52 | 1.35 | 0.77 | 2.37 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Asian/Asian | 1.03 | 0.87 | 1.23 | 0.98 | 0.81 | 1.19 | 1.34 | 0.83 | 2.16 |
| Other/Other | 0.65 | 0.35 | 1.22 | 0.63 | 0.31 | 1.26 | 0.98 | 0.19 | 5.07 |
| White/ Black | 0.74 | 0.52 | 1.05 | 0.69 | 0.46 | 1.04 | 1.10 | 0.51 | 2.41 |
| White/Asian | 0.90 | 0.69 | 1.17 | 0.86 | 0.64 | 1.16 | 1.14 | 0.58 | 2.23 |
| White/Other | 1.00 | 0.66 | 1.50 | 1.06 | 0.68 | 1.64 | 1.13 | 0.36 | 3.59 |
| Black/Asian | 0.31 | 0.10 | 0.94 | 0.23 | 0.07 | 0.80 | [a] | [a] | [a] |
| Black/Other | 0.14 | 0.02 | 1.15 | 0.19 | 0.02 | 1.55 | [b] | [b] | [b] |
| Asian/Other | 1.49 | 0.25 | 8.97 | 2.97 | 0.31 | 28.67 | [b] | [b] | [b] |

[a] Not estimable due to small cell counts
[b] No observations

**Table 4**. Conditional odds ratios (95% CIs) for childhood leukemia and Hispanic origin of child matched on child's age and sex and adjusted for birth order, birth weight, mother's age, and father's education (N=8078). California birth registry, 1986-2007.

| | All types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Hispanic ethnicity of child** | | | | | | | | | |
| Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Hispanic | 1.30 | 1.18 | 1.44 | 1.36 | 1.22 | 1.52 | 1.03 | 0.78 | 1.37 |
| Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| One parent Hispanic | 1.17 | 1.00 | 1.38 | 1.19 | 1.00 | 1.42 | 1.13 | 0.74 | 1.73 |
| Both parents Hispanics [a] | 1.35 | 1.21 | 1.50 | 1.42 | 1.26 | 1.60 | 1.00 | 0.74 | 1.35 |

[a] trend test p-value < 0.0001

**Table 5.** Conditional odds ratios (95% CIs) for childhood leukemia and combined child's race and Hispanic ethnicity, adjusted for birth order, birth weight, mother's age, father's education and matched on child's age and sex (N=7968). California birth registry, 1986-2007.

| **Combined child's race and Hispanic ethnicity** | **All types** | | | **ALL** | | | **AML** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| Non-Hispanic White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Hispanic White | 1.24 | 1.10 | 1.39 | 1.28 | 1.13 | 1.46 | 1.08 | 0.77 | 1.50 |
| Non-Hispanic Black | 0.57 | 0.46 | 0.71 | 0.48 | 0.37 | 0.62 | 1.18 | 0.68 | 2.05 |
| Hispanic Black | 1.04 | 0.65 | 1.68 | 0.94 | 0.54 | 1.64 | 1.70 | 0.56 | 5.20 |
| Non-Hispanic Asian | 1.14 | 0.96 | 1.36 | 1.12 | 0.93 | 1.35 | 1.29 | 0.81 | 2.04 |
| Hispanic Asian | 0.92 | 0.59 | 1.44 | 0.79 | 0.47 | 1.32 | 1.66 | 0.58 | 4.74 |
| Non-Hispanic Other | 0.98 | 0.61 | 1.56 | 1.09 | 0.65 | 1.83 | 1.13 | 0.32 | 4.05 |
| Hispanic Other | 1.13 | 0.69 | 1.87 | 1.20 | 0.70 | 2.07 | 1.08 | 0.26 | 4.55 |

## Appendix

**Table 6**. Conditional odds ratios (95% CIs) based on multiple imputations of missing data for childhood leukemia and mother's, father's race, child's race and Hispanic ethnicity matched on child's age and sex and adjusted for birth order, birth weight, mother's age, and father's education. California birth registry, 1986-2007.

| | **All types** | | | **ALL** | | | **AML** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Mother's race** | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.56 | 0.47 | 0.66 | 0.46 | 0.38 | 0.56 | 1.17 | 0.80 | 1.71 |
| Asian | 1.02 | 0.90 | 1.16 | 0.98 | 0.85 | 1.12 | 1.27 | 0.92 | 1.74 |
| Other | 0.79 | 0.57 | 1.08 | 0.76 | 0.54 | 1.07 | 0.95 | 0.43 | 2.10 |
| **Father's race** | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.60 | 0.51 | 0.71 | 0.50 | 0.42 | 0.60 | 1.30 | 0.90 | 1.87 |
| Asian | 1.06 | 0.92 | 1.21 | 1.00 | 0.86 | 1.16 | 1.40 | 1.00 | 1.96 |
| Other | 1.02 | 0.73 | 1.42 | 0.93 | 0.64 | 1.33 | 1.91 | 0.79 | 4.62 |
| **Child's race** (reconstructed from parental races) | | | | | | | | | |
| White | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Black | 0.59 | 0.51 | 0.69 | 0.50 | 0.42 | 0.60 | 1.21 | 0.86 | 1.72 |
| Asian | 1.04 | 0.92 | 1.18 | 0.99 | 0.87 | 1.14 | 1.31 | 0.96 | 1.78 |
| Other | 0.94 | 0.72 | 1.24 | 0.89 | 0.67 | 1.20 | 1.41 | 0.68 | 2.92 |
| **Child's Hispanic ethnicity** | | | | | | | | | |
| Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Hispanic | 1.28 | 1.18 | 1.38 | 1.31 | 1.20 | 1.43 | 1.04 | 0.83 | 1.29 |
| Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| One parent Hispanic | 1.22 | 1.07 | 1.38 | 1.24 | 1.08 | 1.42 | 1.07 | 0.76 | 1.50 |
| Both parents Hispanics [a] | 1.32 | 1.22 | 1.44 | 1.37 | 1.25 | 1.50 | 1.08 | 0.86 | 1.35 |

[a] trend test p-value < 0.0001

## Chapter 6

# Socio-economic status and childhood leukemia in California

Oksuzyan S[1], Crespi CM[3], Cockburn M[2], Mezei G[4], Vergara X[5], Kheifets L[6]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, soksuzyan@ucla.edu; [2]Department of Preventive Medicine, University of Southern California, Los Angeles, California, myles@med.usc.edu; [3]Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California, ccrespi@ucla.edu; [4]Environment Department, Electric Power Research Institute, Palo Alto, California, gmezei@epri.com; [5]Environment Department, Electric Power Research Institute, Palo Alto, California, xvergara@epri.com; [6]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, kheifets@ucla.edu.

Correspondence: Sona Oksuzyan, MD, MPH, PhD (c), UCLA, Department of Epidemiology.
Address: 17072 Signature Drive, Granada Hills, CA 91344
Phone: 818-731-5854
Email: soksuzyan@ucla.edu

**Abstract**

*Aims:* We conducted a large registry-based study in California to investigate the association between socio-economic status (SES) at individual and community levels and childhood leukemia, focusing on two subtypes: acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML).

*Methods:* We obtained information on 5788 cases and 5788 controls by linking California cancer and birth registries. We evaluated the relative risk of childhood leukemia by both community-level (census-based) and individual-level SES measures (parental education and the source of payment for delivery) using conditional logistic regression, with adjustment for potential confounders.

*Results:* Children in the higher census-based SES categories had a tendency toward slightly decreased risk of total childhood leukemia and ALL compared to children of the lowest SES. Compared to < 12 years of school, 13 and more years of parental education were associated with slightly decreased risk of total childhood leukemia and ALL. However, confidence intervals were wide and included null value. Sizably decreased risk of AML was noted for children with higher educated fathers (12-16 and 17 years) compared to those whose fathers were least educated (< 12 years of school); the association weakened when performing an analysis using multiply imputed data. The source of payment for delivery was not associated with childhood leukemia nor with either subtype.

*Conclusions*: Overall, we found little evidence to support the suggestion that SES, as measured by a variety of proxies, is associated with risk of childhood leukemia or either major subtype. It is likely that results of many previous studies that found an association between childhood leukemia and SES were influenced by selection or ecological bias.

**Key words:** childhood leukemia, socio-economic status, census-based SES, parental education, father's education, parental education, source of payment for delivery

## Introduction

Leukemia accounts for the largest number of cases of childhood cancer in the United States and worldwide.(1) Risk factors for this disease remain mostly unidentified.

The role of socio-economic status (SES) in the development of childhood leukemia remains poorly understood. To elucidate this association, some studies used maternal or paternal education, occupation, or family income as proxies for individual-level SES.(2-6) Other studies have used community-level SES, for example, census tract or ward-based socio-economic characteristics, e.g. income, population density, rent, employment, community-level occupational class, and social deprivation score. (7-12) A small number of studies have looked at SES at both individual and community levels.(13-15) Poole at al. reviewed literature on SES and childhood leukemia and concluded that in studies using individual-level measures (family income, mother's and father's education) that required some subject contact, lower SES was linked to higher risk of leukemia. (16) In contrast, in record-based studies using father's occupational class and in ecologic studies using community-level occupational class, higher SES was associated with higher risk of childhood leukemia. The authors suggested that these apparent contradictions may have resulted from differences in design and/or SES measure, but could not distinguish between these two hypotheses. One study, by MacMahon (1962), used medical payment status as a proxy measure for SES to examine mortality data from 1947 to 1960.(17) Two other studies used the source of payment for prenatal care as a proxy for SES in research on preterm birth.(18, 19)

In our study, using population-based registry data from California not involving contact with subjects, we explore relationships between individual-level SES (father's and mother's education and payment source for medical services) and community-level SES (census block-based SES) with childhood leukemia and examine whether the relationships between individual- and community-level SES and childhood leukemia risk are different from each other, as hypothesized in the review by Poole et al.

## Materials and methods

The population-based California Cancer Registry (CCR) was used to obtain information on all childhood leukemia cases diagnosed between 1988 and 2008 at younger than 16 years of age and born and residing in California at the time of diagnosis. The statewide cancer registry

regularly records age, race/ethnicity, sex, and residence at the time of diagnosis as well as information on cancer types and characteristics. Controls were selected randomly from the California Birth Registry (CBR) for years 1986-2007 and matched to cases (1 to 1) on the basis of date of birth (±6 months) and sex. Thus, controls were children younger than 16 years old at the time of "pseudo-diagnosis," who had not been diagnosed in California with any type of cancer.

Information on gender, parental age (in years), perinatal characteristics (e.g. birth weight, birth order), maternal and paternal years of education, sources of payment for prenatal care and delivery, and race/ethnicity were extracted from California birth records.

## Measures of Socioeconomic Status (SES)

### *Individual-Level Measures*

We used parental education and the source of payment for delivery as proxies for SES at the individual level. We categorized parental years of education into four levels: <12 years (less than high school), 12 years (high school), 13-16 years (some college, college), and 17 years and more (graduate school). Since maternal and paternal years of education were correlated, we used them in different models. We also combined father's and mother's years of education into one variable using principal component analysis (PCA) and then used the resulting principal component score in some models.

There were two sources of payment available from birth records: for prenatal care and for delivery. The source of payment for delivery was considered a better predictor of SES than the source of payment for prenatal care (personal communication, Gerald Kominski, Professor, Department of Health Services, UCLA School of Public Health, Associate Director, UCLA Center for Health Policy Research). In addition, we compared estimates and 95% confidence intervals for sources of payment for prenatal care and delivery for similarity. Because they were nearly identical, the latter was used in all analyses. We categorized the source of payment for delivery using two approaches. In one approach, we created three categories for the source of payment for delivery: low SES (including governmental programs such as Medicare, Medi-Cal, Title V, other governmental programs, Indian Health Services, 'No care' and 'Medically indigent'), middle SES (military and worker's compensations, self pay, CHAMPUS/TRICARE, no charge and other sources of payment) and high SES (Blue Cross/Blue Shield, private

insurance, and health maintenance organization/prepaid health plan). In a second classification, the middle and high categories were combined to yield a total of two categories: low (same as before) and middle-high (combining the highest two in the previous classification) SES.

*Community-Level Measures*

A measure of census-based SES was provided by our collaborators from USC. The variable was derived from U.S. Census data using principal components analysis based on seven indicator variables at a census block group level: education index developed by Liu et al. (20) that weights the proportion of people with a given level of education by a number of years need to attain that level, proportion with a blue-collar job, proportion older than 16 without a job, median household income, proportion below 200% of the federal poverty level, median rent, median home value. Available to us were the quintiles of the PCA scores. (21)

Census-based SES was assessed using the address at birth obtained from CBR and the address at diagnosis from CCR. Since individual-level SES was available only at birth and the address at diagnosis was available only for cases, the main focus of the study was SES at birth. For subjects whose birth address could not be resolved to a census tract, census-based SES could not be obtained and was unknown.

We adjusted all models for child's race. Although racial origin of the child was collected in birth records, more than 50% of values were missing. We used more complete variables encoding racial origin of mother and father from birth records to derive the race of the child and combined those into four main racial groups: White, Black, Asian and Other. More detailed information for race/ethnicity is described elsewhere. (22)

Statistical Analysis

As a part of a descriptive analysis, we examined relationships among father's education, mother's education, the source of payment for delivery and census-based SES using kappa statistics and Spearman correlation and the distribution of individual-level proxies for SES by census-based SES.

We tested the association of childhood leukemia and variables of interest using unadjusted and adjusted conditional logistic regression. (23) Since results were not substantially different, we present results of adjusted analyses only. Various models with different subsets of

covariates were fit and checked for potentially influential observations. For adjusted analysis, the models for SES were chosen based on information on known or potential confounders and model fit statistics; models with the lowest Akaike information criterion **(**AIC) value and the lowest number of potential confounders are presented.

To test whether census-based SES or individual level SES measures were predictive of childhood leukemia risk, we also conducted joint significance tests of the sets of coefficients associated with the SES variables (null hypothesis that all coefficients equal zero vs. alternative that at least one is nonzero).

Since risk factors for the two main subtypes of childhood leukemia, acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML), could be different, analyses were also conducted for these subtypes separately.

Despite the large number of cases and controls, sample sizes were smaller for some analyses due to missing data. Most missingness was due to variation in the variables collected on birth records from year-to-year, and no systematic differences in patterns of missingness were detected between cases and controls. As a sensitivity analysis, multivariate imputation techniques were used to impute missing or unknown values for father's and mother's education, payment source for delivery and census-based SES under a missing at random assumption using the MI procedure in SAS. (24-26) The imputation models included all variables used in the analytic models. Analyses were repeated using the multiply imputed data using the MIANALYZE procedure.(24)

Analyses were conducted using SAS 9.3. (24)

The study was approved by University of California, Los Angeles Office for the Protection of Research Subjects.

**Results**

A total of 6645 childhood leukemia cases were identified from the CCR. Linkage to birth records was successful for 87.1% (5788/6645) of cases. Of the 5788 cases (55.8% males and 44.2% females) included in this analysis, 4721 were ALL cases (56.2% males and 43.8% females), 852 were AML cases (53.3% males and 46.7% females), and 215 were other childhood leukemia subtypes. The median age at diagnosis was 3.8 years with a range of 0 to 15.6 years with the peak for ALL between 2 and 5 years of age and between 0 and 2 years of age for AML.

Table 1 shows other characteristics of study subjects.

To compare individual- and community-level measures, we present the distribution of father's education and mother's education, the source payment for delivery and census-based SES by each other presented in Figure 1. As shown in the figure, the proportion of high census-based SES increased and the proportion of low census-based SES decreased with increasing years of paternal and maternal education. A similar pattern was exhibited for the relationship between source of payment for delivery and census-based SES: the proportion of governmental programs decreased and the proportion of private insurance and other sources increased with increasing census-based SES.

Kappa statistics indicated slight agreement between father's education and census-based SES and between father's education and the source of payment for delivery (weighted kappa= 0.18 for both). The highest weighted kappa was observed between mother's and father's education (kappa=0.60). (Table 2) Similar results were obtained using Spearman correlation coefficients.

In adjusted analysis, we detected a tendency toward slightly decreased risk of total childhood leukemia and ALL in children of the middle and middle-high census-based SES compared to the lowest SES class, but the associations were imprecise. Compared to children whose parents (mother and father) were least educated (< 12 years of school), children with parents having higher education levels (13-16 and ≥17 years) had slightly decreased risk of total childhood leukemia and ALL although confidence intervals were wide. Sizably decreased risk of AML was noted for children with higher educated fathers compared to those whose fathers were least educated (< 12 years of school). (Table 3)

When we used a principal component score for father's and mother's years of education, results did not differ from separate models containing either mother's education or father's education (OR=0.96, 95% CI: 0.89-1.03).

We also conducted joint significance tests of the sets of coefficients associated with the SES variables. P-value for joint significance test for census-based SES was 0.10, for source of payment for delivery p-value=0.45, for father's and mother's education p-value=0.66 and 0.27, respectively.

After performing complete case analyses, analyses were repeated using multiply imputed data. Results, presented in Appendix Table 4, were very similar to complete case analysis, except the association of father's education and AML. The negative association found in the complete case analysis was not replicated in analysis with multiply imputed data. For 13-16 years of paternal education, OR= 0.79 with 95% CI: 0.54-1.18, and for 17 years of education, OR=0.86 with 95% CI: 0.50-1.47.

**Discussion**

We assessed the association of childhood leukemia and socio-economic status (SES) using individual and community level measures. We used father's education and mother's education as individual-level SES proxies similar to previous research. As a departure from previous work in childhood leukemia research, we defined a new way to assess individual-level SES, namely, the source of payment for delivery. We were also able to assess the association between individual-level and community-level measures for this population-based sample.

For census-based SES and for parental education variables, the point estimates suggest possible slight decreased risk for higher levels, but with wide confidence intervals. For AML, children of higher educated fathers (12 and 13-16 years), but not for 17+ were at decreased risk compared to children of lower educated fathers. However, this association appeared to weaken and become less precise in analysis using multiply imputed dataset, suggesting that the initial finding was a chance result. The source of payment for delivery did not show a clear association with childhood leukemia, nor with its main subtypes, ALL and AML.

Many studies that used individual-level proxies for SES have detected an inverse association with childhood leukemia, i.e. low SES was associated with higher risk of leukemia. (2, 4-6, 16, 27-29) The majority of these studies assessed individual-level SES using self-administered questionnaires or interviews. In such studies, selection bias may occur because controls of high SES are usually more likely to participate. (30) A major strength of our study was that data were obtained from population registries with almost complete registration of births and cancers in California. Also, controls were randomly selected from the CBR rather than directly involving participants, as in many case-control studies. Since these registries are independent of each other and participation of subjects was not required for our data collection, selection bias was unlikely.

Although the majority of studies that used community-level proxies for SES found an increased risk of childhood leukemia for high SES, some studies had reported an inverse association with childhood leukemia (higher risk of leukemia for low SES). (7, 8, 14, 16) Many of the studies that detected a positive association were purely ecological in design and thus were prone to ecological bias. We detected a tendency toward negative association between childhood leukemia and census-based SES.

A high risk of childhood leukemia associated with higher SES would be more consistent with the viral hypothesis suggested by Greaves (31, 32). In this hypothesis, late exposure to common infections, potentially associated with higher SES, could lead to abnormal immune response and play an important role in the development of childhood leukemia (10, 11). An inverse association of SES with childhood leukemia is more consistent with a "population mixing" theory (33). According to this theory, some childhood leukemia cases could be a rare immune response to unidentified infections introduced by a high level of personal contacts and /or large migrations of new people into "closed" communities by increased commuting and travel ("population mixing") that is potentially associated with lower SES. (10)

Another advantage of this study was that the large size of the dataset allowed us to carry out analyses for two main subtypes of childhood leukemia, ALL and AML. We did not find a clear association of any SES measure with either subtype, which may further indicate that SES is not a determinant of childhood leukemia.

One of the limitations of this study is potential misclassification of variables of interest and covariates. Misclassification of the outcome was unlikely due to high accuracy of the

CCR.(34) Misclassification of SES was possible due to potential inaccuracy of reporting and/or categorization of parental education and the source of payment for delivery. Parental education, particularly maternal education, was not consistently recorded on birth certificates, its completeness varied from year to year, and we do not know how accurate it is, but we believe that misclassification of parental education was not differential for cases and controls. (35) We also think that the source of payment for delivery was reported fairly accurately because medical entities relied on this information for payment for delivery care. (36) Misclassification could possibly arise from our categorizations of this variable. However, in our sensitivity analysis of various categorizations of these two classifications, our results were unaltered.

Another limitation of the study was missing data. However, since information was missing mainly due to year-to-year differences in the information collected on birth certificates rather than non-response, the potential for systematic biases was probably small and the impact was mainly on the precision of the estimates. For example, information on parental education was available only for more recent years. Since census-based SES did not come from CBR, we assessed if subjects that were missing census-based SES were different from subjects with this variable. We found that subjects with missing census-based SES had slightly younger maternal age at birth, slightly lower paternal education and higher proportion of governmental programs as source of payment for delivery. There was no difference in the pattern of missingness between cases and controls for census-based SES nor for any other variable. Therefore, missingness patterns were not differential and probably did not bias our results. We re-analyzed the data using multiple imputation and obtained very similar results, except for the association for father's education and AML, which became weaker and less precise.

Overall, we found little support for the suggestion that SES, as measured by individual- or community-level proxies, is associated with the risk of childhood leukemia or of its major subtypes. It is likely that previous studies finding an association between childhood leukemia and SES were prone to selection or ecological bias.

**Acknowledgements**

This project was supported by a research contract from the Electric Power Research Institute to UCLA and by UCLA Faculty Grants Program. Crespi was also partially supported by National Institutes of Health grant P30 CA16042.

**Reference:**


1. Ma SK, Wan TS, Chan LC. Cytogenetics and molecular genetics of childhood leukemia. Hematol Oncol. 1999 Sep;17(3):91-105. PubMed PMID: 10641030.
2. Dockerty JD, Elwood JM, Skegg DC, Herbison GP. Electromagnetic field exposures and childhood cancers in New Zealand. Cancer Causes Control. 1998 May;9(3):299-309. PubMed PMID: 9684710.
3. Dockerty JD, Skegg DC, Elwood JM, Herbison GP, Becroft DM, Lewis ME. Infections, vaccinations, and the risk of childhood leukaemia. Br J Cancer. 1999 Jul;80(9):1483-9. PubMed PMID: 10424755.
4. Ma X, Buffler PA, Selvin S, Matthay KK, Wiencke JK, Wiemels JL, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. Br J Cancer. 2002 May 6;86(9):1419-24. PubMed PMID: 11986774. Pubmed Central PMCID: 2375371.
5. McBride ML, Gallagher RP, Theriault G, Armstrong BG, Tamaro S, Spinelli JJ, et al. Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. Am J Epidemiol. 1999 May 1;149(9):831-42. PubMed PMID: 10221320.
6. Reynolds P, Von Behren J, Elkin EP. Birth characteristics and leukemia in young children. Am J Epidemiol. 2002 Apr 1;155(7):603-13. PubMed PMID: 11914187.
7. Muirhead CR. Childhood leukemia in metropolitan regions in the United States: a possible relation to population density? Cancer Causes Control. 1995 Sep;6(5):383-8. PubMed PMID: 8547535.
8. Reynolds P, Elkin E, Scalf R, Von Behren J, Neutra RR. A case-control pilot study of traffic exposures and early childhood leukemia using a geographic information system. Bioelectromagnetics. 2001;Suppl 5:S58-68. PubMed PMID: 11170118.
9. Richardson S, Monfort C, Green M, Draper G, Muirhead C. Spatial variation of natural radiation and childhood leukaemia incidence in Great Britain. Stat Med. 1995 Nov 15-30;14(21-22):2487-501. PubMed PMID: 8711283.
10. Stiller CA, Kroll ME, Boyle PJ, Feng Z. Population mixing, socioeconomic status and incidence of childhood acute lymphoblastic leukaemia in England and Wales: analysis by census ward. Br J Cancer. 2008 Mar 11;98(5):1006-11. PubMed PMID: 18253115.
11. Borugian MJ, Spinelli JJ, Mezei G, Wilkins R, Abanto Z, McBride ML. Childhood leukemia and socioeconomic status in Canada. Epidemiology. 2005 Jul;16(4):526-31. PubMed PMID: 15951671.
12. Ribeiro KB, Buffler PA, Metayer C. Socioeconomic status and childhood acute lymphocytic leukemia incidence in Sao Paulo, Brazil. Int J Cancer. 2008 Oct 15;123(8):1907-12. PubMed PMID: 18688860.
13. Dockerty JD, Draper G, Vincent T, Rowan SD, Bunch KJ. Case-control study of parental age, parity and socioeconomic level in relation to childhood cancers. Int J Epidemiol. 2001 Dec;30(6):1428-37. PubMed PMID: 11821358.
14. McKinney PA, Juszczak E, Findlay E, Smith K, Thomson CS. Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study. Br J Cancer. 1999 Aug;80(11):1844-51. PubMed PMID: 10468308.

15. Smith A, Roman E, Simpson J, Ansell P, Fear NT, Eden T. Childhood leukaemia and socioeconomic status: fact or artefact? A report from the United Kingdom childhood cancer study (UKCCS). Int J Epidemiol. 2006 Dec;35(6):1504-13. PubMed PMID: 16945940.
16. Poole C, Greenland S, Luetters C, Kelsey JL, Mezei G. Socioeconomic status and childhood leukaemia: a review. Int J Epidemiol. 2006 Apr;35(2):370-84. PubMed PMID: 16308412.
17. Macmahon B. Prenatal x-ray exposure and childhood cancer. J Natl Cancer Inst. 1962 May;28:1173-91. PubMed PMID: 14468031.
18. Ghosh JK, Wilhelm MH, Dunkel-Schetter C, Lombardi CA, Ritz BR. Paternal support and preterm birth, and the moderation of effects of chronic stress: a study in Los Angeles county mothers. Archives of women's mental health. 2010 Aug;13(4):327-38. PubMed PMID: 20066551. Pubmed Central PMCID: 2896639.
19. Lewis C, Suffet IH, Hoggatt K, Ritz B. Estimated effects of disinfection by-products on preterm birth in a population served by a single water utility. Environ Health Perspect. 2007 Feb;115(2):290-5. PubMed PMID: 17384780. Pubmed Central PMCID: 1831522.
20. Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). Cancer Causes Control. 1998 Aug;9(4):369-80. PubMed PMID: 9794168.
21. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control. 2001 Oct;12(8):703-11. PubMed PMID: 11562110.
22. Oksuzyan S, Cockburn M, Crespi CM, Mezei G, Vergara X, Kheifets L. Race/ethnicity and childhood leukemia: a case-control study in California. Cancer Causes and Control. 2013, submitted Aug. 25, 2013.
23. Hosmer DW, Jr. , Lemeshow, S. Applied Logistic Regression. New York: Wiley; 2000.
24. SAS Institute I. SAS Statistical Software. 9.1 ed. Cary, NC2006.
25. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991 Apr;10(4):585-98. PubMed PMID: 2057657.
26. Little R, Rubin, D. Statistical Analysis with Missing Data. 2nd Edition ed. New York: Wiley; 2002 2002.
27. Kaye SA, Robison LL, Smithson WA, Gunderson P, King FL, Neglia JP. Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia. Cancer. 1991 Sep 15;68(6):1351-5. PubMed PMID: 1873786.
28. Perrillat F, Clavel J, Auclerc MF, Baruchel A, Leverger G, Nelken B, et al. Day-care, early common infections and childhood acute leukaemia: a multicentre French case-control study. Br J Cancer. 2002 Apr 8;86(7):1064-9. PubMed PMID: 11953850. Pubmed Central PMCID: 2364194.
29. Shu XO, Linet MS, Steinbuch M, Wen WQ, Buckley JD, Neglia JP, et al. Breast-feeding and risk of childhood acute leukemia. J Natl Cancer Inst. 1999 Oct 20;91(20):1765-72. PubMed PMID: 10528028.
30. Mezei G, Kheifets L. Selection bias and its implications for case-control studies: a case study of magnetic field exposure and childhood leukaemia. Int J Epidemiol. 2006 Apr;35(2):397-406. PubMed PMID: 16303812.
31. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer. 2006 Mar;6(3):193-203. PubMed PMID: 16467884.

32. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. Leukemia. 1988 Feb;2(2):120-5. PubMed PMID: 3278171.
33. Kinlen LJ, Stiller C. Population mixing and excess of childhood leukemia. Bmj. 1993 Apr 3;306(6882):930. PubMed PMID: 8490436.
34. The California Cancer Registry. Mission Statement & Purpose Sacramento, CA [March 2009]. Available from: http://www.ccrcal.org/abouttheccr.html.
35. Rothman KJ, Greenland, S., Lash, T.L.,. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
36. Baumeister L, Marchi K, Pearl M, Williams R, Braveman P. The validity of information on "race" and "Hispanic ethnicity" in California birth certificate data. Health services research. 2000 Oct;35(4):869-83. PubMed PMID: 11055453. Pubmed Central PMCID: 1089157.
37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159-74. PubMed PMID: 843571.
38. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. John Wiley & Sons L, editor. New York: John Wiley & Sons, Ltd.; 1981.

**Table 1.** Characteristics of study subjects, California birth registry, 1986-2007.

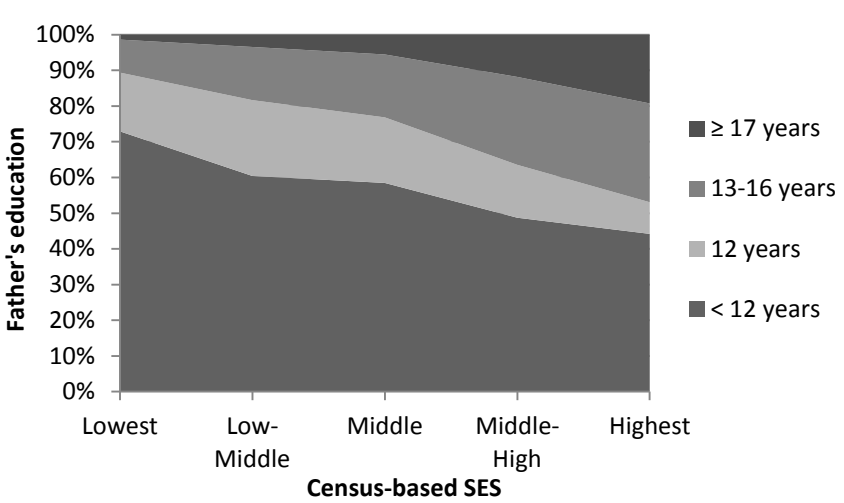
| Variables | Cases (%) | Controls (%) | ALL # cases/controls[a] | AML # cases/controls[a] |
|---|---|---|---|---|
| **All** | 5788 | 5788 | 4721 | 852 |
| **Census-based SES quintile** | | | | |
| Lowest | 1264 (26.6) | 1249 (26.3) | 1023/1009 | 184/195 |
| Lower-middle | 1126 (23.7) | 1131 (23.8) | 918/914 | 167/182 |
| Middle | 969 (20.4) | 990 (20.9) | 816/809 | 128/135 |
| Higher-middle | 713 (15.0) | 749 (15.8) | 559/609 | 122/107 |
| Highest | 672 (14.2) | 628 (13.2) | 557/522 | 87/84 |
| Unknown | 1044 | 1041 | 848/858 | 164/149 |
| **Source for payment for delivery (classification 1)[a]** | | | | |
| Governmental programs (Low SES) | 2173 (44.03) | 2284 (46.43) | 1741/1861 | 346/350 |
| Military, worker's compensation, self pay, other (Medium SES) | 157 (3.18) | 170 (3.46) | 137/135 | 17/28 |
| Private insurance, prepaid plans (High SES) | 2605 (52.79) | 2465 (50.11) | 2159/2031 | 361/339 |
| Missing (Not collected) * | 853 (845) | 869 (850) | 684/694 | 128/135 |
| **Father's education** | | | | |
| <12 years | 2478 (60.7) | 2443 (60.1) | 2014/1996 | 374/349 |
| 12 years | 658 (16.1) | 668 (16.5) | 553/561 | 87/89 |
| 13-16 years | 678 (16.6) | 691 (17.0) | 570/577 | 83/91 |
| ≥17 years | 269 (6.6) | 260 (6.4) | 219/213 | 40/40 |
| Missing (Not collected)* | 1705 (1509) | 1726 (1505) | 1365/1374 | 268/283 |
| **Mother's education** | | | | |
| <12 years | 730 (30.2) | 719 (30.0) | 607/587 | 96/103 |
| 12 years | 708 (29.3) | 671 (28.0) | 593/561 | 98/93 |
| 13-16 years | 737 (30.5) | 760 (31.8) | 605/635 | 103/100 |
| ≥17 years | 239 (9.9) | 244 (10.2) | 196/199 | 35/37 |
| Missing (Not collected)* | 3374 (3335) | 3394 (3340) | 2720/2739 | 520/519 |

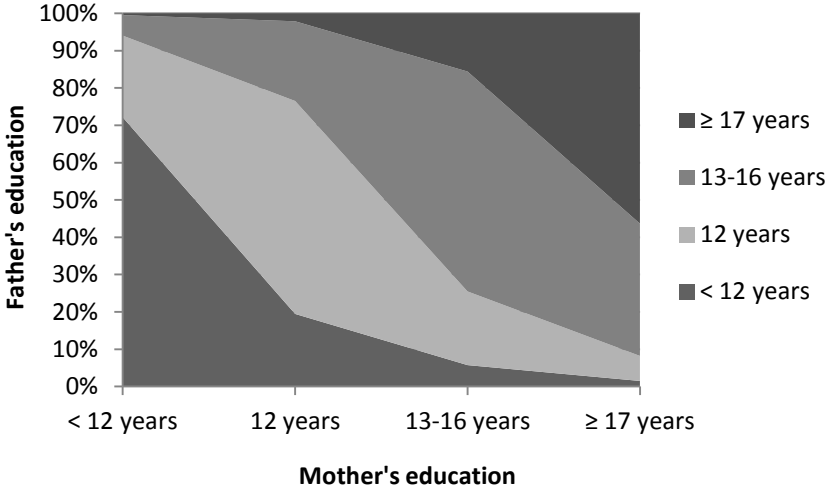
[a] The second classification of source of payment for delivery combined the middle and high SES categories.

* Patterns of missingness varied by year due to differences in data collection year-to-year

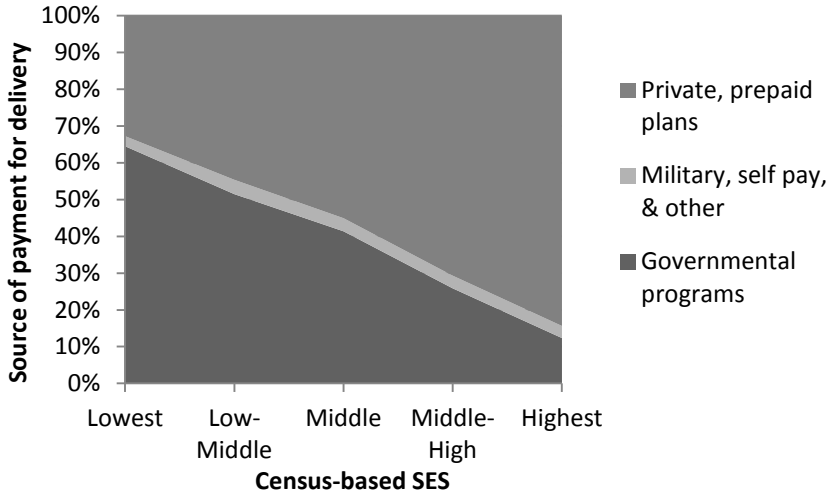
**Figure 1.** The distribution of individual- and community-level SES by each other, California birth registry, 1986-2007.

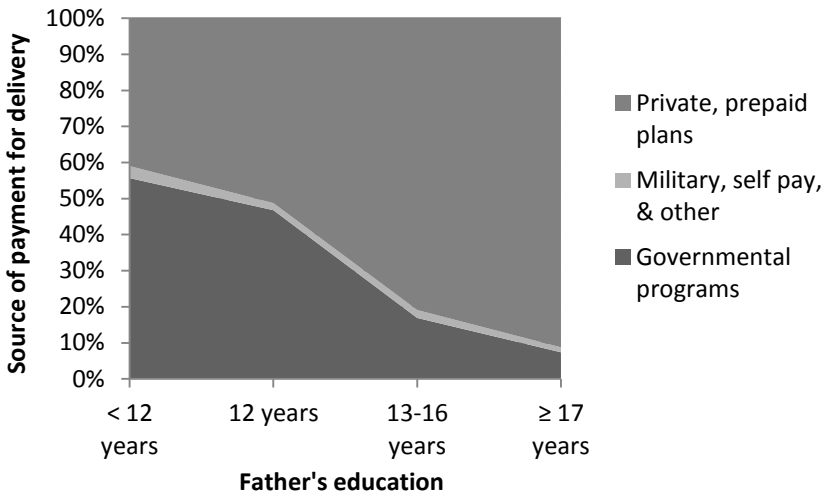
(a) Father's education by census-based SES and by Mother's education



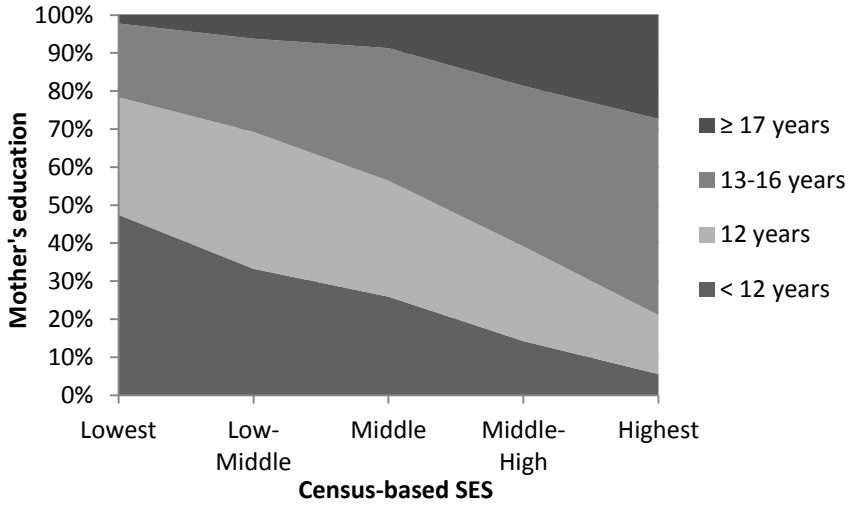




(b) Source of payment for delivery by census-based SES and by Father's education




(c) Mother's education by census-based SES and by Source of payment for delivery



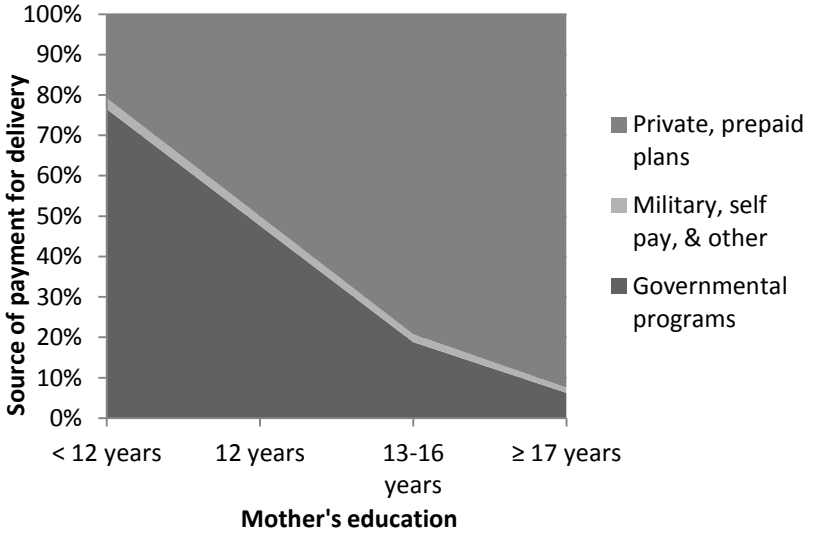



**Table 2**. Kappa statistics for agreement between individual- and community-level measures of socio-economic status of childhood leukemia cases and their matched controls, California birth registry, 1986-2007. [1]

| | Kappa statistics (adjusted) | Level of agreement [by Landis] (37) | Level of agreement [by Fleiss] (38) |
|---|---|---|---|
| Father's education and Source of payment for delivery | 0.18 | Slight | Poor |
| Father's education and Census-based SES | 0.18 | Slight | Poor |
| Father's education and Mother's education | 0.60 | Substantial | Good |
| Source of payment for delivery and Census-based SES | 0.29 | Fair | Poor |
| Mother's education and Source of payment for delivery | 0.29 | Fair | Poor |
| Mother's education and Census-based SES | 0.28 | Fair | Poor |

**Table 3**. Conditional odds ratios (95% CIs) for individual level and community level socio-economic status and childhood leukemia matched on child's age and sex, California birth registry, 1986-2007. [1]

| | All types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| ***Community-Level Measures*** | | | | | | | | | |
| **Census-based SES (5 categories)** | | | | | | | | | |
| Reference - Lowest | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Low-middle | 1.06 | 0.92 | 1.21 | 1.07 | 0.92 | 1.24 | 1.01 | 0.72 | 1.44 |
| Middle | 0.92 | 0.79 | 1.05 | 0.99 | 0.84 | 1.15 | 0.72 | 0.50 | 1.05 |
| Middle-High | 0.87 | 0.74 | 1.01 | 0.83 | 0.70 | 0.98 | 1.13 | 0.76 | 1.68 |
| Highest | 0.95 | 0.81 | 1.11 | 0.95 | 0.80 | 1.14 | 1.07 | 0.69 | 1.66 |
| p-value for trend | 0.093 | | | | | | | | |
| ***Individual-Level Measures*** | | | | | | | | | |
| **Source for payment for delivery (classification 1)** | | | | | | | | | |
| Governmental programs (low SES) | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Private insurance (middle-high SES) | 1.04 | 0.95 | 1.14 | 1.08 | 0.97 | 1.19 | 1.03 | 0.82 | 1.30 |
| **(classification 2)** | | | | | | | | | |
| Governmental programs (low SES) | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Military, self pay, and other (middle SES) | 0.94 | 0.74 | 1.19 | 1.08 | 0.83 | 1.40 | 0.58 | 0.30 | 1.12 |
| Private insurance (high SES) | 1.05 | 0.96 | 1.15 | 1.08 | 0.97 | 1.19 | 1.07 | 0.84 | 1.35 |
| **Father education** | | | | | | | | | |
| <12 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 12 years | 0.98 | 0.83 | 1.16 | 1.03 | 0.86 | 1.23 | 0.63 | 0.39 | 1.00 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 13-17 years | 0.91 | 0.77 | 1.08 | 0.95 | 0.79 | 1.14 | 0.63 | 0.40 | 1.00 |
| 17 and + years | 0.92 | 0.73 | 1.15 | 0.93 | 0.73 | 1.20 | 0.80 | 0.43 | 1.48 |
| p-value for trend | 0.33 | | | | | | | | |
| **Mother education** | | | | | | | | | |
| <12 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 12 years | 1.07 | 0.91 | 1.26 | 1.10 | 0.92 | 1.30 | 0.92 | 0.58 | 1.44 |
| 13-17 years | 0.92 | 0.78 | 1.09 | 0.90 | 0.75 | 1.08 | 0.91 | 0.56 | 1.50 |
| 17 and + years | 0.90 | 0.71 | 1.15 | 0.92 | 0.70 | 1.20 | 0.72 | 0.37 | 1.41 |
| p-value for trend | 0.22 | | | | | | | | |

[1] Adjusted for child's race, birth order, birth weight and mother's age

## Appendix

**Table 4**. Conditional odds ratios (95% CIs) for individual level and community level socio-economic status and childhood leukemia matched on child's age and sex resulting from multiple imputation analysis, California birth registry, 1986-2007. [1]

| | All types | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Census-based SES (5 categories)** | | | | | | | | | |
| Reference - Lowest | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Low-middle | 1.00 | 0.90 | 1.10 | 1.02 | 0.91 | 1.14 | 0.90 | 0.70 | 1.17 |
| Middle | 0.95 | 0.85 | 1.06 | 0.96 | 0.85 | 1.09 | 0.93 | 0.70 | 1.24 |
| Middle-High | 0.91 | 0.81 | 1.02 | 0.87 | 0.76 | 0.99 | 1.12 | 0.84 | 1.51 |
| Highest | 0.98 | 0.86 | 1.11 | 0.97 | 0.85 | 1.12 | 1.02 | 0.72 | 1.44 |
| **Census-based SES (3 categories)** | | | | | | | | | |
| Reference - Low | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Middle | 0.95 | 0.86 | 1.05 | 0.96 | 0.86 | 1.07 | 0.96 | 0.73 | 1.26 |
| High | 0.94 | 0.86 | 1.03 | 0.92 | 0.83 | 1.01 | 1.11 | 0.88 | 1.40 |
| **Source for payment for delivery (classification 1)** | | | | | | | | | |
| Governmental programs (low SES) | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Private insurance (middle-high SES) | 1.04 | 0.96 | 1.14 | 1.06 | 0.96 | 1.16 | 1.04 | 0.83 | 1.29 |
| **Source for payment for delivery (classification 2)** | | | | | | | | | |
| Governmental programs (low SES) | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Military, self pay, and other (middle SES) | 0.93 | 0.74 | 1.16 | 1.02 | 0.80 | 1.31 | 0.61 | 0.33 | 1.14 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Private insurance (high SES) | 1.06 | 0.97 | 1.15 | 1.07 | 0.97 | 1.17 | 1.07 | 0.86 | 1.33 |
| **Father education** | | | | | | | | | |
| <12 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 12 years | 1.01 | 0.87 | 1.16 | 1.03 | 0.88 | 1.21 | 0.85 | 0.57 | 1.27 |
| 13-16 years | 0.95 | 0.82 | 1.10 | 0.98 | 0.83 | 1.15 | 0.79 | 0.54 | 1.18 |
| 17 and + years | 0.96 | 0.78 | 1.17 | 0.98 | 0.79 | 1.23 | 0.86 | 0.50 | 1.47 |
| **Mother education** | | | | | | | | | |
| <12 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 12 years | 1.09 | 0.94 | 1.26 | 1.09 | 0.93 | 1.28 | 1.09 | 0.74 | 1.61 |
| 13-16 years | 0.95 | 0.82 | 1.10 | 0.94 | 0.80 | 1.10 | 1.05 | 0.70 | 1.57 |
| 17 and + years | 0.92 | 0.74 | 1.14 | 0.92 | 0.73 | 1.16 | 0.96 | 0.55 | 1.67 |

[1] Adjusted for child's race, birth order, birth weight, and mother's age

## Chapter 7

# Factors associated with residential mobility of childhood leukemia cases: a population-based study in California

Oksuzyan S[1], Crespi CM[3], Cockburn M[2], Mezei G[4], Vergara X[5], Kheifets L[6]

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, soksuzyan@ucla.edu; [2]Department of Preventive Medicine, University of Southern California, Los Angeles, California, myles@med.usc.edu; [3]Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California, ccrespi@ucla.edu; [4]Environment Department, Electric Power Research Institute, Palo Alto, California, gmezei@epri.com; [5]Environment Department, Electric Power Research Institute, Palo Alto, California, xvergara@epri.com; [6]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, kheifets@ucla.edu.

**Abstract**

*Aims:* We conducted a large registry-based study in California to evaluate socio-demographic, perinatal factors and the distance to the nearest power line that might be associated with the residential mobility in childhood leukemia cases in California.

*Methods:* We obtained information on 5788 cases by linking California cancer and birth registries. The residential mobility was defined based on a distance between geocoded latitude and longitude of birth and diagnosis addresses. The primary analysis method was logistic regression with the binary outcome for the residential mobility. The secondary analysis was multinomial logistic regression.

*Results:* The odds of moving before diagnosis increased strongly with child's age at diagnosis and decreased with maternal age (p-value for trend < 0.0001 for both). Increased odds of moving was detected for Black non-Hispanic race/ethnicity compared to that of White non-Hispanic cases of childhood leukemia and ALL (OR=1.73, 95% CI: 1.11-2.67 and OR=1.88 , 95% CI: 1.09-3.22, respectively). For AML cases of Black non-Hispanic race/ethnicity the odds of moving was more than tripled (OR=3.47, 95% CI: 1.33-9.08). The odds of moving was lower for childhood leukemia cases with higher census-based SES and with private insurance as a source of payment for delivery. In multinomial analysis with 3-category residential mobility increased odds of moving outside of birth neighborhood was found for children who lived very close (< 100 meters) to the nearest power line; associations for the majority of other factors were also more pronounced for the move outside of a neighborhood compared to those who did not move.

*Conclusions*: Overall, the residential mobility of childhood leukemia cases notably varied by child's age and race/ethnicity, maternal age at birth, census-based SES, and the source of payment for delivery; it varied less by a distance to the nearest power line. Our results suggest that even if information on the residential mobility is unavailable, it might be possible and important to examine the distribution of factors that could be associated with the residential mobility and control for them.

**Key words:** childhood leukemia, residential mobility, electromagnetic fields, father's education, source of payment for delivery, census-based SES, perinatal factors

**Introduction**

Many studies examining the association between environmental exposures and childhood leukemia have used residential addresses of study subjects to evaluate risks associated with a proximity of a child's residence to sources of electro-magnetic fields (EMF) (1-4), air pollution (5), traffic emissions (6), or pesticides.(7) The majority of these studies used a single residential address of the child (e.g., birth, diagnosis, longest lived) and did not account for a residential mobility of subjects, which could result in misclassification of exposure and biased results. Misclassification can occur not only because more relevant exposure may have occurred at a different address but also because of potentially different effects of exposure during critical time windows of the child's development.(8)

In addition, the residential mobility might be related to increased exposure to viruses or other infections which could be associated with higher leukemia risk.(9) The distance of move (e.g. within and outside of a neighborhood) could be an indicator for exposure to new infections.

One of the proxies for EMF exposure frequently used in epidemiological research is a distance to overhead power line(s). Calculated fields commonly used in epidemiological research are also based on a distance to overhead power line(s). Not accounting for the residential mobility may results in misclassification of exposure to EMF: if only birth address is available and 'exposed' subjects moved after birth further from power line(s), we might overestimate their exposure; and we might underestimate EMF exposure if 'unexposed' subjects moved closer to power line(s) after birth. This may lead to biased results, particularly if the residential mobility differs for cases and controls (differential misclassification of exposure). In several studies on EMF cases tended to be more residentially mobile than controls prior to diagnosis.(10, 11)

Unfortunately, in many studies only one residential address is available, and thus directly assessing the residential mobility is impossible. In such situations, it may be possible to evaluate the relationships between childhood leukemia and factors thought to be associated with the residential mobility, but not mobility per se.

We conducted a large epidemiologic case-control study in California to examine the association of childhood leukemia and a distance from the birth address to the nearest high voltage overhead transmission line as a replication of the study of Draper et al. (2005) in the United Kingdom. (12) Detailed description of the study design, methods of case ascertainment,

control selection, exposure assessment and data analysis were presented in our previous publication. (13)

In this study both birth and diagnosis addresses were available for cases only. The aims of this analyses were to evaluate socio-demographic, perinatal factors and a distance to the nearest power line that might be associated with the residential mobility in childhood leukemia cases in California.

**Materials and methods**

Eligible cases included in this analysis were childhood leukemia cases diagnosed between 1988 and 2008 in children younger than 16 years who were born in California and who resided in California at the time of diagnosis. Reporting of incident cancer cases is mandatory in the state of California, and the California Cancer Registry (CCR) is the agency responsible for registering incident cases of cancer. The CCR meets all standards of the Surveillance, Epidemiology and End Results (SEER) Program and North American Association of Central Cancer Registries (NAACCR) for the timely and complete reporting of incident cancer cases in the state; it covers more than 99% of incident cancers. (14) Information on child's age, sex, residence at the time of diagnosis as well as information on cancer types and characteristics was extracted from the CCR. Cancer registry data were linked to the California Birth Registry (CBR; California Department of Public Health, Vital Statistics Branch) which is also over 99% complete. (15) Controls were randomly selected from the CBR and were matched to cases (1:1) on the date of birth ±6 months and sex. Details regarding selection of controls have been described in previously published papers. (13, 16) Since we did not have information on "pseudo-diagnosis" address of controls, we performed case-only analysis.

The CBR provided information on socio-demographic and perinatal factors of study subjects, including mother's residential address at time of birth, child's date of birth, sex, child's race and ethnicity, birth weight, birth order, number of live births living, parental ages, parental education, parental race and ethnicity, and source of payment for delivery. The specific variables collected by the CBR changed over time period of the study: certain variables (e.g., maternal and paternal education) were collected only in certain years or their classification was changed. (16)

A total of 5788 childhood leukemia cases and 5788 matched controls were available for this analysis. Birth and diagnosis addresses of cases were geocoded using the University of Southern California (USC) Geographic Information Systems (GIS) Laboratory's open-source geocoder, which uses parcel level data for Los Angeles County and street level data for the whole of California. Details on the geocoder, reference data, and procedures can be found in our previous publication.(17) For the purposes of the current study, we used geocoded latitude and longitude, and the geographic matching feature. The geographic matching feature is the geographic entity to which the address was resolved, which depended on the completeness of the address. The geographic matching features were tax-assessor parcel (the boundary of an inhabited property or dwelling) centroid, street segment or street centroid (most often a specific block of a street), US Postal Service ZIP Code Tabulation Area (18), town/city centroid, county centroid, or state centroid for those with unknown addresses.(13) We included in the analysis only cases with geographic matching to parcel and street/street segment, which generally correspond to more precise geocoding; about 78% of all childhood leukemia cases (4505/5788) had such matches for both birth and diagnosis addresses.

We did not have access to residential addresses of cases, but rather only geocoded latitude and longitude of addresses. Hence the residential mobility was defined based on a distance between geocoded latitude and longitude of birth and diagnosis addresses. For distances less than 100 meters, we mapped birth and diagnosis addresses using Google satellite images and visually inspected them to decide whether the subject truly moved or the distance between birth and diagnosis addresses was most probably due to geocoding variations. Sensitivity analysis was performed with various cut points for the definition of a residential move (30, 40, 50, 60, 70 and 100 meters). Based on a visual inspection and a sensitivity analysis, 50 meters was selected as the main cut point. In addition, Sermage-Faure et al. (2013) reported that uncertainty of 50 meters for a geocoded address corresponded to a street segment of a short street level (19), the uncertainty level for geocoded match that we were willing to accept. For the primary analysis, cases with 0 to 50 meters distance between birth and diagnosis addresses were classified as not having moved between birth and diagnosis.

Some factors, e.g., community SES and a proximity to power lines, may change more or less depending on the distance of a move, particularly with moving outside of a neighborhood. Therefore we also performed analyses with three categories for the residential mobility: 1) cases

with distance 0 to 50 meters between birth and diagnosis addresses were classified as not moved, 2) cases with distance between birth and diagnosis addresses 50 to 2000 meters were classified as moved within a neighborhood, 3) cases with distance 2000 meters or further were classified as moved outside of a neighborhood.

We considered the following variables as covariates: birth weight, gestational age, a number of live births, census-based socio-economic status (SES), race/ethnicity, mother's and father's age, mother's and father's education, mother's place of birth, the payment source for delivery, and the proximity to high voltage power lines.

Perinatal factors such as birth weight, gestational age, and a total number of live births were modeled as categorical variables.

We used parental education and the source of payment for delivery as proxies for SES at an individual level. Maternal educational attainment (in years) was only collected in certain years and missing in more than 50% of subjects; therefore, we used father's educational attainment as SES proxy. It was categorized into four levels: <12 years (less than high school), 12 years (high school), 13-16 years (some college, college), and 17 years and more (graduate school). Sources of payment for delivery were categorized into two categories corresponding to low (governmental programs) and middle-high SES (Blue Cross/Blue Shield, private insurance and all others). Detailed description of father's education and the source of payment for delivery are available in our previous publication. (20)

A measure of census-based SES was provided by our collaborators from USC. The variable was derived from U.S. Census data using principal components analysis based on seven indicator variables at a census block group level: education index developed by Liu et al. (21) that weights the proportion of people with a given level of education by a number of years need to attain that level, proportion with a blue-collar job, proportion older than 16 without a job, median household income, proportion below 200% of the federal poverty level, median rent, median home value. Principal component scores were than categorized into and provided quintiles.(22) Census-based SES at birth was used as one of covariates in models. Census-based SES at birth and SES at diagnosis were compared in a total sample of cases and by residential mobility categories ("did not move" and "moved").

We combined race and Hispanic ethnicity and constructed an 8-category variable for mixed race and ethnicity of child, mother and father consisting of the following combinations of

race and ethnicity: Non-Hispanic White, Hispanic White, Non-Hispanic Black, Hispanic Black, Non-Hispanic Asian, Hispanic Asian, Non-Hispanic Other and Hispanic Other. More detailed information about classification of child's race and ethnicity is available in our previous publication. (23)

Some studies showed that mother's place of birth outside of U.S. was associated with increased mobility of the family. (8) We dichotomized mother's place of birth as U.S. born and non-U.S. born.

A proximity to power lines was defined as a distance from birth address to the nearest power line. The distance from the birth home to nearby overhead transmission lines was ascertained based on GIS databases of main four electric power companies in California. We evaluated distances to power lines up to 2000 m and, when possible we included consideration of lower voltages (60 - 69 kV). (13) The distance was classified into 5 categories: 0-100 meters, 100-199 meters, 200-599 meters, 600-1999 meters, and no lines within 2000 meters. Distances to the nearest power line from birth geocoded address and from diagnosis geocoded address where compared in the total sample of cases and by residential mobility categories. Besides considering a distance to the nearest power line of any voltage, we also performed analysis for a distance to the nearest power line with voltage 200 kV and above.

Statistical analysis

The primary analysis method was logistic regression with the binary outcome for residential mobility, i.e., moved versus did not move between birth and diagnosis. We confirmed the absence of unduly influential observations by fitting a variety of models with different subsets of covariates and examining the results for outliers and influence. Covariates in models were chosen a priori and based on information about known or potential confounders. Main subtypes of childhood leukemia were included in one of the models as a covariate to explore if there was a difference in relationships between the two main subtypes (ALL and AML) and the residential mobility.

The secondary analysis was conducted using multinomial logistic regression with 3-category residential mobility (did not move, moved within neighborhood and moved outside of neighborhood) as an outcome variable using "did not move" as a reference. Adjustment variables

for these models were the same as for the model with dichotomous residential mobility as the outcome.

Analyses were conducted using statistical software SAS 9.3.(24)

This study was approved by University of California, Los Angeles Office for the Protection of Research Subjects.

**Results**

A total of 6645 eligible leukemia cases were identified from the California Cancer Registry. Linkage to birth records was successful for 87.1% (N=5788) of cases. Of the 5788 cases, 4505 were matched at parcel and street segment levels for both birth and diagnosis addresses and were included in analyses. 81.6 % of these cases were acute lymphocytic leukemia (ALL) cases (n=3675), 14.6% were acute myeloid leukemia (AML) cases (n=658), and 3.8% were other childhood leukemia types (n=172). Of the 4505 cases, 2610 (58%) moved between birth and diagnosis; 1895 (42%) did not move birth and diagnosis according to our definition of move. Among those who moved, the majority (n=1992) moved outside their birth neighborhood, i.e., more than 2000 meters away; only 618 moved within their birth neighborhood (< 2000 meters). The median age at diagnosis was 3.8 years with a range of 0 to 15.6 years, with the peak for ALL between 2 and 4 years of age and between 0 and 2 years of age for AML. Additional characteristics of study subjects are presented in Table 1.

As a part of the descriptive analysis, we ran unadjusted logistic regression models for the residential mobility for the factors in Table 1. Results of this analysis revealed that child's age, Black non-Hispanic race/ethnicity, middle-higher and highest census-based SES, higher paternal education, private health insurance, closer proximity to power lines, advanced maternal age, number of live births living, and mother's place of birth were associated with the residential mobility; no association was detected for sex, birth weight and gestational age (results not presented).

Results of adjusted logistic regression analyses are presented in the Table 2. The odds of moving before diagnosis increased strongly with age at diagnosis; the higher was age the higher

was the odds of moving (p-value for trend < 0.0001). The odds of moving was slightly higher for cases who lived <100 m from power lines at birth compared to those who did not have power line(s) within 2000 meters, but the estimate was imprecise. The odds of moving was almost double for Black non-Hispanic race/ethnicity compared to that of White non-Hispanic cases of childhood leukemia and ALL. For AML cases of Black non-Hispanic race/ethnicity compared to that of White non-Hispanic, the odds of moving from the birth address was more than tripled. For total childhood leukemia cases and ALL, the odds of moving decreased with maternal age (p-value for trend < 0.0001). The odds of moving was lower for total childhood leukemia cases, ALL and AML whose source of payment for delivery was private insurance and other sources compared to that of cases with governmental insurance.

We also fit a model with census-based SES. The model was the same as in Table 2, but individual level SES proxies (father's education and source of payment for delivery) were replaced by census-based SES. For cases of childhood leukemia of middle-higher and highest census-based SES, the odds of moving was lower than that of cases of the lowest census-based SES (OR=0.78, 95% CI: 0.61-0.99 and OR=0.63, 95% CI: 0.49-0.82, respectively) with p-value for trend 0.0002. For ALL and AML cases, only the highest census-based SES was associated with decreased odds of moving (OR=0.67, 95% CI: 0.50-0.89 and OR=0.41, 95% CI: 0.18-0.92, respectively). Estimates for all other variables were similar to those in Table 2.

Analysis with high voltage overhead power lines (≥200 kV) showed that the odds of moving was slightly higher for ALL cases who lived <100 m from power lines at birth compared to those who did not have power line(s) within 2000 meters, but the estimate was imprecise. Table 3 presents results only for the nearest distance to high voltage power line. For all other factors associations were similar to those in Table 2.

Results of multinomial (3-category) logistic regression with outcome categories: did not move, moved between 50 and 1999 meters (within neighborhood) and moved 2000 meters and further (outside of neighborhood) are presented in Table 4.

As shown in Table 4, child's age was associated with increased odds of moving from a birth residence both within and outside of a neighborhood (within and outside of 2000 meters) for total childhood leukemia and main two subtypes (results not presented). The association was much stronger for children who moved outside of 2000 meter area (outside of a neighborhood). In contrast to child's age, advanced mother's age compared to < 25 years olds was associated with decreased odds of moving both within and outside of a neighborhood for childhood leukemia and ALL (results not presented); the inverse association was stronger for those who moved outside of a neighborhood. For childhood leukemia cases and main two subtypes (ALL and AML), White Hispanic race/ethnicity was associated with increased odds of moving within a neighborhood; Black non-Hispanic race/ethnicity was associated with increased odds of moving outside of a neighborhood (ref. - White non-Hispanic). Private insurance and other sources of payment for delivery were associated with decreased odds of moving both within and outside of a neighborhood compared to cases with governmental health insurance programs as a source of payment for delivery. Living in close proximity to a power line (within 100 meters) was associated with increased odds of moving outside of a neighborhood for cases of total childhood leukemia and ALL.

In the model with census-based SES instead of father's education and source of payment for delivery, highest SES was associated with decreased odds of moving both within and outside of a neighborhood compared to lowest SES (OR=0.67 ; 95% CI: 0.45-0.99 and OR=0.63 ; 95% CI: 0.47-0.83, respectively), for total childhood leukemia cases. For ALL and AML cases, the association for the highest SES was observed only for moving outside of a neighborhood (results not shown). A proximity to power lines was not associated with increased odds of moving. Estimates for other variables were similar to the model with individual level SES proxies (results not presented).

We were also interested in assessing if either of the main two subtypes of childhood leukemia was associated with the residential mobility more than the other, adjusted for other variables in the model. These results showed that AML was associated with decreased odds of moving from birth address compared to ALL, although the association was not precise (OR=0.85, 95% CI: 0.67-1.08).

Results of comparison of census-based SES at birth and at diagnosis and of distances to the nearest power line from birth and from diagnosis geocoded addresses are presented in Table 5. Correlations between census-based SES at birth and at diagnosis and between distances to the nearest power line from birth geocoded address and from diagnosis geocoded address was good (Spearman r=0.7, kappa=0.5 and Spearman r=0.6, kappa=0.6, respectively).

As shown in Table 5 approximately half of cases (49.2%) were in the same census-based SES quintile at diagnosis as they were at birth; 29.1% were in a higher SES quintile and 21.7% were in a lower SES quintile at diagnosis (Table 5). A change in census-based SES from birth to diagnosis could be explained either by move from birth residence (65%) or by changes in factors used to calculate quintiles for census-based SES over the years (35%), i.e., some census tracts changed SES quintile over the years.

The majority of cases (69.6%) were at the same category for a distance to the nearest power line at birth and at diagnosis geocoded addresses; same percentage of cases (15.2 %) had a longer or a shorter distance from birth address then from diagnosis address. One case in "did not move" category appeared as "moved" since it had a shorter distance to the nearest power line from birth then from diagnosis address. Based on aerial images we concluded that this discrepancy was most probably due to geocoding error of a birth address.

**Discussion**

Our study on the residential mobility of childhood leukemia cases showed that cases were fairly mobile, with 58% of them having changed their residence between birth and diagnosis. Our results are very similar to those from a small pilot study conducted in San Diego, California where about 55% of children less than 5 years of age with childhood leukemia had moved from their birth residence.(6) Another study in Northern California (Urayama et al, 2009) also observed that a high percentage of cases (65.8%) moved at least once between birth and diagnosis.(8) Alexander et al. (1993) conducted a study in Northern England and found that 30.8% of children with leukemia and non-Hodgkin’s lymphoma had changed residence by the

age of 2. The authors of the Northern England study acknowledged that their study population was likely to be less mobile than the general population due to the subject selection process. The following factors were associated with the residential mobility in our study, regardless of the model used: child's age and race/ethnicity, census-based SES, source of payment for delivery, and mother's age. Children diagnosed at older ages had higher odds of moving from their birth residence before diagnosis, as would be expected. Higher census-based SES and older maternal age were associated with decreased odds of moving from the birth residence with significant linear trends. These findings for child's and maternal age and SES were similar to those in another Californian study by Urayama et al. but with a more pronounced effect of child's age and census-based SES. (8) In contrast to the afore-mentioned study, we found that Black non-Hispanic race/ethnicity was associated with higher odds of moving in all our analyses, and this association persisted after controlling for SES. The lack of association between race/ethnicity and mobility of childhood leukemia cases in the study by Urayama et al. could be due to small sample size (n=380), potential selection bias (interview-based study with incomplete participation of cases), and/or a different categorization of race/ethnicity. We also detected that the source of payment for delivery was associated with moving among childhood leukemia cases. Decreased odds of moving in children with a private insurance or other sources of payment for delivery compared to governmental health programs is in line with our findings for census-based SES since people with private and other health insurances tend to be higher SES compared to those with governmental health programs.

Analysis with the nearest power line with high voltage, i.e. $\geq$ 200 kV showed a tendency toward increase in odds of moving for children who lived < 100 meters to the nearest power line only for ALL cases, but the association was imprecise.

In analysis with 3-category residential mobility (none, within and outside of birth neighborhood), we noted differences between those who moved within their birth neighborhood compared to those who moved outside of it. Increased odds of moving outside of birth neighborhood ( $\geq$ 2000 meters) was found for children who lived very close (< 100 meters) to the nearest power line. Associations for the majority of factors that were associated with the residential mobility were more pronounced for move outside of a neighborhood compared to those who did not move.

We performed an exploratory analysis to understand how the association of childhood leukemia and a distance to the nearest power line at birth would change depending on mobility of subjects, if there was an association. (results not shown) There were four possible scenarios: 1) childhood leukemia risk was associated with the distance to the nearest power line at birth, residential mobility was not associated with the risk of childhood leukemia and the residential mobility was not association with distance to power line, 2) childhood leukemia risk was associated with the distance to the nearest power line, the residential mobility was not associated with the risk of childhood leukemia and the residential mobility was association with distance to power line at birth, 3) childhood leukemia risk was associated with the distance to the nearest power line, residential mobility was associated with the risk of childhood leukemia and the residential mobility was not association with distance to power line, and 4) childhood leukemia risk was associated with the distance to the nearest power line, the residential mobility was associated with the risk of childhood leukemia and the residential mobility was association with distance to power line. By our study design controls were more mobile residentially than cases (one of inclusion criteria for cases was being a resident of CA at the time of diagnosis, while for controls there was no such criterion), i.e. we were missing cases that moved out of California and were diagnosed somewhere else. Taking this fact into consideration, under the 2nd and the 4th scenarios the association of childhood leukemia and the distance to the nearest power line at birth address would be underestimated. (results not shown) Under the 2nd scenario the residential mobility acts as an intermediate and under the 4th scenario it is a confounder. The relationships between childhood leukemia and distance to the nearest power line at birth address would not be bias under the 1st and the 3rd scenarios since the residential mobility in these situations is neither a confounder nor an intermediate.

We did not assess the nearest distance to a power line from diagnosis address. However, the time period of EMF exposure is relevant for childhood leukemia risk. Since the peak of childhood leukemia incidence is at very early years of child's life (between 2 and 5 years of age for ALL and between 0 to 2 years for AML), birth address is probably more relevant to exposure than diagnosis address.

We did not find associations between father's education, birth weight, gestational age, birth order or total number of live births with the residential mobility (some results are not shown).

One strength of our study was that data were obtained from population registries. Unlike studies that use interview or questionnaires for their data collection, since participation of subjects was not required for our data collection, selection bias due to differences in participation was improbable. However, to increase accuracy we excluded from our analysis about 22% of our initial sample that had imprecise geocode match (e.g. at zip code or state levels). This could have resulted in selection bias. To assess whether selection bias could be an issue, we compared excluded and included cases on covariates. We found that excluded cases had slightly higher father's education level and census-based SES, slightly lower maternal age at birth and higher percentage of governmental programs as a source of payment for delivery. The difference in proportions was very small, therefore, we do not think it would considerably affect our results.

Another limitation of our study was potential misclassification of the outcome (the residential mobility). We did not have residential history nor exact addresses for cases; we had only the geocoded latitude and longitude coordinates of the birth residence and diagnosis residence. We defined the residential mobility by distance between the geocoded points of birth and diagnosis addresses of cases. Although some misclassification was inevitable, we believe it was small since we manually investigated, mapped and visually inspected all distances between birth and diagnosis residences that were < 100 meters. Based on our visual inspection and geocoding accuracy considerations, we considered a 50 meters cut point an optimal one to decide whether a case moved or not. As a sensitivity analysis we also performed analysis with different cut points for move (30, 40, 60, 70, 100 meters); results did not differ much (see Table A in Appendix). In the analysis with 3-category the residential mobility (did not move, moved within and outside of neighborhood) misclassification of a distance used as the cut point for moving outside of a neighborhood could also be an issue. We were particularly interested in cut point of 2000 meters since in our study on EMF and childhood leukemia we used 2000 meters buffer around a geocoded residence location as the area within which we identified power lines. We have also tried other cut points (3000 and 4000 meters) and did not see any important differences in results (not presented).

Misclassification of covariates was possible, but unlikely in this study. We are confident that child's age at diagnosis, maternal age at birth and source of payment for delivery were recorded quite accurately in CCR and CBR, since the first two are among the main factors affecting treatment of cancer and delivery management and complications and the later one is a

variable of financial interest for the reporting medical facility. A validation study in California has shown that child's race and ethnicity recorded on birth certificates was more than 94% accurate. (25) The distance to the nearest power line provided by main four electric companies in California was validated using Google Earth aerial photos; the Pearson correlation between the two was greater than 0.99. We also have ongoing work to refine the distance measures based on site visits. (13)

The strength of this study was a large sample size with 4505 cases included in the analysis and statewide inclusion of cases. One previous study on the residential mobility in California by Urayama et al. (2009) had a much smaller sample size (n=380), included children only from the northern part of California and was interview-based, leaving potential for selection bias. The large sample size allowed us to look at associations of different factors with the residential mobility for subtypes of leukemia, which was not possible in previous research. Results of analysis by subtypes showed similar association for both subtypes between the residential mobility and child's age, race/ethnicity, source of payment for delivery, and census-based SES. The difference was that in ALL but not AML cases, the residential mobility was also associated with mother's age and the proximity to power lines. We also found that AML was associated with slightly decreased odds of moving compared with ALL, but the estimate was imprecise.

Despite the large sample size, in some analyses the analytic sample was reduced due to missing data. Since information was missing mainly due to differences between years in the collected information on birth certificates rather than non-response, and did not differ between movers and non-movers, the potential for biases was probably small, and the impact was mainly on the precision of the estimates.

In summary, the results of this study indicate that childhood leukemia cases in California are residentially mobile, with 58% of them moving between birth and diagnosis. The residential mobility of childhood leukemia cases notably varied by time interval between birth and diagnosis, child's race/ethnicity, maternal age at birth, census-based SES, and the source of payment for delivery; it varied less by the distance to the nearest power line. Our results suggest that even if information on the residential mobility of subjects is unavailable, it might be possible and important to examine the distribution of factors that could be associated with the residential mobility and control for them.

**Acknowledgements**

This project was supported by a research contract from the Electric Power Research Institute to UCLA and by UCLA Faculty Grants Program. Crespi was also partially supported by National Institutes of Health grant P30 CA16042. The authors want to thank California Department of Public Health for providing support and access to birth registry.

**References**

1. Ahlbom A, Day N, Feychting M, Roman E, Skinner J, Dockerty J, et al. A pooled analysis of magnetic fields and childhood leukaemia. Br J Cancer. 2000 Sep;83(5):692-8. PubMed PMID: 10944614.
2. Greenland S, Sheppard AR, Kaune WT, Poole C, Kelsh MA. A pooled analysis of magnetic fields, wire codes, and childhood leukemia. Childhood Leukemia-EMF Study Group. Epidemiology. 2000 Nov;11(6):624-34. PubMed PMID: 11055621.
3. Kheifets L, Ahlbom A, Crespi CM, Draper G, Hagihara J, Lowenthal RM, et al. Pooled analysis of recent studies on magnetic fields and childhood leukaemia. Br J Cancer. 2010 Sep 28;103(7):1128-35. PubMed PMID: 20877339. Pubmed Central PMCID: 2965855.
4. Kheifets L, Shimkhada R. Childhood leukemia and EMF: review of the epidemiologic evidence. Bioelectromagnetics. 2005;Suppl 7:S51-9. PubMed PMID: 16059924.
5. Raaschou-Nielsen O, Reynolds P. Air pollution and childhood cancer: a review of the epidemiological literature. Int J Cancer. 2006 Jun 15;118(12):2920-9. PubMed PMID: 16425269.
6. Reynolds P, Elkin E, Scalf R, Von Behren J, Neutra RR. A case-control pilot study of traffic exposures and early childhood leukemia using a geographic information system. Bioelectromagnetics. 2001;Suppl 5:S58-68. PubMed PMID: 11170118.
7. Alavanja MC, Ward MH, Reynolds P. Carcinogenicity of agricultural pesticides in adults and children. Journal of agromedicine. 2007;12(1):39-56. PubMed PMID: 18032335.
8. Urayama KY, Von Behren J, Reynolds P, Hertz A, Does M, Buffler PA. Factors associated with residential mobility in children with leukemia: implications for assigning exposures. Ann Epidemiol. 2009 Nov;19(11):834-40. PubMed PMID: 19364662. Pubmed Central PMCID: 2761989.
9. Sahl JD. Viral contacts confound studies of childhood leukemia and high-voltage transmission lines. Cancer Causes Control. 1994 May;5(3):279-83. PubMed PMID: 7710486.
10. Green LM, Miller AB, Villeneuve PJ, Agnew DA, Greenberg ML, Li J, et al. A case-control study of childhood leukemia in southern Ontario, Canada, and exposure to magnetic fields in residences. Int J Cancer. 1999 Jul 19;82(2):161-70. PubMed PMID: 10389746.
11. McBride ML, Gallagher RP, Theriault G, Armstrong BG, Tamaro S, Spinelli JJ, et al. Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. Am J Epidemiol. 1999 May 1;149(9):831-42. PubMed PMID: 10221320.
12. Draper G, Vincent T, Kroll ME, Swanson J. Childhood cancer in relation to distance from high voltage power lines in England and Wales: a case-control study. Bmj. 2005 Jun 4;330(7503):1290. PubMed PMID: 15933351.

13. Kheifets L, Crespi, CM, Hooper, C, Oksuzyan, S, Cockburn, M, Ly, T, Mezei, G, . Epidemiologic study of residential proximity to transmission lines and childhood cancer in California: Description of design, epidemiologic methods and study population. Journal Of Exposure Science And Environmental Epidemiology. 2013, in print.
14. The California Cancer Registry. Mission Statement & Purpose Sacramento, CA [March 2009]. Available from: http://www.ccrcal.org/abouttheccr.html.
15. Schoendorf KC, Branum AM. The use of United States vital statistics in perinatal and obstetric research. American journal of obstetrics and gynecology. 2006 Apr;194(4):911-5. PubMed PMID: 16580275.
16. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Kheifets L. Birth weight and other perinatal characteristics and childhood leukemia in California. Cancer epidemiology. 2012 Dec;36(6):e359-65. PubMed PMID: 22926338.
17. Goldberg D, Cockburn, M,. Improving Geocode Accuracy with Candidate Selection Criteria. Transactions in GIS. 2010;14(1):149-76.
18. United States Census Bureau. Geography: Zip Code Tabulation Areas 2012 [cited 2012 April 2]. Available from: http://www.census.gov/geo/reference/zctas.html.
19. Sermage-Faure C, Demoury C, Rudant J, Goujon-Bellec S, Guyot-Goubin A, Deschamps F, et al. Childhood leukaemia close to high-voltage power lines--the Geocap study, 2002-2007. Br J Cancer. 2013 May 14;108(9):1899-906. PubMed PMID: 23558899. Pubmed Central PMCID: 3658518.
20. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Vergara X, Kheifets L. Socio-economic status and childhood leukemia in California. Journal of Hematological Malignancies. 2013;(submitted).
21. Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). Cancer Causes Control. 1998 Aug;9(4):369-80. PubMed PMID: 9794168.
22. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control. 2001 Oct;12(8):703-11. PubMed PMID: 11562110.
23. Oksuzyan S, Cockburn M, Crespi CM, Mezei G, Vergara X, Kheifets L. Race/ethnicity and childhood leukemia: a case-control study in California. Cancer Causes and Control. 2013, submitted Aug. 25, 2013.
24. SAS Institute I. SAS Statistical Software. 9.1 ed. Cary, NC2006.
25. Baumeister L, Marchi K, Pearl M, Williams R, Braveman P. The validity of information on "race" and "Hispanic ethnicity" in California birth certificate data. Health services research. 2000 Oct;35(4):869-83. PubMed PMID: 11055453. Pubmed Central PMCID: 1089157.

**Table 1.** Socio-demographic and other characteristics of childhood leukemia cases, California birth and cancer registries, 1986-2007.

| Variables | Did not move | Row % | Moved | Row % |
|---|---|---|---|---|
| **Leukemia main subtypes** | | | | |
| ALL | 1503 | 40.90 | 2172 | 59.10 |
| AML | 303 | 46.05 | 355 | 53.95 |
| Other | 89 | 51.74 | 83 | 48.26 |
| **Child's sex** | | | | |
| Male | 1033 | 41.75 | 1441 | 58.25 |
| Female | 862 | 42.44 | 1169 | 57.56 |
| **Child's age** | | | | |
| < 1 years | 239 | 79.14 | 63 | 20.86 |
| 1-5 years | 1362 | 46.69 | 1555 | 53.31 |
| 6-9 years | 205 | 27.52 | 540 | 72.48 |
| 10-15 years | 89 | 16.45 | 452 | 83.55 |
| **Child's race** | | | | |
| Non-Hispanic White | 633 | 48.84 | 663 | 51.16 |
| Hispanic White | 859 | 38.55 | 1369 | 61.45 |
| Non-Hispanic Black | 53 | 27.6 | 139 | 72.4 |
| Hispanic Black | 11 | 28.21 | 28 | 71.79 |
| Non-Hispanic Asian | 233 | 51.43 | 220 | 48.57 |
| Hispanic Asian | 17 | 36.17 | 30 | 63.83 |
| Non-Hispanic Other | 19 | 47.5 | 21 | 52.5 |
| Hispanic Other | 9 | 31.03 | 20 | 68.97 |
| Missing | 61 | | 120 | |
| **Father's education** | | | | |
| <12 years | 791 | 39.93 | 1190 | 60.07 |
| 12 years | 224 | 47.16 | 251 | 52.84 |
| 13-16 years | 260 | 52.85 | 232 | 47.15 |

| | | | | |
|---|---|---|---|---|
| ≥17 years | 122 | 61.62 | 76 | 38.38 |
| Missing* | 498 | | 861 | |
| **Source for payment for delivery** | | | | |
| Governmental programs | 551 | 33.21 | 1108 | 66.79 |
| Other insurance | 1110 | 51.03 | 1065 | 48.97 |
| Missing* | 234 | | 437 | |
| **Census-based SES** | | | | |
| Lowest | 392 | 38.39 | 629 | 61.61 |
| Lower-Middle | 349 | 40.11 | 521 | 59.89 |
| Middle | 294 | 41.18 | 420 | 58.82 |
| Middle-higher | 265 | 48.89 | 277 | 51.11 |
| Highest | 300 | 58.37 | 214 | 41.63 |
| Unknown | 295 | | 549 | |
| **Mother's place of birth** | | | | |
| U.S. | 1058 | 43.41 | 1379 | 56.59 |
| Non-U.S. | 837 | 40.47 | 1231 | 59.53 |
| **Mother's age** | | | | |
| < 25 years | 428 | 30.1 | 994 | 69.9 |
| 25-34 years | 1054 | 44.06 | 1338 | 55.94 |
| ≥35 years | 413 | 59.86 | 277 | 40.14 |
| Missing | | | 1 | |
| **Number of live birth living** | | | | |
| 1 | 658 | 38.46 | 1053 | 61.54 |
| 2 | 630 | 44.15 | 797 | 55.85 |
| 3 | 332 | 43.63 | 429 | 56.37 |
| 4+ | 275 | 45.38 | 331 | 54.62 |
| **Proximity of birth address to nearest high voltage power line** | | | | |
| ≥ 2000 m | 865 | 42.86 | 1153 | 57.14 |
| 600-1999 m | 675 | 43.19 | 888 | 56.81 |
| 200-599 m | 217 | 38.14 | 352 | 61.86 |

| | | | | |
|---|---|---|---|---|
| 100-199 m | 70 | 40.46 | 103 | 59.54 |
| < 100 m | 68 | 37.36 | 114 | 62.64 |

* Patterns of missingness varied by year due to differences in data collection for different years.

**Table 2**. Adjusted odds ratios for the residential mobility and selected socio-demographic and other factors in childhood leukemia cases, California birth and cancer registries, 1986-2007.

| Variables | Total | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | **OR** | **95% lower CI** | **95% upper CI** | **OR** | **95% lower CI** | **95% upper CI** | **OR** | **95% lower CI** | **95% upper CI** |
| **Child's age** | | | | | | | | | |
| < 1 year | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 1-5 years | 4.53 | 3.24 | 6.34 | 4.89 | 2.95 | 8.10 | 3.04 | 1.69 | 5.44 |
| 6-9 years | 10.60 | 7.19 | 15.63 | 10.43 | 6.04 | 18.01 | 13.21 | 5.32 | 32.81 |
| 10-15 years | 19.59 | 12.38 | 31.01 | 19.19 | 10.36 | 35.55 | 19.84 | 7.80 | 50.46 |
| p-value for trend | <0.0001 | | | | | | | | |
| **Child's race/ethnicity combined** | | | | | | | | | |
| White Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| White Hispanic | 1.13 | 0.93 | 1.38 | 1.03 | 0.83 | 1.28 | 2.18 | 1.20 | 3.94 |
| Black Non-Hispanic | 1.73 | 1.11 | 2.67 | 1.88 | 1.09 | 3.22 | 3.47 | 1.33 | 9.08 |
| Black Hispanic | 1.07 | 0.48 | 2.35 | 0.97 | 0.40 | 2.36 | 1.18 | 0.15 | 9.35 |
| Asian Non-Hispanic | 0.95 | 0.72 | 1.25 | 0.89 | 0.66 | 1.20 | 1.39 | 0.65 | 2.97 |
| Asian Hispanic | 0.99 | 0.49 | 2.01 | 1.03 | 0.44 | 2.37 | 2.13 | 0.46 | 9.99 |
| Other Non-Hispanic | 0.80 | 0.35 | 1.84 | 0.87 | 0.35 | 2.15 | 0.63 | 0.05 | 8.24 |
| Other Hispanic | 1.14 | 0.40 | 3.23 | 0.97 | 0.32 | 2.92 | 2.26 | 0.11 | 47.32 |
| **Source of payment for delivery** | | | | | | | | | |
| Governmental programs | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Private insurance and other sources | 0.61 | 0.51 | 0.73 | 0.61 | 0.50 | 0.74 | 0.59 | 0.36 | 0.98 |
| **Father's education** | | | | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| < 12 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 12 years | 1.01 | 0.81 | 1.26 | 1.08 | 0.85 | 1.37 | 0.69 | 0.36 | 1.33 |
| 13-16 years | 1.06 | 0.85 | 1.33 | 1.05 | 0.82 | 1.34 | 0.91 | 0.45 | 1.84 |
| ≥ 17 years | 0.84 | 0.60 | 1.18 | 0.82 | 0.57 | 1.18 | 1.17 | 0.45 | 3.07 |
| **Mother's age** | | | | | | | | | |
| < 25 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 25-34 years | 0.68 | 0.56 | 0.83 | 0.63 | 0.51 | 0.79 | 2.18 | 1.20 | 3.94 |
| ≥35 years | 0.41 | 0.31 | 0.54 | 0.37 | 0.27 | 0.49 | 3.47 | 1.33 | 9.08 |
| p-value for trend | <0.0001 | | | | | | | | |
| **Total live birth children** | | | | | | | | | |
| 1 | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 2 | 0.90 | 0.75 | 1.09 | 0.89 | 0.72 | 1.09 | 1.02 | 0.59 | 1.77 |
| 3 | 0.95 | 0.75 | 1.20 | 1.00 | 0.78 | 1.30 | 0.64 | 0.32 | 1.26 |
| 4+ | 0.82 | 0.63 | 1.08 | 0.87 | 0.64 | 1.16 | 0.54 | 0.25 | 1.21 |
| **Distance to the nearest power line** | | | | | | | | | |
| No lines within 2000 m | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 600-1999 m | 0.96 | 0.81 | 1.14 | 0.95 | 0.79 | 1.15 | 1.16 | 0.72 | 1.88 |
| 200-599 m | 1.12 | 0.88 | 1.43 | 1.08 | 0.83 | 1.41 | 1.61 | 0.75 | 3.44 |
| 100-199 m | 1.09 | 0.72 | 1.63 | 1.06 | 0.68 | 1.65 | 1.55 | 0.50 | 4.78 |
| < 100 m | 1.44 | 0.97 | 2.13 | 1.52 | 0.98 | 2.33 | 0.51 | 0.13 | 2.08 |

* The variables in the model are adjusted for all other

**Table 3**. Adjusted odds ratios for the residential mobility and a distance to the nearest power line of high voltage (≥200 kV) in childhood cases, California birth and cancer registries, 1986-2007.

| Distance to nearest power line with voltage ≥200 kV | Total | | | ALL | | | AML | | |
|---|---|---|---|---|---|---|---|---|---|
| | OR | 95% lower CI | 95% upper CI | OR | 95% lower CI | 95% upper CI | OR | 95% lower CI | 95% upper CI |
| No lines within 2000 m | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 600-1999 m | 0.98 | 0.83 | 1.16 | 0.97 | 0.80 | 1.17 | 1.20 | 0.74 | 1.93 |
| 200-599 m | 1.06 | 0.83 | 1.37 | 1.02 | 0.78 | 1.34 | 1.59 | 0.74 | 3.44 |
| 100-199 m | 1.24 | 0.80 | 1.94 | 1.22 | 0.76 | 1.97 | 1.79 | 0.49 | 6.55 |
| < 100 m | 1.45 | 0.92 | 2.27 | 1.64 | 0.99 | 2.71 | 0.17 | 0.03 | 1.19 |

[a] Adjusted for child's age and race/ethnicity, mother's age, father's education, source of payment for delivery, and total number of live births

**Table 4**. Adjusted odds ratios from multinomial logistic regression for the residential mobility and select socio-demographic and other factors in childhood leukemia cases. California birth and cancer registries, 1986-2007.

| | Move within neighborhood (50-1999 meters) vs, did not move | | | Move outside of neighborhood (≥2000 meters) vs, did not move | | |
|---|---|---|---|---|---|---|
| **Variables** [a] | **OR** | **95% lower CI** | **95% upper CI** | **OR** | **95% lower CI** | **95% upper CI** |
| **Child's age** | | | | | | |
| < 1 year | 1.00 | - | - | 1.00 | - | - |
| 1-5 years | 2.96 | 1.79 | 4.87 | 5.51 | 3.67 | 8.27 |
| 6-9 years | 7.82 | 4.49 | 13.65 | 12.37 | 7.85 | 19.48 |
| 10-15 years | 9.38 | 4.96 | 17.76 | 26.28 | 15.68 | 44.04 |
| **Source of payment for delivery** | | | | | | |
| Governmental programs | 1.00 | - | - | 1.00 | - | - |
| Private insurance and other sources | 0.53 | 0.41 | 0.69 | 0.64 | 0.52 | 0.77 |
| **Father's education** | | | | | | |
| < 12 years | 1.00 | - | - | 1.00 | - | - |
| 12 years | 1.05 | 0.77 | 1.45 | 1.00 | 0.79 | 1.27 |
| 13-16 years | 0.72 | 0.49 | 1.06 | 1.18 | 0.93 | 1.50 |
| ≥ 17 years | 1.02 | 0.60 | 1.72 | 0.79 | 0.55 | 1.15 |
| **Child's race/ethnicity combined** | | | | | | |
| White Non-Hispanic | 1.00 | - | - | 1.00 | - | - |
| White Hispanic | 1.70 | 1.25 | 2.32 | 0.99 | 0.80 | 1.22 |
| Black Non-Hispanic | 1.27 | 0.62 | 2.60 | 1.86 | 1.18 | 2.92 |
| Black Hispanic | 0.59 | 0.12 | 2.77 | 1.17 | 0.52 | 2.63 |
| Asian Non-Hispanic | 1.02 | 0.65 | 1.62 | 0.93 | 0.69 | 1.24 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Asian Hispanic | 2.10 | 0.84 | 5.28 | 0.73 | 0.33 | 1.62 |
| Other Non-Hispanic | 1.82 | 0.64 | 5.24 | 0.56 | 0.22 | 1.45 |
| Other Hispanic | 1.09 | 0.21 | 5.62 | 1.16 | 0.39 | 3.43 |
| **Mother's age** | | | | | | |
| < 25 years | 1.00 | - | - | 1.00 | - | - |
| 25-34 years | 0.71 | 0.54 | 0.95 | 0.67 | 0.54 | 0.82 |
| ≥35 years | 0.55 | 0.37 | 0.83 | 0.37 | 0.27 | 0.49 |
| **Total live birth children** | | | | | | |
| 1 | 1.00 | - | - | 1.00 | - | - |
| 2 | 0.97 | 0.73 | 1.29 | 0.88 | 0.72 | 1.08 |
| 3 | 1.25 | 0.90 | 1.73 | 0.85 | 0.66 | 1.10 |
| 4+ | 0.84 | 0.56 | 1.24 | 0.82 | 0.62 | 1.10 |
| **Distance from birth residence to nearest power line** | | | | | | |
| No lines within 2000 m | 1.00 | - | - | 1.00 | - | - |
| 600-1999 m | 0.95 | 0.74 | 1.23 | 0.96 | 0.80 | 1.16 |
| 200-599 m | 1.15 | 0.80 | 1.63 | 1.11 | 0.86 | 1.45 |
| 100-199 m | 1.54 | 0.90 | 2.63 | 0.93 | 0.59 | 1.45 |
| < 100 m | 1.11 | 0.61 | 2.04 | 1.55 | 1.03 | 2.34 |

[a] Odds ratios for each variable are adjusted for all other variables.

**Table 5**. Census-based SES and distances to the nearest power line from birth and diagnosis geocoded addresses by residential mobility, California birth and cancer registries, 1986-2007.

| Variable | Did not move | Moved | Total |
|---|---|---|---|
| **Census -based SES** | | | |
| Same SES at birth and at diagnosis | 949 | 852 | 1801 |
| Lower SES at birth then at diagnosis | 369 | 696 | 1065 |
| Higher SES at birth then at diagnosis | 282 | 513 | 795 |
| Total* | 1600 | 2061 | 3661 |
| **Distance to the nearest power line** | | | |
| Same distance from birth and from diagnosis address | 1894 | 1243 | 3137 |
| Longer distance from birth then from diagnosis address | 0 | 683 | 683 |
| Shorter distance from birth then from diagnosis address | 1 | 684 | 685 |
| Total | 1895 | 2610 | 4505 |

*844 cases were missing information on either birth or diagnosis census-based SES

## Appendix

**Table A**. Adjusted odds ratios using different cut points for dichotomous residential mobility and selected socio-demographic and other factors in childhood leukemia cases, California birth and cancer registries, 1986-2007

| Distance for cut point | 30 m | | | 40 m | | | 60 m | | | 70 m | | | 100 m | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variables** | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | | **OR** | **95% CI** | |
| **Child's age** | | | | | | | | | | | | | | | |
| < 1 year | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 1-5 years | 4.53 | 3.25 | 6.33 | 4.61 | 3.29 | 6.45 | 4.62 | 3.30 | 6.48 | 4.61 | 3.29 | 6.47 | 4.53 | 3.23 | 6.35 |
| 6-9 years | 10.62 | 7.22 | 15.63 | 10.87 | 7.37 | 16.04 | 10.92 | 7.40 | 16.12 | 10.85 | 7.35 | 16.02 | 10.72 | 7.27 | 15.82 |
| 10-15 years | 21.10 | 13.28 | 33.53 | 21.02 | 13.24 | 33.39 | 19.99 | 12.64 | 31.60 | 19.59 | 12.41 | 30.92 | 19.70 | 12.48 | 31.10 |
| **Source of payment for delivery** | | | | | | | | | | | | | | | |
| Governmental programs | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| Private insurance and other sources | 0.58 | 0.49 | 0.70 | 0.59 | 0.49 | 0.70 | 0.60 | 0.51 | 0.72 | 0.60 | 0.50 | 0.72 | 0.61 | 0.51 | 0.73 |
| **Father's education** | | | | | | | | | | | | | | | |
| < 12 years | | | | | | | | | | | | | | | |
| 12 years | 1.05 | 0.84 | 1.31 | 1.04 | 0.83 | 1.30 | 1.06 | 0.85 | 1.32 | 1.07 | 0.86 | 1.34 | 1.10 | 0.88 | 1.37 |
| 13-16 years | 1.11 | 0.88 | 1.39 | 1.10 | 0.88 | 1.38 | 1.13 | 0.90 | 1.41 | 1.14 | 0.91 | 1.43 | 1.16 | 0.92 | 1.45 |
| ≥ 17 years | 0.92 | 0.66 | 1.30 | 0.90 | 0.64 | 1.26 | 0.88 | 0.62 | 1.23 | 0.89 | 0.63 | 1.25 | 0.90 | 0.64 | 1.27 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Child's race/ethnicity combined** | | | | | | | | | | | | | | | |
| White Non-Hispanic | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| White Hispanic | 1.16 | 0.95 | 1.41 | 1.15 | 0.94 | 1.40 | 1.15 | 0.94 | 1.39 | 1.14 | 0.94 | 1.39 | 1.13 | 0.92 | 1.37 |
| Black Non-Hispanic | 1.78 | 1.15 | 2.77 | 1.79 | 1.15 | 2.78 | 1.76 | 1.14 | 2.73 | 1.76 | 1.14 | 2.73 | 1.76 | 1.14 | 2.72 |
| Black Hispanic | 1.24 | 0.55 | 2.78 | 1.06 | 0.48 | 2.34 | 1.08 | 0.49 | 2.37 | 0.93 | 0.43 | 2.02 | 0.92 | 0.43 | 2.01 |
| Asian Non-Hispanic | 0.96 | 0.73 | 1.26 | 0.95 | 0.72 | 1.25 | 0.96 | 0.73 | 1.27 | 0.94 | 0.71 | 1.23 | 0.94 | 0.71 | 1.23 |
| Asian Hispanic | 0.99 | 0.49 | 2.01 | 0.99 | 0.49 | 2.01 | 1.01 | 0.50 | 2.06 | 1.02 | 0.50 | 2.06 | 0.90 | 0.44 | 1.82 |
| Other Non-Hispanic | 0.81 | 0.35 | 1.86 | 0.81 | 0.35 | 1.86 | 0.82 | 0.36 | 1.89 | 0.83 | 0.36 | 1.89 | 0.82 | 0.36 | 1.89 |
| Other Hispanic | 1.50 | 0.51 | 4.40 | 1.13 | 0.40 | 3.21 | 1.16 | 0.41 | 3.29 | 1.17 | 0.41 | 3.29 | 0.90 | 0.32 | 2.49 |
| **Mother's age** | | | | | | | | | | | | | | | |
| < 25 years | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 25-34 years | 0.69 | 0.57 | 0.84 | 0.69 | 0.57 | 0.84 | 0.68 | 0.56 | 0.83 | 0.69 | 0.57 | 0.84 | 0.69 | 0.56 | 0.84 |
| ≥35 years | 0.42 | 0.32 | 0.55 | 0.42 | 0.32 | 0.55 | 0.41 | 0.31 | 0.54 | 0.41 | 0.32 | 0.54 | 0.41 | 0.31 | 0.53 |
| **Total live births** | | | | | | | | | | | | | | | |
| 1 | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |
| 2 | 0.89 | 0.74 | 1.08 | 0.90 | 0.74 | 1.08 | 0.89 | 0.74 | 1.07 | 0.90 | 0.74 | 1.08 | 0.90 | 0.74 | 1.08 |
| 3 | 0.93 | 0.74 | 1.18 | 0.93 | 0.73 | 1.17 | 0.93 | 0.74 | 1.17 | 0.93 | 0.74 | 1.17 | 0.95 | 0.75 | 1.19 |
| 4+ | 0.84 | 0.64 | 1.10 | 0.82 | 0.63 | 1.07 | 0.83 | 0.64 | 1.09 | 0.84 | 0.64 | 1.10 | 0.83 | 0.64 | 1.09 |
| **Distance to the nearest power line** | | | | | | | | | | | | | | | |
| No lines within 2000 m | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - | 1.00 | - | - |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 600-1999 m | 0.97 | 0.82 | 1.16 | 0.98 | 0.82 | 1.16 | 0.95 | 0.80 | 1.13 | 0.95 | 0.80 | 1.13 | 0.95 | 0.80 | 1.13 |
| 200-599 m | 1.15 | 0.90 | 1.47 | 1.16 | 0.91 | 1.49 | 1.11 | 0.87 | 1.42 | 1.12 | 0.87 | 1.42 | 1.13 | 0.88 | 1.44 |
| 100-199 m | 1.16 | 0.77 | 1.75 | 1.13 | 0.75 | 1.70 | 1.05 | 0.70 | 1.57 | 1.05 | 0.70 | 1.58 | 1.02 | 0.68 | 1.53 |
| < 100 m | 1.48 | 0.99 | 2.20 | 1.50 | 1.01 | 2.23 | 1.35 | 0.91 | 1.99 | 1.31 | 0.89 | 1.93 | 1.28 | 0.87 | 1.89 |

[a] Odds ratios for each variable are adjusted for all other variables

# Chapter 8

## Overall summary and conclusions

Childhood leukemia and CNS/brain tumors are very rare diseases, and an implementation of prospective study design is not feasible. Therefore, all studies are case-control in design. In addition, in case-control studies with a rare outcome the odds ratio is a good approximation of the rate ratio.

These five studies were part of a large epidemiologic study of residential proximity to transmission lines and childhood cancer in California, a replication of Draper's study [91]. As Draper's study, this project was registry-based with information obtained from California Cancer and California Birth Registries.

One of the major strengths of registry-based studies is the absence of differential information bias (recall bias) which remains a problem in most interview-based studies. Since California Cancer and California Birth registries are independent of each other and participation of subjects was not required for data collection, selection bias was also unlikely in our studies.

Misclassification of cases of childhood leukemia and CNS/brain tumors was unlikely due to the completeness and high accuracy of the California Cancer Registry. Validation studies where birth certificate data were compared with structured post-partum interviews, the sensitivity of birth records to correctly identify most racial and ethnic groups was greater than 94% with the exception of Native Americans.[92] The issue of misclassification of independent variables such as birth weight, gestational age, parental age, parental education, payment source for delivery and census-based SES was partially addressed by using various cut points for these variables.

All analyses were adjusted for known and potential confounders. Nevertheless, information on some potential confounders, such as traffic density, parental occupation, nutritional factors, parental smoking status, substances use and other, was limited or unavailable. Thus, residual confounding was possible; but for the majority of aforementioned factors it was unlikely to bias the results or it would bias the results toward the null.

Despite a large sample size that allowed us to conduct analyses for main subtypes of childhood leukemia and CNS/brain tumors, the number of cases for some subtypes for childhood leukemia and brain tumors was small. As a results, many odds ratios for subtypes were not stable with wide 95% confidence intervals.

In addition, a significant proportion of values for some variables was missing. The majority of previous studies performed complete-case analyses without addressing a problem of missing data. To overcome this problem in our studies multiple imputations were performed and analyses was repeated using multiply imputed datasets and compared to complete-case analyses. No important differences were detected between these two analyses.

Summarizing results, we detected an increased risk of total childhood leukemia and acute lymphocytic leukemia (ALL) for high birth weight and large-for-gestational age (LGA) children and children with advanced paternal age. A decreased risk of total childhood leukemia and ALL was observed for small-for-gestational age (SGA) children. We also observed that being first-born was associated with a slightly decreased risk of total childhood leukemia and AML.

Our results indicate that maternal genital herpes, blood and immunological disorders during pregnancy, newborn CNS abnormalities, and advanced parental age were associated with an increased risk of CNS/brain tumors. Maternal infections during pregnancy were associated with a decreased risk of CNS tumors.

Our findings for subtypes suggest that different risk factors may play a role in etiology of subtypes for leukemia and CNS/brain tumors.

We found ethnic and racial differences in the incidence of childhood leukemia: children of Black race were at lower risk of childhood leukemia and ALL; children of Hispanic ethnicity were at high risk of childhood leukemia and ALL. The highest risk of childhood leukemia was observed for Hispanic White children and the lowest risk for non-Hispanic Black children. Asian race was associated with an increased risk of AML only. These differences in the incidence of childhood leukemia indicate that some genetic or environmental factors may be involved in etiology of childhood leukemia.

We found no evidence to support the suggestion that SES, as measured by variety of proxies is a determinant of childhood leukemia and of its both subtypes. We hypothesize that results of many previous studies that found an association between childhood leukemia and SES were largely influenced by selection or ecological bias.

Results of our research on residential mobility of childhood leukemia cases indicate that the residential mobility of childhood leukemia cases notably varied by child's age and race/ethnicity, maternal age at birth, census-based SES, and the source of payment for delivery; it varied less by the distance to the nearest power line. Our results suggest that even if

information on the residential mobility of subjects is unavailable, it might be possible and important to examine the distribution of factors that are associated with the residential mobility and to control for them.

The main theory behind the association of high birth weight and childhood tumors is the association of high birth weight with insulin-like growth factor 1 (IGF-1), a known procarcinogenic agent; thus, further research should be focused on the association of IGF-1 and childhood cancer via direct measuring of IGF-1 or through detection of a gene related to IGF-1 concentrations in blood. Racial and ethnic variations in the incidence of childhood cancers are also likely to be related to genetic differences. Therefore, further research on risk factors of childhood leukemia and CNS tumors should be directed more toward genetic and/or epigenetic determinants of these factors.

## Ethical Considerations

The study was approved by University of California, Los Angeles Office for the Protection of Research Subjects as well as by California Committee for the Protection of Human Subject (CPHS) and by Vital Statistics Advisory Committee of the California Department of Public Health.

**References:**


1. Ma, S.K., T.S. Wan, and L.C. Chan, *Cytogenetics and molecular genetics of childhood leukemia.* Hematol Oncol, 1999. **17**(3): p. 91-105.
2. Gurney, G., Smith, MA, Bunin, GR,, *CNS and Miscellaneous Intracranial and intraspinal neoplasms* in *SEER Pediatric Monograph*, N.C. Institute, Editor 2001, SEER: Bethesda. p. 51-64.
3. Children with leukaemia. *What is leukaemia*. About Leukaemia 2007 [cited 2008 July 30, 2008]; Available from: http://www.leukaemia.org/about-leukaemia/what-is-leukaemia.
4. National Children's Leukemia Foundation. *What Is Leukemia?* Education 2008 [cited 2008 July 30, 2008]; Available from: http://www.leukemiafoundation.org/cancer-and-leukemia-info/5.shtml.
5. American Cancer Society. *Types of Leukemia in Children*. Cancer Reference Information 2006 [cited 2007.
6. St. Jude Children's Research Hospital. *Leukemias / Lymphomas: Acute Lymphoblastic Leukemia (ALL)* Disease Information 2005 [cited 2007.
7. Linet, M.S., et al., *Cancer surveillance series: recent trends in childhood cancer incidence and mortality in the United States.* J Natl Cancer Inst, 1999. **91**(12): p. 1051-8.
8. Rubnitz, J.E. and W.M. Crist, *Molecular genetics of childhood cancer: implications for pathogenesis, diagnosis, and treatment.* Pediatrics, 1997. **100**(1): p. 101-8.
9. Belson, M., B. Kingsley, and A. Holmes, *Risk factors for acute leukemia in children: a review.* Environ Health Perspect, 2007. **115**(1): p. 138-45.
10. Kim, A.S., D.A. Eastmond, and R.J. Preston, *Childhood acute lymphocytic leukemia and perspectives on risk assessment of early-life stage exposures.* Mutat Res, 2006. **613**(2-3): p. 138-60.
11. Ries, L.A.G., Smith, M.A., Gurney JG, Linet, M., Tamra, T., Young, J.L., Bunin, G.R., *Cancer Incidence and Survival Among Children and Adolescents: United States SEER Program 1975-1995*, in *SEER Pediatric Monograph*, N.C. Institute, Editor 1999, SEER: Bethesda. p. 17-34.
12. Pollack, L.A., Stewart, S.L., Thompson, T.D.,, *Trends in Childhood Cancer Mortality -- United States, 1990--2004*, in *Morbidity and Mortality Weekly Report (MMWR)*2007, CDC. p. 1257-1261.
13. Campleman, S.L., Wright, W.E.,, *Childhood Cancer in California 1988 to 1999. Volume I: Birth to Age 14.* , in *Cancer Surveillance Section*, C.D.o.H. Services, Editor, California Cancer Registry: Sacramento, CA.
14. Robison, L.L. and J.P. Neglia, *Epidemiology of Down syndrome and childhood acute leukemia.* Prog Clin Biol Res, 1987. **246**: p. 19-32.
15. Westergaard, T., et al., *Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study.* J Natl Cancer Inst, 1997. **89**(13): p. 939-47.
16. Harrison, C.J., *The genetics of childhood acute lymphoblastic leukaemia.* Baillieres Best Pract Res Clin Haematol, 2000. **13**(3): p. 427-39.
17. Greaves, M., *Childhood leukaemia.* Bmj, 2002. **324**(7332): p. 283-7.

18. Roman, E., et al., *Childhood acute lymphoblastic leukemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study.* Am J Epidemiol, 2007. **165**(5): p. 496-504.
19. Baldus, C.D., et al., *High expression of the ETS transcription factor ERG predicts adverse outcome in acute T-lymphoblastic leukemia in adults.* J Clin Oncol, 2006. **24**(29): p. 4714-20.
20. Greaves, M., *Infection, immune responses and the aetiology of childhood leukaemia.* Nat Rev Cancer, 2006. **6**(3): p. 193-203.
21. Greaves, M.F., *Speculations on the cause of childhood acute lymphoblastic leukemia.* Leukemia, 1988. **2**(2): p. 120-5.
22. Smith, M.A., et al., *Evidence that childhood acute lymphoblastic leukemia is associated with an infectious agent linked to hygiene conditions.* Cancer Causes Control, 1998. **9**(3): p. 285-98.
23. Blattner W, C.M., *Human T-cell lymphotropic virus types I and II. In: Mandell, Douglas, and Bennett⁄s Principles and Practice of Infectious Diseases*, ed. 5th. 2005.: New York:Elsevier/Churchill Livingstone. 2098–2118
24. Greaves, M.F. and J. Wiemels, *Origins of chromosome translocations in childhood leukaemia.* Nat Rev Cancer, 2003. **3**(9): p. 639-49.
25. Kinlen, L.J. and C. Stiller, *Population mixing and excess of childhood leukemia.* Bmj, 1993. **306**(6882): p. 930.
26. Koushik, A., W.D. King, and J.R. McLaughlin, *An ecologic study of childhood leukemia and population mixing in Ontario, Canada.* Cancer Causes Control, 2001. **12**(6): p. 483-90.
27. Rudant, J., et al., *Childhood acute leukemia, early common infections, and allergy: The ESCALE Study.* Am J Epidemiol, 2010. **172**(9): p. 1015-27.
28. Richardson, R.B., *Promotional etiology for common childhood acute lymphoblastic leukemia: the infective lymphoid recovery hypothesis.* Leuk Res, 2011. **35**(11): p. 1425-31.
29. Urayama, K.Y., et al., *A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia.* Int J Epidemiol, 2010. **39**(3): p. 718-32.
30. Hjalgrim, L.L., et al., *Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland.* J Natl Cancer Inst, 2004. **96**(20): p. 1549-56.
31. Hjalgrim, L.L., et al., *Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies.* Am J Epidemiol, 2003. **158**(8): p. 724-35.
32. Ou, S.X., et al., *Birth characteristics, maternal reproductive history, hormone use during pregnancy, and risk of childhood acute lymphocytic leukemia by immunophenotype (United States).* Cancer Causes Control, 2002. **13**(1): p. 15-25.
33. Schuz, J. and M.R. Forman, *Birthweight by gestational age and childhood cancer.* Cancer Causes Control, 2007. **18**(6): p. 655-63.
34. Milne, E., et al., *Fetal growth and acute childhood leukemia: looking beyond birth weight.* Am J Epidemiol, 2007. **166**(2): p. 151-9.
35. Dorak, M.T., et al., *Examination of gender effect in birth weight and miscarriage associations with childhood cancer (United Kingdom).* Cancer Causes Control, 2007. **18**(2): p. 219-28.
36. Roman, E., et al., *Perinatal and reproductive factors: a report on haematological malignancies from the UKCCS.* Eur J Cancer, 2005. **41**(5): p. 749-59.

37. Ross, J.A., et al., *Epidemiology of childhood leukemia, with a focus on infants.* Epidemiol Rev, 1994. **16**(2): p. 243-72.
38. Ross, J.A., et al., *Evaluating the relationships among maternal reproductive history, birth characteristics, and infant leukemia: a report from the Children's Cancer Group.* Ann Epidemiol, 1997. **7**(3): p. 172-9.
39. Sandler, D.P. and J.A. Ross, *Epidemiology of acute leukemia in children and adults.* Semin Oncol, 1997. **24**(1): p. 3-16.
40. Borugian, M.J., et al., *Childhood leukemia and socioeconomic status in Canada.* Epidemiology, 2005. **16**(4): p. 526-31.
41. Poole, C., et al., *Socioeconomic status and childhood leukaemia: a review.* Int J Epidemiol, 2006. **35**(2): p. 370-84.
42. Ribeiro, K.B., P.A. Buffler, and C. Metayer, *Socioeconomic status and childhood acute lymphocytic leukemia incidence in Sao Paulo, Brazil.* Int J Cancer, 2008. **123**(8): p. 1907-12.
43. Okcu, M.F., et al., *Birth weight, ethnicity, and occurrence of cancer in children: a population-based, incident case-control study in the State of Texas, USA.* Cancer Causes Control, 2002. **13**(7): p. 595-602.
44. Howard, S.C., et al., *Childhood cancer epidemiology in low-income countries.* Cancer, 2008. **112**(3): p. 461-72.
45. Mahoney, M.C., et al., *The Chernobyl childhood leukemia study: background & lessons learned.* Environ Health, 2004. **3**(1): p. 12.
46. Ron, E., *Ionizing radiation and cancer risk: evidence from epidemiology.* Radiat Res, 1998. **150**(5 Suppl): p. S30-41.
47. Feychting, M., A. Ahlbom, and L. Kheifets, *EMF and health.* Annu Rev Public Health, 2005. **26**: p. 165-89.
48. Kheifets, L., M. Feychting, and J. Schuz, *Childhood cancer and power lines: results depend on chosen control group.* Bmj, 2005. **331**(7517): p. 635; discussion 636; author reply 636-7.
49. Ahlbom, A., et al., *A pooled analysis of magnetic fields and childhood leukaemia.* Br J Cancer, 2000. **83**(5): p. 692-8.
50. Greenland, S., et al., *A pooled analysis of magnetic fields, wire codes, and childhood leukemia. Childhood Leukemia-EMF Study Group.* Epidemiology, 2000. **11**(6): p. 624-34.
51. Hatch, E.E., et al., *Association between childhood acute lymphoblastic leukemia and use of electrical appliances during pregnancy and childhood.* Epidemiology, 1998. **9**(3): p. 234-45.
52. Ritz, B. and R.P. Rull, *Assessment of environmental exposures from agricultural pesticides in childhood leukaemia studies: challenges and opportunities.* Radiat Prot Dosimetry, 2008. **132**(2): p. 148-55.
53. Ahlbom, A., et al., *A pooled analysis of magnetic fields and childhood leukaemia.* Br J Cancer, 2000. **83**(5): p. 692-8.
54. Kheifets, L., et al., *Pooled analysis of recent studies on magnetic fields and childhood leukaemia.* Br J Cancer, 2010. **103**(7): p. 1128-35.
55. Kheifets, L. and R. Shimkhada, *Childhood leukemia and EMF: review of the epidemiologic evidence.* Bioelectromagnetics, 2005. **Suppl 7**: p. S51-9.

56. Raaschou-Nielsen, O. and P. Reynolds, *Air pollution and childhood cancer: a review of the epidemiological literature.* Int J Cancer, 2006. **118**(12): p. 2920-9.
57. Reynolds, P., et al., *A case-control pilot study of traffic exposures and early childhood leukemia using a geographic information system.* Bioelectromagnetics, 2001. **Suppl 5**: p. S58-68.
58. Alavanja, M.C., M.H. Ward, and P. Reynolds, *Carcinogenicity of agricultural pesticides in adults and children.* J Agromedicine, 2007. **12**(1): p. 39-56.
59. McWhirter, W.R., *The relationship of incidence of childhood lymphoblastic leukaemia to social class.* Br J Cancer, 1982. **46**(4): p. 640-5.
60. Zipursky, A., et al., *Leukemia and/or myeloproliferative syndrome in neonates with Down syndrome.* Semin Perinatol, 1997. **21**(1): p. 97-101.
61. Paltiel, O., et al., *Birth weight and other risk factors for acute leukemia in the Jerusalem Perinatal Study cohort.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(6): p. 1057-64.
62. Adam, M., et al., *Childhood leukaemia and socioeconomic status: what is the evidence?* Radiat Prot Dosimetry, 2008. **132**(2): p. 246-54.
63. Kaye, S.A., et al., *Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia.* Cancer, 1991. **68**(6): p. 1351-5.
64. Cnattingius, S., et al., *Prenatal and neonatal risk factors for childhood myeloid leukemia.* Cancer Epidemiol Biomarkers Prev, 1995. **4**(5): p. 441-5.
65. Ross, J.A., et al., *Periconceptional vitamin useand leukemia risk in children with Down syndrome: a Children's Oncology Group study.* Cancer, 2005. **104**(2): p. 405-10.
66. Schuz, J., T. Weihkopf, and P. Kaatsch, *Medication use during pregnancy and the risk of childhood cancer in the offspring.* Eur J Pediatr, 2007. **166**(5): p. 433-41.
67. Wen, W., et al., *Parental medication use and risk of childhood acute lymphoblastic leukemia.* Cancer, 2002. **95**(8): p. 1786-94.
68. MacArthur, A.C., et al., *Risk of childhood leukemia associated with vaccination, infection, and medication use in childhood: the Cross-Canada Childhood Leukemia Study.* Am J Epidemiol, 2008. **167**(5): p. 598-606.
69. Jensen, C.D., et al., *Maternal dietary risk factors in childhood acute lymphoblastic leukemia (United States).* Cancer Causes Control, 2004. **15**(6): p. 559-70.
70. Peters, J.M., et al., *Processed meats and risk of childhood leukemia (California, USA).* Cancer Causes Control, 1994. **5**(2): p. 195-202.
71. Sarasua, S. and D.A. Savitz, *Cured and broiled meat consumption in relation to childhood cancer: Denver, Colorado (United States).* Cancer Causes Control, 1994. **5**(2): p. 141-8.
72. Whitworth, K.W., E. Symanski, and A.L. Coker, *Childhood lymphohematopoietic cancer incidence and hazardous air pollutants in southeast Texas, 1995-2004.* Environ Health Perspect, 2008. **116**(11): p. 1576-80.
73. Pearson, R.L., H. Wachtel, and K.L. Ebi, *Distance-weighted traffic density in proximity to a home is a risk factor for leukemia and other childhood cancers.* J Air Waste Manag Assoc, 2000. **50**(2): p. 175-80.
74. Shu, X.O., et al., *Chloramphenicol use and childhood leukaemia in Shanghai.* Lancet, 1987. **2**(8565): p. 934-7.
75. Gurney, G., Smith, MA, Bunin, GR,, *CNS and Miscellaneous Intracranial and intraspinal neoplasms* in *SEER Pediatric Monograph*, N.C. Institute, Editor 1999, SEER: Bethesda. p. 51-64.

76. National Cancer Institute, *A snapshot of brain and central nervous system cancers*, in *Snapshots*2008, National Cancer Institute, Office of Science Planning and Assessment
77. Kaatsch, P., et al., *Population-based epidemiologic data on brain tumors in German children.* Cancer, 2001. **92**(12): p. 3155-64.
78. Central Brain Tumor Registry of the United States. *FactSheet*. Factsheets [cited 2009 April 23,]; Available from: http://www.cbtrus.org/factsheet/factsheet.html.
79. Baldwin, R.T. and S. Preston-Martin, *Epidemiology of brain tumors in childhood--a review.* Toxicol Appl Pharmacol, 2004. **199**(2): p. 118-31.
80. Smith, M.A., et al., *Trends in reported incidence of primary malignant brain tumors in children in the United States.* J Natl Cancer Inst, 1998. **90**(17): p. 1269-77.
81. Kheifets, L., et al., *A pooled analysis of extremely low-frequency magnetic fields and childhood brain tumors.* Am J Epidemiol, 2010. **172**(7): p. 752-61.
82. Matasar, M.J., et al., *Incidence rates of the major leukemia subtypes among US Hispanics, Blacks, and non-Hispanic Whites.* Leuk Lymphoma, 2006. **47**(11): p. 2365-70.
83. Yamamoto, J.F. and M.T. Goodman, *Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002.* Cancer Causes Control, 2008. **19**(4): p. 379-90.
84. The California Cancer Registry. *Mission Statement & Purpose*. March 2009]; Available from: http://www.ccrcal.org/aboutthecr.html.
85. California Department of Public Health. *Available Birth and Fetal Death Variables*. March, 2009]; Available from: http://www.cdph.ca.gov/data/dataresources/requests/Pages/OHIRBirthVariables.aspx.
86. Goldberg, D., Cockburn, M,, *Improving Geocode Accuracy with Candidate Selection Criteria.* Transactions in GIS, 2010. **14**(1): p. 149-176.
87. Rubin, D.B. and N. Schenker, *Multiple imputation in health-care databases: an overview and some applications.* Stat Med, 1991. **10**(4): p. 585-98.
88. Little, R., Rubin, D., *Statistical Analysis with Missing Data*. 2nd Edition ed. 2002, New York: Wiley.
89. Hosmer, D.W., Jr. , Lemeshow, S., *Applied Logistic Regression*. 2000, New York: Wiley.
90. SAS Institute, I., *SAS Statistical Software*, 2006: Cary, NC.
91. Draper, G., et al., *Childhood cancer in relation to distance from high voltage power lines in England and Wales: a case-control study.* Bmj, 2005. **330**(7503): p. 1290.
92. Baumeister, L., et al., *The validity of information on "race" and "Hispanic ethnicity" in California birth certificate data.* Health Serv Res, 2000. **35**(4): p. 869-83.