Breastfeeding and Risk of Acute Lymphoblastic Leukemia Within the Context of Immune Related Factors

by

Julia Sergeevna Selezneva

A thesis submitted in partial satisfaction of the

Requirements for the degree of

Master of Science

In

Health and Medical Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Lisa F. Barcellos, PhD MPH
Professor Maureen Lahiff, PhD
Professor John Balmes, MD

Spring 2014

Breastfeeding and Risk of Acute Lymphoblastic Leukemia

Within the Context of Immune Related Factors

By Julia Sergeevna Selezneva

Dedication

I would like to dedicate this thesis to my grandmother Valentina Postolenko for her unconditional love and support throughout my entire life.

Acknowledgements

I would first like to thank Maureen Lahiff. Without her encouragement, guidance, and occasional tough love, this thesis would not have been completed. To Patricia Buffler, who started the California Childhood Leukemia Study, and whose support for and kindness to students doing research with CCLS was extraordinary. She is sorely missed. I would also like to thank the other members of my committee, John Balmes and Lisa Barcellos, not only for their edits on this thesis, but for rallying behind me after the tragic and unexpected death of my committee chair and mentor, Dr. Patricia Buffler.

Thank you to the other researchers of the CCLS team – Anand Chokkalingham, Joseph Weimels, and Kevin Urayama, for supporting my research ideas and for their feedback. To the UC Berkeley-UCSF Joint Medical Program faculty and staff, especially Collette Auerswald, Douglas Jutte, Harrison Alter, Ralph Catalano, Tracey Jones, and my thesis working group classmates for their contributions over the last 2 years. This work would not have been possible without financial support from the Helen Schoeneman Research Grants. To all the patients and parents who participated in this study – I hope your child is alive and well, and would like to thank you for this contribution to help understand how we can best prevent and treat more children affected by leukemia.

Finally, to my wonderful, loving boyfriend Yoganand Chillarige, who has spent countless hours listening to my ideas regarding this thesis, as well as making sure I took the time to breathe and enjoy life.

Table of Contents

List of Tables

Paper 1

Introduction

Today the treatment of childhood leukemia is one of the greatest success stories in the field of oncology[1]. With aggressive chemotherapy and supportive care, 75-85% of children with leukemia are cured[2]. Nevertheless, childhood leukemia remains the leading cause of cancer deaths among children[3]. Furthermore, there are continuing challenges in understanding the etiology behind childhood leukemia[4]. The goal of this paper is to review the epidemiology, etiology, and risk factors of childhood leukemia and to discuss the role of breastfeeding in the infectious etiology hypothesis in childhood leukemogenesis.

Background

The Epidemiology of Childhood Leukemia

Cancer is the second most common cause of death among children between 1-14 years of age in the United States, with the most common cause of death being accidents[3,5]. The overall incidence rate for cancer under the age of 14 has increased by 0.5% each year since 1975 (not statistically significant); while the death rate for childhood cancer decreased from 4.9 per 100,000 in 1975 to 2.2 in 2008[5-7]. In 2008, cancer accounted for 1,284 deaths among children in the US between the ages of one and fourteen years [6]. Leukemia is responsible for approximately 30% of diagnoses of cancer among children under the age of 14, with other common childhood cancers being: cancers of the brain and nervous system (27%), soft tissue sarcomas (7%), neuroblastoma (7%), renal tumors (5%), and Hodgkin/non-Hodgkin lymphomas (4% each)[6,8]. Among childhood leukemias, acute lymphoblastic leukemia (ALL), occurs five times more frequently than acute myelogenous leukemia (AML) and accounts for 78% of childhood leukemia diagnoses[5,9]. Each year in the US, there are 2500 to 3500 new cases of childhood ALL diagnosed, with an incidence of 2.8 cases per 100,000[10].

The Biology of Childhood Leukemia

Leukemias arise from the malignant transformation and proliferation of stem or progenitor cells involved in the process of hematopoiesis that eventually produces lymphoid cells (B- and T-cells) and myeloid cells (granulocytic, monocytic, erythroid and megakaryocytic cells)[11]. This process occurs in the bone marrow, lymph nodes, and/or other lymphoid tissue with immune function. As a result, childhood leukemia is made of several subtypes that vary in phenotype and age incidence, with the broadest division being the lymphoid and myeloid split[11].

Acute Lymphoblastic Leukemia

ALL is a cancer of lymphoid progenitor cells and is composed of immature B (pre-B) or T (pre-T) cells[12]. About 85% of ALL is B-ALL; the less common T-ALL tends to present in adolescent males as thymic lymphomas. B-cell ALL peaks in incidence between the ages of 2-5[12]. It occurs more frequently among boys than in girls[6,9]. Hispanics have the highest incidence among any

ethnic group; ALL presents in Hispanics 1.2 times as often as in whites, and twice as often as in blacks.[4].

ALL is associated with certain genetic and immunodeficiency syndromes such as Down syndrome, Neurofibromastosis type 1, Bloom syndrome, and ataxia telangiectasia[9,13]. Exposure to ionizing radiation also increases risk for ALL[9,13]. However, ionizing radiation and congenital genetic syndromes explain less than 10% of all cases [4].

The classification of ALL is based on morphologic, immunologic, biochemical and cytogenetic features[10]. Morphologic classification of ALL is based on bone marrow aspirate evaluation using the French-American-British system. Eighty-five to 89% percent of children with ALL are classified as having FAB L1 (lymphoblasts that are small cells with scant cytoplasm, condensed nuclear chromatin, and indistinct nucleoli)[10,14]. Immunophenotypic characteristics are determined using a panel of monoclonal antibodies to cell surface "cluster of differentiation" (CD) markers that distinguish at which stage of pre-B cell development the leukemic cells have been arrested[11 14]. Seventy to 80% of cases of childhood ALL are of B-precursor lineage (early pre-B ALL). B-precursor leukemia is CD10+ and CD19+ [15]. Very immature B-ALLs are CD10 negative. More mature "late pre-B" ALL expresses CD10+, CD19+, CD20+, and cytoplasmic IgM heavy chain (μ chain) [15]. Additional immunostaining for terminal deoxynucleotidyl-transferase (TdT), a specialized DNA polymerase that is expressed only in pre-B and pre-T lymphoblasts, is positive in more than ninety-five percent of cases of ALL[11].

Cytogenetically ALL is classified through karyotyping, fluorescence in situ hybridization, and other molecular techniques used to identify numerical and structural chromosomal abnormalities[15]. Chromosomal abnormalities include high hyperdiploidy of 54 to 58 chromosomes and hypodiploidy (fewer than 45 chromosomes). Structural abnormalities are translocations between chromosomes or rearrangements. The most common one is t(12;21) TEL-AML1 translocation, which occurs in 20-25% of childhood ALL cases. Other common translocations include t(4;11) MLL/AF4 rearrangements present in 5% of pediatric ALL and 60% of infant ALL patients, and t(9;22) BCR/ABL present in 3-4% of ALL (often occurs in older children)[10,15].

Acute Myeloid Leukemia

On the other hand, acute myelogenous/myeloid leukemia (AML) is a leukemia that involves the clonal proliferation of myeloid precursors and accounts for less than 10% of acute leukemias in children less than 10 years of age[11,14,16]. In contrast, among adults, AML accounts for approximately 80% of acute leukemia[3]. In contrast to ALL, the incidence of AML increases with age, with 1.3 and 12.2 cases per 100,000 population for those under or over 65 years, respectively, and the median age at diagnosis is 65 years[3,13]. It occurs more commonly in males than females, with the male to female ratio being 5:3[13,17]. The incidence of childhood AML is similar among persons of different ethnic groups in the US, although there seems to be an increased incidence among Hawaiians and increased incidence among Hispanics for a specific subtype of AML, acute promyelocytic leukemia[3,18]. However, among adults, non-Hispanic

whites have a slightly higher incidence than Hispanics, blacks and Asian Pacific Islanders (4 cases versus 3 cases per 100,000 population)[17,19].

AML is also associated with congenital genetic syndromes such as Down syndrome, Bloom syndrome, Fanconi's anemia, and familial RUNX1 mutations[13,17]. It is also associated with environmental factors, such as exposure to radiation, chemicals, tobacco, or chemotherapy drugs, as well as other hematologic diseases (such as paroxysmal nocturnal hemoglobinuria, aplastic anemia, myelodysplastic syndrome and myeloproliferative disorders)[13].

Presentation

Although ALL and AML are genetically and immunophenotypically distinct, they have a similar clinical presentation[11]. In both leukemias, the accumulation of neoplastic immature "blast" cells in the bone marrow suppresses the normal hematopoietic process through physical crowding, competition for growth factors, and other mechanisms that are not well understood[10,11,17]. The suppression of bone marrow function results in fatigue due to anemia, fever due to infections secondary to neutropenia, and bleeding due to thrombocytopenia. Bone pain occurs because of neoplastic proliferation, bone marrow expansion and infiltration of the subperiostium[11,20]. Additional symptoms include generalized lymphadenopathy, splenomegaly, hepatomegaly, testicular enlargement, and central nervous system manifestations such as headache, vomiting and nerve palsies due to meningeal spread. In acute leukemias, these symptoms come in an abrupt stormy onset within days to weeks of the first symptoms[11,20].

Prognosis

The prognosis for ALL is very good[6,21]. With aggressive chemotherapy and/or bone marrow transplant, about 95% of children with ALL obtain a complete remission and 75-85% are cured. Nevertheless, ALL remains the leading cause of cancer deaths in children. Favorable prognostic markers are: age between 2-10 years; a low white cell count; hyperploidy; trisomy of chromosomes 4, 7, or 10; and the presence of t(12;21) TEL-AML1 translocation[2,10]. Poor prognostic markers are: age under 2 (largely due to the strong association of infantile ALL with translocations involving the MLL gene) or presentation in adolescence or adulthood; peripheral white cell count greater than 100,000 reflecting a higher tumor burden; and the presence of certain cytogenetic translocations such as the t(9;22) BCR-ABL fusion (also known as the Philadelphia chromosome), which is present in 3% of childhood ALL[2,10].

On the other hand, the prognosis for AML is relatively poor, with 5-year-survival rates averaging at 54% for children under the age of 13, and 46% for adolescents aged 13 to under 21 years of age[16,21,22]. However, survival varies depending on certain cytogenetic and molecular abnormalities, with 5-year survival ranging from 22% to 90% among different cytogenetic changes [23]. Unlike adult AML, age is not an independent prognostic factor in infants and adolescents, and very high blast counts at diagnosis have been associated with increased risk of death and nonresponse[23].

Etiology of Childhood Leukemias

The genesis of cancer is typically a multistep process[24]. Similarly for childhood leukemia, single mutations are not sufficient to produce ALL[4,12]. Current evidence suggests that the "first hit" in the genesis of leukemia occurs *in utero*, establishing a pre-leukemic clone[1]. Given that childhood ALL usually appears in patients between the ages of 2-5 years, this lengthy prodrome is consistent with the pre-leukemic clone acquiring additional mutations that lead to the development of overt leukemia[1].

Evidence supporting this comes from studies that demonstrate that key translocation events occur prenatally for specific molecularly defined ALL subtypes, such as infant leukemia characterized by MLL-AF4 t4[4;11] translocation, childhood ALL characterized by TEL-AML1 t[12;21] translocation, and childhood AML characterized by AML1-ETO translocation[25,26]. Specifically, studies on identical twins with concordant B-ALL demonstrated that both twins shared a common chromosomal aberration[27]. Among triplets, different post-natal genetic events occurred despite three triplets demonstrating the presence of TEL-AML1 fusion sequences in their blood spots[28]. Additional studies with neonatal blood spots, Guthrie cards, have identified the presence at birth of common mutations such as t(12;21) TEL-AML1, t(8;21) AML1-ETO[29,30]. However, Mori et al showed a 100-fold higher occurrence of TEL-AML1 or AML-ETO fusion genes in cord blood than the incidence of childhood ALL or AML[31]. The mutations associated with leukemia are insufficient to cause disease by themselves, as these translocations occur at a rate of 1% or more in the normal population[31,32]. This suggested that the critical rate-limiting step in the production of overt leukemia occurs after birth[4]. Other mutations, such as t(1;19) E2A-PBX1, FLT3, and RAS are postnatal[33,34]. MLL translocations (11q23) appear to occur within temporal proximity to diagnosis, meaning that infants (<1 year old) have prenatal translocations, and children beyond 2 years of age at diagnosis have postnatal translocations[35]. The identity of the "second-hit" mutations is incomplete, but aberrations that increase growth and survival, such as activating mutations in tyrosine kinases, are commonly present[13].

Since fetal and child hematopoiesis exhibit a high degree of cellular proliferation, it is a window of time where environmental exposure to chemicals may induce mutations[4]. For example, children with t(12;21) TEL-AML1 translocations had 4 times the odds of being born from mothers who were exposed to paints during their pregnancy when compared to controls and this increased risk was not noted for other cytogenetic subtypes of leukemia[36]. Environmental exposures including ionizing radiation for diagnostic imaging during pregnancy and atomic bomb exposure during childhood and young adulthood have been demonstrated to have a strong causative association[37,38]. Other environmental contributors include the diet of the mother and child, parental smoking, pesticides and household chemicals, traffic fumes, immunologic modifiers[4,39-42]. In addition, variation in several genes implicated in B-cell development (ARID5B, IKZF1, CEBPA), and cell cycle regulation/DNA repair (CDKN2A/B) are confirmed genetic risk factors for childhood ALL. Several HLA haplotypes associations with childhood leukemia have been noted[43-46].

Infectious Exposures and Immune System Development

Of environmental exposures that have been shown to be associated with the progression of leukemia, those related to immunologic development, specifically infectious diseases, have been the most pertinent[4]. The hypothesis that exposure to infections and the development of the immune system play a role in the development of leukemia was proposed by two separate researchers with a different take on possible infectious etiology[47-49]. The "population mixing" hypothesis advanced by Kinlen proposes that a specific viral infection is responsible for the leukemia "outbreaks" that occur during population mixing[48]. This hypothesis came from observations that the leukemia incidence often increased when children and families were moved and mixed with the removal of children during central London during World War II, and the creation of new towns in Britain in the decades following the war[48]. However, another hypothesis was posed by Greaves, who noted that children who received a lower level of immune stimulation during early childhood developed a higher risk of leukemia, and as a result, a normal course of infection in early childhood was protective against the development of leukemia[47]. The lack of immune stimulation in children who are relatively isolated followed by an aberrant over response to common infections later in childhood is hypothesized as the promotion of the second "hit" in children who have a pre-leukemic mutation[50].

In support of Kinlen's hypothesis, a number of studies have shown elevated childhood ALL in geographic regions with high levels of population mixing (e.g., previously isolated areas in which there was a recent increase in population density, areas of population growth, and regions with population movements during wartime, increased social contact during commuting, or mass tourism)[51]. However, screening for four lymphotropic herpes virus genomes in leukemic samples using conventional molecular techniques and sensitive real-time PCR revealed no novel herpesvirus genomes. Samples from children with ALL also demonstrated no evidence of genomes of JC and BK polyoma viruses[13]. Nevertheless it has been suggested that there is an infectious etiology behind the recent cluster at Fallon, Nevada, in which 17 children developed cancer between 1997 to 2004, especially given the geographic patterning, narrow time window of the cluster, along with similar age and immunophenotype of the leukemias [51,52].

Greaves' hypothesis is supported by a number of studies evaluating indirect measures linking infectious agents with childhood ALL, such as childcare history, vaccinations, infections, maternal infection during pregnancy, and allergies/asthma[53-57]. Childcare history, or day care, is considered a proxy measure for exposure to infection since the more contacts a child has in day care setting the better chance s/he has for exposure to new infections. Studies of day care and leukemia sometimes demonstrated no difference in frequency, or more often a reduced level of day care in children who contract leukemia compared to controls[53]. A meta-analysis by Urayama and colleagues compiled evidence from 15 studies and yielded a combined OR of 0.76 (95% CI 0.67-0.87), indicating a reduced level of prior childhood contacts in leukemia patients via day care settings in the majority of studies[58]. Studies have also shown that higher birth order and vaccinations such as those against *Haemophilus influenzae* (HiB), have been associated with reduced risks[59]. These studies suggest that an increased opportunity for early childhood infections protects against leukemia[50]. Household crowding, medical conditions suggesting poor hygiene, and household pets have not been adequately studied[13]. There is some evidence that measures of higher socioeconomic status are linked with increased risk of childhood ALL[60]. A

history of one or more allergic disorders has also been linked with a significantly reduced risk of childhood ALL[57]. This inverse relationship, however, may be due to information bias, due to the nature of case-control studies that assess allergies through parental recall[4]. Parents of children with leukemia may ruminate about factors that may have affected their child's risk to leukemia leading to false positive associations, while families of children included as controls, tend to misreport allergies that may have occurred after a ''reference date'' (diagnostic date for the corresponding leukemic children), therefore, over-reporting exposures, infections, and medical conditions[4]. Two other studies that utilized medical record abstraction rather than patient interview found that allergy was a risk factor for leukemia[57,61]. It is possible that allergies and leukemia share similar risk factors related to immune system development. The hygiene hypothesis proposed by Strachan to explain the rising prevalence of allergy in Western populations is very similar to Greaves' hypothesis[62]. The hygiene hypothesis states that early childhood infections may be protective against allergy, but that declining family size and improved sanitation may have reduced exposure to infectious agents during early childhood thus resulting in immune dysregulation and a rise in the prevalence of allergy[4,57,62]. Given the generally sporadic patterning of leukemia incidence among populations, one could argue that prevailing evidence favors Greaves' hypothesis of abnormal immune development rather than the involvement of a specific leukemia virus for the induction of the second "hit"[4].

Dysregulated Immune System and Childhood Leukemia

It is thought that children who develop leukemia may be born with a dysregulated immune system, and while immune stimulation by early life exposure to infections and vaccinations appear to be protective, those born with an aberrant immune system may respond to infections more vigorously – predisposing them to the induction of a second "hit" to the pre-leukemic clone[4,50]. It has been demonstrated that children diagnosed with ALL had significantly more clinically diagnosed infectious episodes in the first year of life compared to controls[63-65]. The number of infectious episodes in children with ALL increased with increasing indices of infectious exposure (birth order, regular social activity outside the home, and social deprivation at birth), a phenomenon not observed among healthy control children[65]. These studies also demonstrated fewer social contacts for children contracting leukemia, indicating that overall exposure to infections were likely lower than controls[64].

Two separate phenomena may be influencing leukemia risk: (i) a reduced repertoire of infections during early immune development which would increase the risk and (ii) a congenitally altered immune system response to infection, also increasing the risk of leukemia[4].

Chang et al demonstrated that IL-10, a key anti-inflammatory cytokine secreted by monocytes and lymphocytes and critical in regulating the intensity and duration of immune responses to infections, is deficient among children who later developed leukemia[66]. It is possible that children with dysregulated immune function at birth are at higher risk for developing leukemia due to constitutively lower expression of IL10[66]. Chang et al also analyzed 29 adaptive immune function genes for polymorphisms among 377 children with ALL and 448 matched controls and found that a polymorphism in IL12A, the main driver of T_H1 immunity, was significantly associated with increased risk of ALL (OR: 1.52, 1.25-1.85, $P = 2.9 \times 10^{-5}$) and this risk was

strongest among first-born children and non-Hispanic children with less day care attendance[67]. Infections in childhood combined with a dysregulated immune response may result in rapid expansion of a pre-leukemic clone, leading to an increased opportunity for acquiring a second mutation required for the development of childhood leukemia[4].

Breastfeeding and the Risk of Childhood Leukemia

Within the context of the Greaves' hypothesis, it is postulated that maternal breastfeeding may protect against childhood ALL by modulating the child's immune system early in life to respond effectively during exposure to common infections later in life[50]. Breastfeeding has been demonstrated to be a relevant exposure during a critical developmental period, offering both nutritive and immunologic benefits for newborns and young infants[68]. Studies have shown that it is protective against diarrhea, respiratory diseases, otitis media, and necrotizing enterocolitis[69]. For example, infants not breastfed have a higher risk of being hospitalized for and dying from pneumonia than those breastfed[70]. In addition, breastfeeding has been linked to long-term health benefits such as reduced risk of hyperlipidemia, hypertension, diabetes mellitus type 2, and asthma[69].

Passive protection from infections and immune modulation

The immunological function of breastfeeding can be summarized in two main categories: the passive protection of newborn from infection, and the modulation of the immune system[71]. Breast milk is known to protect newborns from infection. The immunologic contents of breast milk are responsible for stimulating intestinal maturation, enhancing production of specific antibodies, controlling and preventing inflammation, encouraging proliferation of commensal bacteria, and facilitating the survival of substances that protect the infant[68]. Breast milk is rich in secretory IgA (sIgA) antibodies that bind to microbes, such as *E. coli* and *Campylobacter*, in the infant intestine and prevent attachment to the walls of the gut[72]. Very little sIgA is produced in the neonatal period, and as such, breast milk is the neonate's predominant source of the immunoglobulin[73,74].

In addition to antibodies, breast milk contains other immunoactive compounds. Breast milk contains probiotic oligosaccharides, called "bifidogenic" oligosaccharides ("bifidus factor"), that promote propagation of commensal gut bacteria, such as *Bifidobacterium and Lactobacillus bifidus,* which decrease intestinal pH by producing lactic acid[75,76]. These benign bacteria compete with pathogenic bacteria and limit the available nutrients for the latter. In turn, *Lactobacillus* and other commensal bacteria have been shown to stimulate gastrointestinal plasma cells to synthesize sIgA[75]. Breast milk also has over 130 different non-digestible oligosaccharides that competitively inhibit the binding of pathogens or their toxins to the respiratory, gastrointestinal, and urinary tracts[68,75]. For example, some prevent binding of *Streptococcus pneumoniae* to the respiratory epithelium, while others inhibit *E. coli* and *Vibrio cholerae*[68]. In addition, breast milk contains lysozyme, an enzyme that attacks the peptidoglycan layer in the cell wall of certain pathogenic bacteria[68].

Although it has been shown that the immune system is complete at birth, the exposures during infancy and early childhood are essential for the expansion and the priming of adaptive cell immunity[71]. As such, breast milk has been demonstrated to have components that modulate the

immune system by both inhibiting the inflammatory response and promoting it when necessary for development[71]. The immature intestinal epithelium of the neonate has been shown to mount an excessive inflammatory response to both endogenous and exogenous bacteria[77]. This propensity to generate an over-reactive inflammatory response leaves the neonate, and particularly the preterm infant, at risk of chronic inflammation[77]. The anti-inflammatory contents of breast milk help mitigate the risk of an overactive inflammatory response, which may be responsible for the pathogenesis of inflammatory bowel diseases in the preterm infant[71,78]. There are many different classes of anti-inflammatory agents found in human breast milk, including enzymes that degrade inflammatory mediators, epithelial growth factors, antioxidants, and substances that bind to toxins[68,71]. One example is lactoferrin, which binds to iron and reduces the amount of iron available to pathogenic bacteria (thus stemming their proliferation). Iron also prevents leukocytes from releasing pro-inflammatory cytokines like interleukin-1β (IL-1β), tumor necrosis factor-α (TNF-α), and IL-8[68,71]. Breast milk contains other anti-inflammatory cytokines, including IL-10, platelet activating factor (PAF), and transforming growth factor-β (TGF-β), that have been shown to reduce expression of IL-8 and other pro-inflammatory cytokines[68,71].

In addition to anti-inflammatory cytokines, breast milk also contains substances that promote synthesis and activation of endogenous inflammatory and immune mediators that allow the infant to mount an appropriate inflammatory response when necessary[68,79]. For example, intestinal secretions of newborns lack CD-14, an essential cytokine in the innate immune system's inflammatory response. Breast milk, however, contains CD-14, and its levels are highest in colostrum and early milk[77]. Therefore, breast-fed infants are more able to mount a defense against intestinal pathogens. In addition, because breast milk contains a diverse range of anti-inflammatory and immune-modulatory agents, including T_H1 promoting cytokines (INF-ϒ, IL-12), and T_H2 type promoting cytokines (IL-6, Il-10), breastfeeding may play an important biological role in priming the newborn immune system[79]. Other constituents of breast milk have been shown to activate components of the endogenous immune system including macrophages, T cells, and neutrophils[73,80].

As a result, breastfeeding has been supported as an important component for neonatal immune system development[68,72-74]. This can be demonstrated by the role of breastfeeding in the rapid maturation of the neonatal TLR system in the first month of life [74,81]. During this important period, TLR-mediated cytokine responses mature rapidly from a Th2-biased profile toward increased Th1-cell polarizing responses characteristic of later life[74,82]. As such, the innate immune responses at the age of 1 month were markedly different between neonates who received exclusive breastfeeding during the first month of life and those who did not[82]. Several explanations may account for the association between breastfeeding and low TLR7-mediated IL-10 production. The first explanation is that breast milk contains multiple immune modulatory compounds that directly influence TLR-mediated immune responses, including immunoglobulins, nucleotides, oligosaccharides and antimicrobial proteins/peptides[82]. *In vitro* studies have shown that breast milk increases monocyte production of IL-10 while decreasing production of IFN-γ in response to lipopolysaccharide, mitogen and allergen[82]. Immune modulation by breast milk is TLR-specific, as breast milk suppresses IL-8 production to agonists for TLR2 and TLR3, while enhancing responses to TLR4 and TLR5. Researchers have proposed that suppression of TLR7-mediated IL-10 production may promote neonatal defense against viral

infections, while induction of TLR4-mediated IL-10 production serves to promote neonatal tolerance to bacterial colonization[82]. However, the alternative explanation is that the decreased TLR7-mediated IL-10 production at the age of 1 month might be due to decreased incidence of respiratory viral infections in breastfed neonates. Both respiratory syncytial virus (RSV) and rhinovirus, the most common causes of neonatal respiratory tract infections, are single-stranded RNA viruses that trigger TLR7[82]. Severe RSV infection is associated with increased production of IL-10 during and after infection. Breastfeeding protects against severe RSV infections and reduces RSV-mediated immune activation[82]. Reduced incidence of severe RSV infections in the first month of life may account for the lower IL-10 production in breastfed neonates. However, one study found no difference in TLR7-mediated IL-10 between neonates with and without history of respiratory tract infections in the first month of life ($p = 0.96$). In addition, as RSV respiratory tract infections are relatively rare in the first 2–3 months, other mechanisms are likely to play a role. Furthermore, decreased TLR7-mediated IL-10 production (and increased TLR3-mediated IL-12) production in breastfed infants may also reflect more rapid transition to a Th1-polarized innate immune system. However, breastfeeding did not affect cytokine responses to TLR4 and TLR9[82].

As such, the content of breast milk primes the infant for immunologic maturation, and for growth of the immature epithelium in the respiratory and gastrointestinal tracts, protecting the neonate and infant from infections[68]. Within the context of Greaves' infectious etiology hypothesis, breastfeeding may play a significant role in the etiology of childhood leukemia[50].

Association with Childhood Leukemia Risk

As breastfeeding is noted for providing the newborn with passive immunity, protecting from some early infections, as well as playing a role in immune system development, investigators have hypothesized that breastfeeding could reduce the risk of childhood leukemia. However, the results of studies of breastfeeding and its role in the risk of childhood leukemia have been inconsistent.

Guise et al. published a meta-analysis in 2005 reviewing case-control studies related to breastfeeding and risk of childhood leukemia[83]. Of the 10 studies reviewed, only two were considered to provide "good" quality evidence regarding the association between maternal breastfeeding and childhood leukemia[83]. In the two studies rated as good, breastfeeding was associated with a significant risk reduction in one study [OR: 0.80 (95% CI: 0.69–0.93) controlling for maternal education, race, income], as well as decreased risk in a dose-response with breastfeeding – one month of breastfeeding demonstrated no reduced risk, while 6 months of breastfeeding reduced risk for ALL and AML [70,83]. The other study rated as "good", the United Kingdom Childhood Cancer Study (UKCCS), reported a non-significant but suggestive difference (OR: 0.91 (95% CI: 0.81 –1.04)), and reported that breastfeeding for greater than 6 months had an OR of 0.65 (95% CI: 0.43–1.0)[84]. In the two studies rated as of fair quality, one was associated with risk reduction. Taken together, half of the studies associated breastfeeding with a lower risk of ALL[83].

Two other meta-analyses, performed by Beral et al. and Kwan et al., respectively, reported a reduced risk association between ever been breastfed and childhood leukemia[84]. Beral et al.

reported an OR for ever breastfeeding of 0.86; 95% CI: 0.81–0.92[84]. Breastfeeding for 6 months seemed to confer somewhat greater protection with an OR of 0.78 (95% CI: 0.71–0.85)[84]. The meta-analysis by Kwan et al. used 14 case-control studies and concluded that both short-term (less than 6 months) and long-term (greater than 6 months) breastfeeding was associated with a reduced odds of both ALL (OR: 0.76; 95% CI: 0.68 – 0.84) and AML (OR: 0.85; 95% CI: 0.73–0.98)[85].

Guise et al. noted that in the review performed by Beral et al. that included the UKCCS study, the protective effect of breastfeeding was uniform for all pediatric cancers, and this could represent either a universal effect of an immunologic influence across all cancers or an inherent bias in the control ascertainment, indicative of confounding[83]. In regards to the meta-analysis performed by Kwan et al., they noted that the lack of a relationship specific to ALL and the lack of a duration effect of breastfeeding may be indicative of bias, particularly confounding by socioeconomic status[83]. Many included studies failed to adjust for socioeconomic status, and the imbalance in this factor between cases and controls as well as participation bias may have confounded the relationship with breastfeeding[83]. Mothers who breastfeed for 6 months differ from those who breastfeed less, in more ways than just having a different socioeconomic status, such as age, pre-pregnancy BMI, income, education, race, and gravidity[83]. Risk for ALL may be associated with one of these factors, and the effect may not be removed entirely by adjusting for socioeconomic status[83].

Childhood leukemia is the most common type of cancer diagnosed in children, and it is estimated that the United States spends $1.4 billion dollars annually for the treatment of this disease[13,86]. An ability to prevent 10% to 20% of the 3000 cases each year in the USA through breastfeeding would be a health and fiscal benefit. Despite the public health importance of identifying potential interventions that could prevent the onset of childhood leukemia, the current literature regarding the preventive role of breastfeeding has substantial limitations[83]. Given the burden of disease and potential cost-effectiveness, conducting high-quality research should be a high priority[83].

The California Childhood Leukemia Study

The Northern California Childhood Leukemia Study (NCCLS), now called the California Childhood Leukemia Study (CCLS) is an ongoing population-based, matched case–control study in California[87]. It began in 1995 with 17 counties in Northern California with the objective of identifying the etiology of childhood leukemia. The study ran from 1995-1999 in the San Francisco Bay Area, and then expanded in 1999 to include an additional 18 counties in the California Central Valley County. In 2008, the study expanded to include counties from Southern California. Cases are eligible for the study if they were <15 years of age at diagnosis, had an English or Spanish-speaking parent or guardian, lived in one of the counties that comprised the population base at the time of diagnosis and had never been previously diagnosed with leukemia. Cases were ascertained within 72 hours after diagnosis at the Northern and Central California hospitals participating in the study. The control subjects were randomly selected from groups of four birth certificates obtained through the California Office of Vital Records, and one or two control subjects were matched to case subjects on child's date of birth (within 10 days), sex, Hispanic status (defined as either one or both parents being Hispanic, as indicated on the birth certificate record) and maternal race (as indicated on the birth certificate

record). Using CCLS study data, Ma et al. compared birth certificate control subjects with 'ideal' control subjects (California birth certificated records that were exactly population based, for individuals that did not need to be traced) and found little difference in demographic characteristics between the two, suggesting that the CCLS is approximately population based. Case and control participants are similar with respect to matching characteristics, but differ by household income, maternal education and maternal age at birth, all higher among controls[87].

Kwan et al. analyzed breastfeeding and risk of childhood leukemia using CCLS data available from 1995-2002[88]. Information regarding breastfeeding and complementary feeding characteristics was collected by an in-home interview and a self- administered questionnaire, respectively[88]. Most often, the biological mother provided the information on both instruments (95%). Respondents were asked if they ever breastfed their child for at least 1 day (ever/never) and for how long (in months, weeks, or days)[88]. Specific feeding characteristics of interest were the age the child started drinking milk or formula, the type of milk or formula consumed at or before 6 months and after 6 months of age, the age at which the child started eating solid foods, and the type of solid food consumed[88].

After adjusting for household income and maternal education, ever compared to never breastfed (OR 0.99; 95% CI 0.64 – 1.55) and breastfeeding duration in months (OR 1.00; 95% CI 0.98 – 1.02) were not associated with risk of ALL[88]. Similarly, when compared to no breastfeeding, breastfeeding less than or equal to 3 months, 4–6 months, 7–12 months, and greater than 13 months were not associated with ALL risk, and the P-value for trend across the categories was not significant[88]. In addition, exclusivity of breastfeeding was examined to assess the independent biological effects of breast milk on the risk of ALL[88]. Feeding only breast milk for any length of time was not associated with ALL risk, and no significant trend across the categories existed[88]. Restricting the analysis to ALL cases and their respective controls diagnosed from age 2 – 5 years revealed elevated, statistically non-significant effect estimates for breastfeeding and risk of disease. This study provides no evidence that breastfeeding is associated with the risk of childhood ALL[88].

Conclusion

Research has shown that a child's early life exposure to infectious diseases can affect risk of childhood leukemia – early childcare attendance, birth order, vaccination history, and ear infections have all been shown to be protective[53,55,65,69]. This supports Greaves' hypothesis that exposure to early common infectious diseases provides early immune system modulation, and a delay in exposure to nonspecific common infections increases the risk of an adverse immunologic response later in life and developing childhood leukemia[50].

As demonstrated, breastfeeding has been associated with reduced risk of ear infections, gastrointestinal, and respiratory infections, as well as an increased immune response to vaccines in breastfed versus formula-fed babies[69]. Furthermore, the diverse range of anti-inflammatory and immunomodulatory agents in breast milk may play an important biological role in priming the newborn immune system[68,74]. Meta-analyses have demonstrated association of breastfeeding

with a reduced risk of ALL but no association was observed in the CCLS with data collected from 1995-2002[83,85,88].

The association between breastfeeding and risk of childhood leukemia as examined by Kwan et al. in 2005 is limited in several ways[88]. First, it has a relatively small sample size. As of 2012, the CCLS has data available from 2003-2008 for analysis. As such, there is now a larger sample size with greater statistical power to assess the association. Furthermore, a potential association between breastfeeding during the infants first 6 months of life and childhood leukemia may have been missed in the analysis[88]. Second, no additional interactions were examined with other immunomodulatory factors, such as proxy measures of early life exposure to infectious diseases that have been previously shown to be associated with reduction in risk of ALL[53,88].

Given the evidence for the hypothesis that childhood ALL may result from an adverse immunologic response to a delay in exposure to nonspecific common infections, and the evidence of role of breast milk in protecting newborns from infections as well as priming the immune system, an additional analysis examining breastfeeding and risk of ALL is warranted. Specifically, utilizing the larger sample size available through the CCLS, it is imperative to examine the interaction of the effect of breastfeeding with proxy measures of early life exposure to infectious diseases (day care attendance, ear infections, birth order, vaccination history) on the risk of childhood leukemia. Stratification of analysis for the risk for ALL and AML can also be conducted using the larger data set. Based on the immunomodulatory components of breast milk, as well as data supporting that a IL12A polymorphism is associated with increased risk of ALL, an analysis examining whether the effects of breastfeeding are modified by genotypes of adaptive immunity genes would also make a contribution to our knowledge of ALL[67,71].

Paper 1 References

1. Greaves M. Molecular genetics, natural history and the demise of childhood leukaemia. *Eur. J. Cancer*. 1999;35(14):1941–1953.

2. Pui C-H, Mullighan CG, Evans WE, Relling MV. Pediatric acute lymphoblastic leukemia: where are we going and how do we get there? *Blood*. 2012;120(6):1165–1174. doi:10.1182/blood-2012-05-378943.

3. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012;62(1):10–29. doi:10.3322/caac.20138.

4. Wiemels J. Perspectives on the causes of childhood leukemia. *Chem. Biol. Interact.* 2012;196(3):59–67. doi:10.1016/j.cbi.2012.01.007.

5. Smith MA, Seibel NL, Altekruse SF, et al. Outcomes for Children and Adolescents With Cancer: Challenges for the Twenty-First Century. *Journal of Clinical Oncology*. 2010;28(15):2625–2634. doi:10.1200/JCO.2009.27.0421.

6. Siegel R, DeSantis C, Virgo K, et al. Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin*. 2012;62(4):220–241. doi:10.3322/caac.21149.

7. Li J, Thompson TD, Miller JW, Pollack LA, Stewart SL. Cancer Incidence Among Children and Adolescents in the United States, 2001-2003. *Pediatrics*. 2008;121(6):e1470–e1477. doi:10.1542/peds.2007-2964.

8. Bleyer A, O'Leary M, Barr R, Ries L. Cancer epidemiology in older adolescents and young adults 15 to 29 years of age, including SEER incidence and survival: 1975-2000. *Cancer epidemiology in older adolescents and young adults 15 to 29 years of age, including SEER incidence and survival: 1975-2000*. 2006.

9. Belson M, Kingsley B, Holmes A. Risk factors for acute leukemia in children: a review. *Environ. Health Perspect.* 2007;115(1):138–145.

10. Horton TM, Philip SC. Overview of the presentation and classification of acute lymphoblastic leukemia in children. Park JR, Connor RF, Basow DS, eds. Available at: http://www.uptodate.com. Accessed 2013.

11. Kumar V, Abbas AK, Fausto N, Aster JC. *Robbins and Cotran Pathologic Basis of Disease, Professional Edition*. Saunders; 2009.

12. Pui C-H, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet*. 2008;371(9617):1030–1043. doi:10.1016/S0140-6736(08)60457-2.

13. David Schottenfeld John G. Searle Professor of Epidemiology University of Michigan School of Public Health (Emeritus), Division of Cancer Epidemiology and Genetics National Cancer Institute Joseph F. Fraumeni Jr. Director. *Cancer Epidemiology and Prevention*. Oxford University Press, USA; 2006.

14. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. In: Vol 114. 2009:937–951. doi:10.1182/blood-2009-03-209262.

15. Kaushansky K, Williams WJWJ. *Williams Hematology*. McGraw-Hill Medical Publishing; 2010.

16. Rubnitz JE, Gibson B, Smith FO. Acute myeloid leukemia. *Pediatr. Clin. North Am.* 2008;55(1):21–51– ix. doi:10.1016/j.pcl.2007.11.003.

17. Schiffer CA, Anastasi J. Clinical manifestations, pathologic features, and diagnosis of acute myeloid leukemia. Larson RA, ed. 2013:1–13. Available at: http://uptodate.com. Accessed 2013.

18. Puumala SE, Ross JA, Aplenc R, Spector LG. Epidemiology of childhood acute myeloid leukemia. *Pediatr Blood Cancer*. 2013;60(5):728–733. doi:10.1002/pbc.24464.

19. 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;119(1):34–43. doi:10.1182/blood-2011-04-347872.

20. Estlin E, Gilbertson R, Wynn R. *Pediatric Hematology and Oncology*. Wiley-Blackwell; 2011.

21. Pui C-H, Carroll WL, Meshinchi S, Arceci RJ. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *Journal of Clinical Oncology*. 2011;29(5):551–565. doi:10.1200/JCO.2010.30.7405.

22. Schiffer CA. Prognosis of acute myeloid leukemia. Larson RA, Connor RF, eds. 2013:1–25. Available at: http://uptodate.com. Accessed 2013.

23. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood*. 2012;120(16):3187–3205. doi:10.1182/blood-2012-03-362608.

24. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57–70.

25. Zelent A, Greaves M, Enver T. Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. *Oncogene*. 2004;23(24):4275–4283. doi:10.1038/sj.onc.1207672.

26. Wiemels J. Chromosomal translocations in childhood leukemia: natural history, mechanisms, and epidemiology. *J. Natl. Cancer Inst. Monographs*. 2008;(39):87–90. doi:10.1093/jncimonographs/lgn006.

27. Greaves MF, Maia AT, Wiemels JL, Ford AM. Leukemia in twins: lessons in natural history. *Blood*. 2003;102(7):2321–2333. doi:10.1182/blood-2002-12-3817.

28. Maia AT, Ford AM, Jalali GR, et al. Molecular tracking of leukemogenesis in a triplet

pregnancy. *Blood*. 2001;98(2):478–482.

29. Wiemels JL, Xiao Z, Buffler PA, et al. In utero origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia. *Blood*. 2002;99(10):3801–3805.

30. Gale KB, Ford AM, Repp R, et al. Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc. Natl. Acad. Sci. U.S.A.* 1997;94(25):13950–13954.

31. Mori H, Colman SM, Xiao Z, et al. Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc. Natl. Acad. Sci. U.S.A.* 2002;99(12):8242–8247. doi:10.1073/pnas.112218799.

32. Zuna J, Madzo J, Krejci O, et al. ETV6/RUNX1 (TEL/AML1) is a frequent prenatal first hit in childhood leukemia. *Blood*. 2011;117(1):368–9– author reply 370–1. doi:10.1182/blood-2010-09-309070.

33. Wiemels JL, Leonard BC, Wang Y, et al. Site-specific translocation and evidence of postnatal origin of the t(1;19) E2A-PBX1 fusion in childhood acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 2002;99(23):15101–15106. doi:10.1073/pnas.222481199.

34. Wiemels JL, Kang M, Chang JS, et al. Backtracking RAS mutations in high hyperdiploid childhood acute lymphoblastic leukemia. *Blood Cells, Molecules, and Diseases*. 2010;45(3):186–191. doi:10.1016/j.bcmd.2010.07.007.

35. Urayama KY, Chokkalingam AP, Manabe A, Mizutani S. Current evidence for an inherited genetic basis of childhood acute lymphoblastic leukemia. *Int. J. Hematol.* 2013;97(1):3–19. doi:10.1007/s12185-012-1220-9.

36. Scélo G, Metayer C, Zhang L, et al. Household exposure to paint and petroleum solvents, chromosomal translocations, and the risk of childhood leukemia. *Environ. Health Perspect.* 2009;117(1):133–139. doi:10.1289/ehp.11927.

37. Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. Diagnostic X-rays and risk of childhood leukaemia. *International Journal of Epidemiology*. 2010;39(6):1628–1637. doi:10.1093/ije/dyq162.

38. Little MP. Leukaemia following childhood radiation exposure in the Japanese atomic bomb survivors and in medically exposed groups. *Radiation Protection Dosimetry*. 2008;132(2):156–165. doi:10.1093/rpd/ncn264.

39. Chang JS. Parental Smoking and the Risk of Childhood Leukemia. *American Journal of Epidemiology*. 2006;163(12):1091–1100. doi:10.1093/aje/kwj143.

40. Turner MC, Wigle DT, Krewski D. Residential pesticides and childhood leukemia: a systematic review and meta-analysis. *Environ. Health Perspect.* 2010;118(1):33–41. doi:10.1289/ehp.0900966.

41. Petridou E, Ntouvelis E, Dessypris N, Terzidis A, Trichopoulos D, Childhood Hematology-Oncology Group. Maternal diet and acute lymphoblastic leukemia in young children. *Cancer Epidemiol. Biomarkers Prev.* 2005;14(8):1935–1939. doi:10.1158/1055-9965.EPI-05-0090.

42. Amigou A, Sermage-Faure C, Orsi L, et al. Road traffic and childhood leukemia: the ESCALE study (SFCE). *Environ. Health Perspect.* 2011;119(4):566–572. doi:10.1289/ehp.1002429.

43. Treviño LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat. Genet.* 2009;41(9):1001–1005. doi:10.1038/ng.432.

44. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat. Genet.* 2009;41(9):1006–1010. doi:10.1038/ng.430.

45. Sherborne AL, Hosking FJ, Prasad RB, et al. Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nat. Genet.* 2010;42(6):492–494. doi:10.1038/ng.585.

46. Chokkalingam AP, Buffler PA. Genetic susceptibility to childhood leukaemia. *Radiation Protection Dosimetry*. 2008;132(2):119–129. doi:10.1093/rpd/ncn255.

47. Greaves MF, Alexander FE. An infectious etiology for common acute lymphoblastic leukemia in childhood? *Leukemia*. 1993;7(3):349–360.

48. Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. *Lancet*. 1988;2(8624):1323–1327.

49. Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer*. 1995;71(1):1–5.

50. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006;6(3):193–203. doi:10.1038/nrc1816.

51. Kinlen L, Doll R. Population mixing and childhood leukaemia: Fallon and other US clusters. *Br J Cancer*. 2004;91(1):1–3. doi:10.1038/sj.bjc.6601982.

52. Francis SS, Selvin S, Yang W, Buffler PA, Wiemels JL. Unusual space-time patterning of the Fallon, Nevada leukemia cluster: Evidence of an infectious etiology. *Chem. Biol. Interact.* 2012;196(3):102–109. doi:10.1016/j.cbi.2011.02.019.

53. Urayama KY, Ma X, Selvin S, et al. Early life exposure to infections and risk of childhood acute lymphoblastic leukemia. *Int. J. Cancer*. 2010;128(7):1632–1643. doi:10.1002/ijc.25752.

54. Ma X, Urayama K, Chang J, Wiemels JL, Buffler PA. Infection and pediatric acute lymphoblastic leukemia. *Blood Cells, Molecules, and Diseases*. 2009;42(2):117–120. doi:10.1016/j.bcmd.2008.10.006.

55. Ma X. Vaccination history and risk of childhood leukaemia. *International Journal of Epidemiology*. 2005;34(5):1100–1109. doi:10.1093/ije/dyi113.

56. Chang JS, Buffler PA, Metayer C, et al. Maternal Immunoglobulin E and Childhood Leukemia. *Cancer Epidemiology Biomarkers & Prevention*. 2009;18(8):2221–2227. doi:10.1158/1055-9965.EPI-09-0212.

57. Chang JS, Tsai Y-W, Tsai C-R, Wiemels JL. Allergy and risk of childhood acute lymphoblastic leukemia: a population-based and record-based study. *American Journal of Epidemiology*. 2012;176(11):970–978. doi:10.1093/aje/kws263.

58. Urayama KY, Buffler PA, Gallagher ER, Ayoob JM, Ma X. A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia. *International Journal of Epidemiology*. 2010;39(3):718–732. doi:10.1093/ije/dyp378.

59. Groves F, Auvinen A, Hakulinen T. Haemophilus influenzae type b vaccination and risk of childhood leukemia in a vaccine trial in finland. *Ann Epidemiol*. 2000;10(7):474.

60. Ribeiro KB, Buffler PA, Metayer C. Socioeconomic status and childhood acute lymphocytic leukemia incidence in São Paulo, Brazil. *Int. J. Cancer*. 2008;123(8):1907–1912. doi:10.1002/ijc.23738.

61. Spector L, Groves F, DeStefano F, et al. Medically recorded allergies and the risk of childhood acute lymphoblastic leukaemia. *Eur. J. Cancer*. 2004;40(4):579–584. doi:10.1016/j.ejca.2003.08.024.

62. Strachan DP. Hay fever, hygiene, and household size. *BMJ*. 1989;299(6710):1259–1260.

63. Roman E, Simpson J, Ansell P, et al. Childhood acute lymphoblastic leukemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study. *American Journal of Epidemiology*. 2007;165(5):496–504. doi:10.1093/aje/kwk039.

64. Crouch S, Lightfoot T, Simpson J, Smith A, Ansell P, Roman E. Infectious illness in children subsequently diagnosed with acute lymphoblastic leukemia: modeling the trends from birth to diagnosis. *American Journal of Epidemiology*. 2012;176(5):402–408. doi:10.1093/aje/kws180.

65. Simpson J, Smith A, Ansell P, Roman E. Childhood leukaemia and infectious exposure: a report from the United Kingdom Childhood Cancer Study (UKCCS). *Eur. J. Cancer*. 2007;43(16):2396–2403. doi:10.1016/j.ejca.2007.07.027.

66. Chang JS, Zhou M, Buffler PA, Chokkalingam AP, Metayer C, Wiemels JL. Profound Deficit of IL10 at Birth in Children Who Develop Childhood Acute Lymphoblastic Leukemia. *Cancer Epidemiology Biomarkers & Prevention*. 2011;20(8):1736–1740. doi:10.1158/1055-9965.EPI-11-0162.

67. Chang JS, Wiemels JL, Chokkalingam AP, et al. Genetic Polymorphisms in Adaptive Immunity Genes and Childhood Acute Lymphoblastic Leukemia. *Cancer Epidemiology Biomarkers & Prevention*. 2010;19(9):2152–2163. doi:10.1158/1055-9965.EPI-10-0389.

68. Labbok MH, Clark D, Goldman AS. Breastfeeding: maintaining an irreplaceable immunological resource. *Nat. Rev. Immunol.* 2004;4(7):565–572. doi:10.1038/nri1393.

69. Ip S, Chung M, Raman G, Trikalinos TA, Lau J. A summary of the Agency for Healthcare Research and Quality's evidence report on breastfeeding in developed countries. *Breastfeed Med*. 2009;4 Suppl 1:S17–30. doi:10.1089/bfm.2009.0050.

70. Shu XO, Linet MS, Steinbuch M, et al. Breast-feeding and risk of childhood acute leukemia. *J. Natl. Cancer Inst.* 1999;91(20):1765–1772.

71. Palmer AC. Nutritionally mediated programming of the developing immune system. *Adv Nutr*. 2011;2(5):377–395. doi:10.3945/an.111.000570.

72. Jackson KM, Nazar AM. Breastfeeding, the immune response, and long-term health. *J Am Osteopath Assoc*. 2006;106(4):203–207.

73. Goldman AS. The immune system in human milk and the developing infant. *Breastfeed Med*. 2007;2(4):195–204. doi:10.1089/bfm.2007.0024.

74. Clinton C. Development of the Infant Immune Function and the Effects of Breast Milk. *Natural Medicine Journal*. 2010:1–4.

75. McVeagh P, Miller JB. Human milk oligosaccharides: only the breast. *J Paediatr Child Health*. 1997;33(4):281–286.

76. Liepke C, Adermann K, Raida M, Mägert H-J, Forssmann W-G, Zucht H-D. Human milk provides peptides highly stimulating the growth of bifidobacteria. *Eur. J. Biochem.* 2002;269(2):712–718.

77. Walker A. Breast milk as the gold standard for protective nutrients. *J. Pediatr.* 2010;156(2 Suppl):S3–7. doi:10.1016/j.jpeds.2009.11.021.

78. Morgan JA, Young L, McGuire W. Pathogenesis and prevention of necrotizing enterocolitis. *Curr. Opin. Infect. Dis.* 2011;24(3):183–189. doi:10.1097/QCO.0b013e328345d5b5.

79. Garofalo R. Cytokines in human milk. *J. Pediatr.* 2010;156(2 Suppl):S36–40. doi:10.1016/j.jpeds.2009.11.019.

80. Macfarlane TV, Seager AL, Moller M, Morgan G, Thornton CA. Thymic stromal lymphopoietin is present in human breast milk. *Pediatr Allergy Immunol*. 2010;21(2 Pt 2):e454–6. doi:10.1111/j.1399-3038.2009.00916.x.

81. Belderbos ME, Houben ML, van Bleek GM, et al. Breastfeeding modulates neonatal innate immune responses: a prospective birth cohort study. *Pediatr Allergy Immunol*. 2012;23(1):65–74. doi:10.1111/j.1399-3038.2011.01230.x.

82. M'Rabet L, Vos AP, Boehm G, Garssen J. Breast-feeding and its role in early development of the immune system in infants: consequences for health later in life. *J. Nutr.* 2008;138(9):1782S–

1790S.

83. Guise J-M, Austin D, Morris CD. Review of case-control studies related to breastfeeding and reduced risk of childhood leukemia. *Pediatrics*. 2005;116(5):e724–31. doi:10.1542/peds.2005-0636.

84. UK Childhood Cancer Study Investigators. Breastfeeding and childhood cancer. *Br J Cancer*. 2001;85(11):1685–1694. doi:10.1054/bjoc.2001.2110.

85. Kwan ML, Buffler PA, Abrams B, Kiley VA. Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Rep*. 2004;119(6):521–535. doi:10.1016/j.phr.2004.09.002.

86. Smith MA, Ries LAG, Gurney JG, Ross JA. Leukemia - SEER Pediatric Monograph. 1999:1–18.

87. Ma X, Buffler PA, Layefsky M, Does MB, Reynolds P. Control selection strategies in case-control studies of childhood diseases. *American Journal of Epidemiology*. 2004;159(10):915–921.

88. Kwan ML, Buffler PA, Wiemels JL, et al. Breastfeeding patterns and risk of childhood acute lymphoblastic leukaemia. *Br J Cancer*. 2005;93(3):379–384. doi:10.1038/sj.bjc.6602706.

Paper 2

Introduction

Childhood cancer is the second most common cause of death among children aged 1-14 years in the US after accidents, with leukemia accounting for approximately 30% of cancer diagnoses in this age range[1]. Acute lymphoblastic leukemia (ALL) occurs five times more frequently than acute myelogenous leukemia (AML) and accounts for 78% of all childhood leukemia diagnoses[1]. While the treatment of childhood leukemia is one of the greatest success stories in the field of oncology, the etiology of childhood leukemia still remains unclear, with ionizing radiation and congenital genetic syndromes explaining less than 10% of ALL cases[2].

Childhood leukemia most likely follows a 'two-hit' mechanism, in which an initial 'hit' occurs during pregnancy establishing a pre-leukemic clone, followed by a second mutation in early childhood that leads to the development of leukemia[2]. One leading hypothesis suggests that the second 'hit' may occur when an underdeveloped immune system – a result of reduced exposure to nonspecific common infections early in life – responds aberrantly to infection later in life, hence increasing the risk of a second mutation[3]. Evidence in support of this hypothesis comes from a substantial body of literature using surrogate measures of exposure to infections, which show a reduced risk of ALL associated with increasing birth order, a child's history of ear infection, and history of day care attendance [2-4]. The results are the most significant when the exposure occurred in the first 6 months of the child's life[2]. This is in line with research showing that while a newborn's immune system is complete at birth, it is significantly down-regulated and skewed towards a $CD4^+$ T helper 2 (Th2) phenotype[5]. During the first 6 months, the immune system undergoes expansion and the priming of adaptive cell immunity to achieve a balance between $CD4^+$ T helper 1 (Th1) and Th2 activities[6,7]. As such, there is evidence to support a mechanism by which reduced external priming of the adaptive immune system in the first 6 months of life predisposes the child toward abnormal immune responses and proliferation of a pre-leukemic clone later in life[3]. On the other hand, an alternative hypothesis is that some children may already be born with a dysregulated immune system[8]. These children are more likely to develop severe and frequent infections, placing them at a higher risk of developing the second "hit", as suggested by studies reporting that an increased frequency of being ill in the first year of life is associated with an increased risk of ALL[9,10]. In addition, children with low levels of the anti-inflammatory cytokine interleukin-10 (IL-10) at birth have an increased risk of developing ALL, further supporting the hypothesis that some children are born with an underdeveloped immune system, increasing risk of developing ALL[11]. These opposite hypotheses likely represent two distinct groups of children – one lacking early immune system modulation from early life exposure to infections and another born with a dysregulated immune system more prone to infections, both of which may promote the second 'hit' of ALL[2].

An early life exposure that may play a significant role in the etiology of ALL is breastfeeding. Breastfeeding has been demonstrated to offer nutritional and immunologic benefits for newborns and young infants, including stimulating intestinal maturation, enhancing production of specific antibodies, controlling and preventing inflammation, encouraging proliferation of commensal bacteria, and priming the immune system[7,12,13]. Based on the documented short and long-term advantages of breastfeeding, the American Academy of Pediatrics (AAP) recommends exclusive

breastfeeding for 6 months followed by continued breastfeeding as complementary foods are introduced for a year or longer[14]. However, the results of studies of breastfeeding and its role in the risk of childhood leukemia have been inconsistent. Reports of some studies and meta-analyses show a significantly reduced risk reduction of ALL if the child has ever been breastfed, while other reports show no association between breastfeeding and risk of childhood leukemia[15-18]. In addition, a previous study by the California Childhood Leukemia Study (CCLS), using data collected from 1995-2002, showed no evidence of significant association between breastfeeding and ALL[19]. However, the study had a small number of cases (n = 311) and only controlled for income in the analysis of breastfeeding and risk of ALL[19].

This study analyzes the relationship between breastfeeding and the risk of ALL and the three most common-subtypes of ALL using data collected from the CCLS from 1995-2008. In particular, this study looks at breastfeeding and risk of ALL controlling for other early life (first 6 months) immune-modulating exposures (day care attendance, birth order, birth type), immune status (frequency of infections in first 6 months), and parental socioeconomic status and behavioral variables (smoking). With data from the CCLS from 1995-2008, there are 722 cases of non-infant ALL, providing greater statistical power to detect differences between breastfeeding patterns among cases and controls.

Methods

Participants and Procedures

The CCLS is an ongoing, population-based matched case-control study that began in 1995 to investigate the etiology of pediatric leukemias. Beginning in 1995, newly diagnosed childhood leukemia cases were ascertained at the time of diagnosis from major pediatric hospitals located in a 17-county San Francisco Bay area study region, which expanded in 1999 to 35 counties in Northern and Central California. Comparison with the California Cancer Registry (1997-2003) showed that the CCLS case ascertainment protocol captured about 95% of children diagnosed with leukemia in the participating study hospitals. For each eligible case, statewide birth records maintained by the Center for Health Statistics of the California Department of Public Health were used to generate a list of randomly selected controls that matched the case on child's date of birth, sex, Hispanic status (a biologic parent who is Hispanic), and maternal race. Information was obtained from the birth certificates and commercially available searching tools were used to trace and enroll 1 or 2 matched controls for each case. Cases and controls were considered eligible if they were under 15 years of age at date of diagnosis for cases (or corresponding reference date for controls), resided in the study region at the date of diagnosis, had a parent or guardian who spoke either English or Spanish, and had no prior history of malignancy. Approximately 85% of eligible cases consented to participate. Among all eligible controls who were contacted, 86% consented to participate. A previous evaluation showed that participating controls in the CCLS are representative of the sampled population with respect to parental age, parental education, and mother's reproductive history[20].

The current analysis includes ALL cases and control subjects recruited between 1995 and 2008. ALL cases and their matched controls who were below the age of one at diagnosis/reference date were excluded because of growing evidence that leukemias in this age range may be etiologically

distinct from leukemia diagnosed at later ages[2]. Cases and controls breastfed for >1 year were also excluded from analysis to investigate a possible causal effect; breastfeeding should precede the diagnosis of childhood leukemia. This resulted in a dataset with 669 ALL cases and 925 matched controls.

The study protocol was approved by the Institutional Review Boards of the University of California, Berkeley and all collaborating institutions. Written informed consent was obtained from all participating parents and guardians.

Data collection

Breastfeeding characteristics were collected by an in-home interview and a self-administered standardized questionnaire given to the parents/guardians of each child, with the biological mother providing information 95% of the time. Breastfeeding information was collected first as a binary variable (never/ever), and also as a continuous variable indicating the number of weeks that the child was breastfed (in weeks).

Data on the child's social contacts with children inside and outside the home (birth order and day care attendance) and common infections during the first year of life were also obtained using a standardized questionnaire. The child's birth order was determined based on a reproductive history of the biologic mother and used as a measure of number of siblings a child had. History of day care and preschool attendance, including age of attendance, duration of attendance (hours per week), and number of other children present, were used to calculate child-hours of day care attendance. As previously described in other publications[4], total child-hours of exposure at each day care center was calculated as: (number of months attending the day care center) × (mean hours per week at this day care center) × (number of other children at this day care center) × (4.35 weeks per month). For each child, the child-hours in each day care setting were summed to obtain the total child-hours of exposure, and for this study, restricted to day care child-hours in the first 6 month of life. The method of birth delivery (birth type), caesarian vs. vaginal birth, was extracted from birth certificates of cases and controls.

History of common illnesses focused on infections that the child had during the first year of life, such as severe diarrhea/vomiting, ear infection, persistent cough, mouth and eye infection, influenza, and unspecified “other infection” with an emphasis on the timing and frequency. For this study, immune status was defined as frequency of being ill more than 5 times in the first 6 months of life. Smoking history was obtained via self-administered questionnaire and in-person interview. Ever smoking was defined as ever having smoked 100 cigarettes before the case child’s diagnosis of leukemia. Smoking exposure specifically in the first 6 months was not available, and in this study, parental smoking was defined as if either the mother or the father ever smoked prior to the child’s diagnosis.

Immunophenotypes and cytogenetic data analyses for cases of ALL were performed either at the University of California, Berkeley or at a collaborating hospital. Immunophenotype was determined for ALL cases using flow cytometry profiles. Those positive for CD19 or CD10 were classified as B-lineage and were used to classify “common ALL” (c-ALL) subtype of ALL defined as CD19+, CD10+ among the ages of 2-5. For classification of the other common

subtypes, B-cell ALL with t(12;21) translocations (TEL-AML1) and high-hyperdiploid ALL (defined by a karyotype with 51-68 chromosomes), records were obtained from cytogenetic reports from clinical laboratories where classical banding methods were generally applied. All abstracted data were reviewed for accuracy by a consulting clinical oncologist. For cases without any cytogenetic reports or a "normal" karyotype, further karyotyping was done. Fluorescence in situ hybridization (FISH) with gene-loci specific probes for chromosomes 12 (TEL) and 21 (AML1) was used to identify TEL-AML1 translocations. Classifications were reviewed by CCLS and inconsistencies were resolved after discussion by experts, blinded to ethnicity.

Statistical analysis

The primary outcome variable of interest was overall ALL, with secondary analyses of the three common subtypes of ALL (common ALL, TEL-AML1 ALL, and high hyperdiploid ALL). Breastfeeding was modeled in three different ways; as a binary variable "never/ever", as a continuous variable (time in weeks), and as a categorical variable in exclusive intervals of 6 months (never, <6 months, 6-12 months). The 6-month intervals were chosen based on the analysis of prior literature which indicated that modulating exposures are significant in the first 6-months of exposure and based on AAP recommendations of breastfeeding exclusively for 6 months and afterwards, up to 1 year of age or longer not-exclusively. Other variables representing immune-modulating exposures include the number of day care child-hours in the first 6 months, birth order (no siblings or ≥ 1 sibling) and birth type (vaginal or caesarian). Frequency of infections (≥ 5 infections) in the first 6 months was used as a surrogate measure of reduced immune function.

Demographic characteristics included in the analysis were based on literature review and include maternal age at child's birth, maternal educational level, household income, and child's birth weight. Parental smoking was included in the analysis since paternal smoking has been associated with increased risk of ALL and maternal smoking was associated with lower rate of breastfeeding and of shorter duration[21,22]. The demographic characteristics, smoking, and immune-related variables are reported for cases and controls for overall ALL, as well as for the three common subtypes of ALL (common ALL, TEL-AML1 ALL, and high hyperdiploid ALL). Univariable conditional logistic regression was used to assess differences in the distribution of characteristics between the cases and controls for ALL and the subtypes of ALL, as a substitution for the McNemar's Test.

To assess the association of breastfeeding and risk of ALL and subtypes of ALL, a conditional logistic regression analysis was performed to estimate the odds ratios (ORs) and 95% confidence intervals (CIs). First, breastfeeding on risk of ALL and subtypes of ALL was examined as never/ever, continuous, and as exclusive intervals of 6 months (never, < 6 months, and 6-12 months), not adjusting for demographics, smoking, and immune-related variables. Then, a multivariable conditional logistic analysis was conducted to investigate the association between ALL and breastfeeding, with breastfeeding coded with indicators (never, <6 months, 6-12 months), adjusting for both immune-modulating exposures and demographics and select characteristics (maternal age, maternal education, income, child's birth weight, and parental smoking). All of the covariates were included in the multivariable model in order to compare the effect of breastfeeding across all models (ALL and its subtypes). To address the issue of power

when analyzing breastfeeding and risk of individual subtypes of ALL, breastfeeding categories "never" and "<6 months" were combined, since the impact of never and <6 months were similar and not significantly associated with risk of ALL. The binary breastfeeding variable (<6 months vs. ≥ 6 months) was then assessed for risk of each of the three subtypes of ALL, adjusting for immune-related variables, demographics, and smoking. Last, interaction terms describing the two-way multiplicative interactions between breastfeeding and all the covariates (including matching covariates – age, sex, Hispanic ethnicity/maternal race) were assessed for overall significance.

P-values of <0.05 were considered significant in the conditional logistic analyses, as well as for the overall Wald test for the set of breastfeeding indicators. Interactions with a p-value of <0.20 were considered significant. All statistical analyses were done using Stata version 12.1.

Results

Demographics and Smoking Distributions among Cases and Controls (Tables 1A and 1B)

Demographic characteristics of cases and controls are presented in Tables 1A and 1B for overall ALL and separately for the three most common subtypes of ALL: c-ALL, TEL-AML1 ALL, and high hyperdiploid ALL. For overall ALL, compared to controls, more cases came from families with a lower annual household income ($p < 0.001$), were born to mothers with fewer years of education ($p = 0.001$), were born with a higher birthweight ($p = 0.046$) and were born to families in which one or both parents had previously smoked cigarettes ($p = 0.008$). White/Caucasians were the largest racial group in the study population (53.1%), of which 46.2% were Hispanic and 36.3% were Non-Hispanic White. Mixed or 'other' made the third largest subgroup of the study (32.7%), followed by Asian/Pacific Islanders (8.9%), African Americans (2.8%) and Native American (1.2%). There was a higher prevalence of ALL among boys (56.8%) than girls (43.2%), and the average age of diagnosis of ALL was 5.6 years (SD ± 3.4).

For c-ALL, cases and controls differed significantly with respect to household income ($P <0.001$) and parental smoking ($p = 0.032$). TEL-AML1 ALL cases and controls differed significantly for household income only ($p = 0.002$). For high-hyperdiploid ALL, cases and controls differed only by childbirth weight ($p = 0.016$). For all three subtypes of ALL, there was a higher prevalence of leukemia in boys when compared to girls, with the average age of diagnosis being 3.8 years for c-ALL, 4.3 years for TEL-AML1 ALL, and 5.2 years for high-hyperdiploid ALL.

Maternal age at the child's birth did not statistically differ in distribution between cases and controls for overall ALL and for any of the subtypes.

Distributions of Immune-Related Variables among Cases and Controls (Tables 2A and 2B)

The distribution of other immune-modulating exposures are as follows: cases had fewer child-hours of day care by 6 months (324.4 SD±1281) than controls (469 SD±1477), were more likely to be a first-born child or an only child (73.7% vs. 73.2% among controls), more likely to have had a Caesarian section for birth (22.4%) versus controls (20.6%), and more likely to have been reported to have been sick more than 5 times in the first 6 months of life (8.5%) than controls

(6.4%). However, only birth type (Caesarian section vs. vaginal birth) was significantly different between cases and controls (p = 0.047). Among subtypes of ALL, cases on average, had fewer child day care hours by 6 months of age, were more likely to be first born or an only child, were more likely to have been born by cesarean section, and were more likely to have been sick in the first 6 months of life. However, these distributions between cases and controls among subtypes of ALL were not significant.

Distribution of Breastfeeding among Cases and Controls (Tables 3A)

The frequency of breastfeeding was high among cases and controls for overall ALL and its subtypes, with >90% having ever breastfed in any duration. For ALL overall, the prevalence of ever breastfeeding was 90.3% among cases and 93.2% among controls. On average, cases were breastfed for shorter periods (28.4 weeks, SD±16.2) than controls (31.2 weeks, SD±14.5) and only 64.0% of cases were breastfed for greater than 6 months, compared to 72.9% of controls.

Unadjusted Models of Breastfeeding and Risk of ALL and Subtypes of ALL (Table 3B)

When compared to never being breastfed, breastfeeding was significantly protective against overall ALL in the unadjusted model (OR: 0.63, 95% CI: 0.40-1.00, p = 0.048). Breastfeeding modeled as a continuous variable, was similarly protective against overall ALL (OR: 0.98, 95% CI: 0.97-0.99, p = 0.001). Furthermore, when compared to never breastfeeding, breastfeeding for <6 months was not significant (though the point estimate suggests a protective effect), but breastfeeding for 6-12 months was significantly associated with a reduced risk of ALL (OR: 0.40, 95% CI: 0.23-0.68, p = 0.001).

When looking at subtypes of ALL, unadjusted breastfeeding was not significant in c-ALL or TEL-AML1, although point estimates suggest a protective effect. However, breastfeeding, as a continuous variable, was significantly associated with a reduced risk of high-hyperdiploid ALL (OR: 0.98, 95% CI: 0.96-1.00, p = 0.025). In addition, compared to never, breastfeeding for <6 months was not significant, but breastfeeding for 6-12 months was associated with a significantly reduced risk of high-hyperdiploid ALL (OR: 0.37, 95% CI: 0.14 - 0.96, p = 0.04).

Breastfeeding and Risk of ALL, Adjusting for Immune-related Variables, Demographics, and Smoking (Table 4A and 4B)

Adjusting for demographics and other selected covariates (mother's age, maternal education, household income, child's birth weight, parental smoking), and for immune-related variables exposures (day care, birth order, birth type, total times ill in the 6 months), breastfeeding 6-12 months, compared to never, was significantly associated with a decreased risk of ALL (OR: 0.49, 95% CI: 0.26-0.95, p = 0.033). Compared to <6 months of breastfeeding, breastfeeding for 6-12 was also significantly associated with a reduced risk of ALL (OR: 0.56, 95%CI: 0.34-0.92, p = 0.023), after adjusting for all of the covariates (not shown in Table 4A). However, when looking at subtypes of ALL, breastfeeding (coded as indicators) was not significantly associated with risk of common ALL, TEL-AML1 ALL, or high-hyperdiploid ALL, even though the point estimates suggest that breastfeeding for 6-12 months confers a protective effect.

Breastfeeding (as a binary variable) and risk of ALL, Adjusting for Immune-related Factors, Demographics, and Smoking (Table 5A)

To address the issue of reduced power when assessing the association of breastfeeding and subtypes of ALL, and since never and < 6 months did not have a statistically significant difference in outcomes, breastfeeding was coded as a binary variable as <6 months breastfeeding (including never) vs. ≥ 6 months of breastfeeding. Adjusting for immune-modulating exposures and covariates, compared to breastfeeding for <6 months, breastfeeding for ≥6 months was significantly associated with a reduced risk of high-hyperdiploid ALL (OR: 0.40, 95% CI: 0.17-0.96, p = 0.040). Breastfeeding was not significantly associated with common ALL or TEL-AML1 ALL, although point estimates suggest a protective effect when breastfeeding occurs for ≥6 months.

Interaction models

Two-way cross products between breastfeeding and immune-modulating exposures, as well as the selected covariates were created to assess for interaction between breastfeeding and the other variables on risk of ALL and its subtypes. There was no significant interaction between breastfeeding and any of the immune-modulating exposures (day care attendance in the first 6 months, birth order, birth type, and frequency of infections in the first 6 months) on risk of ALL. In addition, there was no significant interaction between breastfeeding and matching variables (gender, race, ethnicity) or covariates (maternal education, household income, child's birth weight, and parental smoking).

Discussion

This study found an association between breastfeeding for 6-12 months, compared to never and < 6 months, and a reduced risk of overall ALL in subjects enrolled in the CCLS from 1995-2008, controlling for other immune-modulating exposures, demographic characteristics and parental smoking. Specifically, compared to children who were never breastfed, children who were breastfed for 6-12 months had 0.49 times the odds of developing leukemia (95% CI: 0.26-0.95, p = 0.033), and compared to children who were breast fed for <6 months, children who were breastfed for 6-12 months had 0.56 times the odds of developing leukemia (95%CI: 0.34-0.92, p = 0.023). Furthermore, since ALL is a heterogeneous disease, stratified analysis by the three most common subtypes demonstrated that breastfeeding for ≥ 6 months, compared to < 6 months, was significantly associated with a reduced risk of high-hyperdiploid ALL (OR: 0.40, 95% CI: 0.17-0.96, p = 0.040). These are new findings for the CCLS, as a previous report showed no association between any duration of breastfeeding and risk of ALL.

Our results are supported by other studies that showed breastfeeding for greater than 6 months is significantly associated with a decreased risk of ALL. The United Kingdom Childhood Cancer Study (UKCSS) reported that infants breastfed for 6 months, compared to never, had an OR of 0.65 (95% CI: 0.43–1.0)[15]. A meta-analysis by Beral et al. reported an OR for 6 months of breastfeeding of 0.78 (95% CI: 0.71–0.85), compared to less than 6 months. A meta-analysis by Kwan et al. showed that breastfeeding for greater than 6 months compared to less than 6 months

of breastfeeding was associated with a reduced odds of ALL (OR: 0.76; 95% CI: 0.68 – 0.84)[16,23].

The association between breastfeeding and reduced risk of high-hyperdiploid ALL is also indirectly supported by a recent study that showed that high-hyperdiploid ALL was strongly associated with a variant of the class II HLA gene, HLA-DPI, and that this variant also had a significant interaction with breastfeeding on the risk of ALL ($P_{interaction} = 0.094$)[24]. That is, children with the HLA-DP1 supertype who were not breastfed had 3 times the odds of developing ALL versus children breastfed and had the HLA-DP1 supertype (OR = 3.04; 95% CI, 1.26-7.30)[24].

With breastfeeding significantly associated with reduced risk of overall ALL and high-hyperdiploid ALL, we hypothesized that there would be interaction between breastfeeding and other protective immune-modulating exposures on risk of overall ALL and ALL subtypes. However, none of the interactions were significant. The lack of evidence for interactions between breastfeeding with day care child-hours, birth order, birth type, and frequency of infections, may be due to limited statistical power from the smaller cell sizes created by stratification.

Our results suggest that breastfeeding and other immune-modulating exposures are independently associated with a protective effect for overall ALL and its subtypes, specifically high-hyperdiploid ALL. The significant association with breastfeeding for 6-12 months, that is children that have been breastfed for the full 6 months, suggests that breastfeeding plays a significant role in priming an infant's immune system. This supports the leading hypothesis that "priming" the immune system in the early stages of life decreases risk of an aberrant immune response to infection later in life, hence decreasing the risk of getting a "second" hit. However, another hypothesis could explain the association between breastfeeding and the reduced risk of high-hyperdiploid ALL.

High-hyperdiploid ALL, which accounts for 20-30% of childhood ALL, is associated with extra chromosomes such as 4, 6, 10, 14, 18, 21 and X, resulting in anywhere from 51-68 chromosomes in one hyperdiploid clone[25]. Often the extra chromosomes result in trisomy, rather than tetrasomy[25]. Studies have shown that all the chromosomes are gained in a single abnormal mitosis that occurs prenatally, that is, a high hyperdiploid B-cell is the "first" hit in the process of leukemogenesis[2,25]. High-hyperdiploid B-cells have been shown by several studies to have a propensity for apoptosis, which may account for the high cure rates among patients with high-hyperdiploid ALL[26,27]. Another hypothesis for the significant protective association between breastfeeding and risk of high-hyperdiploid ALL is that early and constant immune priming through breastfeeding in the critical first 6 months of life causes stimulation of B-cells, and the replication of the high-hyperdiploid B-cell clone triggers apoptosis since many of the trisomic chromosome copies can not be viably sustained during mitotic divisions. This hypothesis may be extended to other immune-modulating exposures, such as birth type, where vaginal birthing exposes the infant to vaginal microflora that then play a role in immune stimulation and inducement of apoptosis among the high-hyperdiploid B-cells. One way of further testing this hypothesis is to examine whether the Ras oncogene mutation, shown to be a "second hit" in the genesis of high-hyperdiploid ALL, varies at all by breastfeeding status among cases of high-

hyperdiploid ALL[28,29]. The association of reduced risk of high-hyperdiploid ALL with being breastfed for greater than 6 months is quite interesting, and it is plausible that immune system priming in the first 6 months of life through breastfeeding may negatively select against the hyperdiploid B-cells, hence preventing the development of secondary mutations such as Ras oncogene.

Our study presents several strengths and limitations. An inherent limitation of the CCLS is recall bias. The exposure histories were obtained by self-report from the parent/guardian, and after diagnosis. This is a potential drawback of any case-control study since biologic mothers of cases may recall certain exposures differently and may be more influenced by the knowledge of their child's diagnosis than biologic mothers of controls. CCLS attempted to minimize reporting differences by mailing preparatory materials to serve as 'aid memoires' for the respondents prior to the in-home interview. Furthermore, this analysis is limited by the fact that not all cases had matched controls, and were dropped by the conditional logistic regression model. There were 117 cases dropped from the analysis that did not have a matching control, in addition to 53 cases being excluded who breastfed for >1 year, leading to a final sample of 669 cases. This created a reduction in power when running the analyses by subtypes of ALL.

Strengths of the CCLS include a population-based selection of controls and detailed exposure assessment. The method of selecting controls from the population-based statewide birth registry ensures that controls are identified from the same study base as cases (as detailed above). In addition, by controlling for socioeconomic status through the demographic variables, as well as for immune-related factors, and smoking, this study truly aimed to isolate the effect of breastfeeding on risk of ALL, eliminating other potential confounders and contributors to ALL.

In conclusion, this investigation was a comprehensive examination of breastfeeding patterns and the risk of childhood ALL and its three common subtypes in a population-based case-control study. We observed a significantly reduced risk of overall childhood ALL among participants who were breastfed for 6-12 months compared to those who were never breastfed or were breastfed for less than 6 months. In addition, we found a significantly reduced risk of high-hyperdiploid ALL among children breastfeed for 6-12 months compared to less than 6 months. No significant association between breastfeeding and risk of the c-ALL or TEL-AML1 ALL subtypes was observed. Overall, our results support the hypothesis that breastfeeding protects against the risk of childhood ALL.

Paper 2 References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012;62(1):10–29. doi:10.3322/caac.20138.

2. Wiemels J. Perspectives on the causes of childhood leukemia. *Chem Biol Interact*. 2012;196(3):59–67. doi:10.1016/j.cbi.2012.01.007.

3. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006;6(3):193–203. doi:10.1038/nrc1816.

4. Urayama KY, Ma X, Selvin S, et al. Early life exposure to infections and risk of childhood acute lymphoblastic leukemia. *Int J Cancer*. 2010;128(7):1632–1643. doi:10.1002/ijc.25752.

5. Levy O. Innate immunity of the newborn: basic mechanisms and clinical correlates. *Nat Rev Immunol*. 2007;7(5):379–390. doi:10.1038/nri2075.

6. Palmer AC. Nutritionally mediated programming of the developing immune system. *Adv Nutr*. 2011;2(5):377–395. doi:10.3945/an.111.000570.

7. Clinton C. Development of the Infant Immune Function and the Effects of Breast Milk. *Natural Medicine Journal*. 2010:1–4.

8. Urayama KY, Chokkalingam AP, Manabe A, Mizutani S. Current evidence for an inherited genetic basis of childhood acute lymphoblastic leukemia. *Int J Hematol*. 2013;97(1):3–19. doi:10.1007/s12185-012-1220-9.

9. Chang JS, Tsai C-R, Tsai Y-W, Wiemels JL. Medically diagnosed infections and risk of childhood leukaemia: a population-based case-control study. *International Journal of Epidemiology*. 2012;41(4):1050–1059. doi:10.1093/ije/dys113.

10. Simpson J, Smith A, Ansell P, Roman E. Childhood leukaemia and infectious exposure: a report from the United Kingdom Childhood Cancer Study (UKCCS). *Eur J Cancer*. 2007;43(16):2396–2403. doi:10.1016/j.ejca.2007.07.027.

11. Chang JS, Zhou M, Buffler PA, Chokkalingam AP, Metayer C, Wiemels JL. Profound Deficit of IL10 at Birth in Children Who Develop Childhood Acute Lymphoblastic Leukemia. *Cancer Epidemiology Biomarkers & Prevention*. 2011;20(8):1736–1740. doi:10.1158/1055-9965.EPI-11-0162.

12. Ip S, Chung M, Raman G, Trikalinos TA, Lau J. A summary of the Agency for Healthcare Research and Quality's evidence report on breastfeeding in developed countries. *Breastfeed Med*. 2009;4 Suppl 1:S17–30. doi:10.1089/bfm.2009.0050.

13. Labbok MH, Clark D, Goldman AS. Breastfeeding: maintaining an irreplaceable immunological resource. *Nat Rev Immunol*. 2004;4(7):565–572. doi:10.1038/nri1393.

14. SECTION ON BREASTFEEDING. Breastfeeding and the Use of Human Milk. *Pediatrics*. 2012;129(3):e827–e841. doi:10.1542/peds.2011-3552.

15. UK Childhood Cancer Study Investigators. Breastfeeding and childhood cancer. *Br J Cancer*. 2001;85(11):1685–1694. doi:10.1054/bjoc.2001.2110.

16. Kwan ML, Buffler PA, Abrams B, Kiley VA. Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Rep*. 2004;119(6):521–535. doi:10.1016/j.phr.2004.09.002.

17. Guise J-M, Austin D, Morris CD. Review of case-control studies related to breastfeeding and reduced risk of childhood leukemia. *Pediatrics*. 2005;116(5):e724–31. doi:10.1542/peds.2005-0636.

18. Shu XO, Linet MS, Steinbuch M, et al. Breast-feeding and risk of childhood acute leukemia. *J Natl Cancer Inst*. 1999;91(20):1765–1772.

19. Kwan ML, Buffler PA, Wiemels JL, et al. Breastfeeding patterns and risk of childhood acute lymphoblastic leukaemia. *Br J Cancer*. 2005;93(3):379–384. doi:10.1038/sj.bjc.6602706.

20. Ma X, Buffler PA, Layefsky M, Does MB, Reynolds P. Control selection strategies in case-control studies of childhood diseases. *American Journal of Epidemiology*. 2004;159(10):915–921.

21. Weiser TM, Lin M, Garikapaty V, Feyerharm RW, Bensyl DM, Zhu B-P. Association of maternal smoking status with breastfeeding practices: Missouri, 2005. *Pediatrics*. 2009;124(6):1603–1610. doi:10.1542/peds.2008-2711.

22. Metayer C, Zhang L, Wiemels JL, et al. Tobacco smoke exposure and the risk of childhood acute lymphoblastic and myeloid leukemias by cytogenetic subtype. *Cancer Epidemiology Biomarkers & Prevention*. 2013;22(9):1600–1611. doi:10.1158/1055-9965.EPI-13-0350.

23. Breastfeeding and childhood cancer. 2001:1–10.

24. Urayama KY, Chokkalingam AP, Metayer C, et al. HLA-DP genetic variation, proxies for early life immune modulation and childhood acute lymphoblastic leukemia risk. *Blood*. 2012. doi:10.1182/blood-2012-01-404723.

25. Paulsson K, Johansson B. High hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2009;48(8):637–660. doi:10.1002/gcc.20671.

26. Ito C, Kumagai M, Manabe A, et al. Hyperdiploid acute lymphoblastic leukemia with 51 to 65 chromosomes: a distinct biological entity with a marked propensity to undergo apoptosis. *Blood*. 1999;93(1):315–320.

27. Zhang Y, Lu J, van den Berghe J, Lee SH. Increased incidence of spontaneous apoptosis

in the bone marrow of hyperdiploid childhood acute lymphoblastic leukemia. *Exp Hematol*. 2002;30(4):333–339.

28. Wiemels JL, Zhang Y, Chang J, et al. RAS mutation is associated with hyperdiploidy and parental characteristics in pediatric acute lymphoblastic leukemia. *Leukemia*. 2005;19(3):415–419. doi:10.1038/sj.leu.2403641.

29. Wiemels JL, Kang M, Chang JS, et al. Backtracking RAS mutations in high hyperdiploid childhood acute lymphoblastic leukemia. *Blood Cells, Molecules, and Diseases*. 2010;45(3):186–191. doi:10.1016/j.bcmd.2010.07.007.

Tables

Table 1A
Demographic and select characteristics of cases and control participating in CCLS (1995-2008) for Acute Lymphoblastic Leukemia (ALL)

	ALL		
	Cases n = 669 n (%)	**Controls n = 925 n (%)**	***p value cases vs. controls*[a]**
Child's sex			-[b]
Female	289 (43.2)	392 (42.4)	
Male	380 (56.8)	533 (57.6)	
Child age at diagnosis			-[b]
1 year	58 (8.67)	77 (8.32)	
2-5 years	384 (57.4)	543 (58.7)	
6-10 years	152 (22.7)	206 (22.3)	
11-14 years	72 (10.8)	95 (10.27)	
Missing	3 (0.5)	4 (0.4)	
Mean ± SD	5.6 ± 3.4	5.6 ± 3.4	
Child's ethnicity			-[b]
Hispanic	309 (46.2)	422 (45.6)	
Non-Hispanic white	243 (36.3)	352 (38.1)	
Non-Hispanic other	116 (17.3)	151 (16.3)	
Maternal race			-[b]
White/Caucasian	355 (53.1)	482 (52.1)	
African American	19 (2.8)	23 (2.5)	
Native American	8 (1.2)	5 (0.5)	
Asian/Pacific Islander	60 (8.9)	68 (7.4)	
Mixed/others	210 (32.7)	344 (37.2)	
Missing	8 (1.2)	3 (0.3)	
Mother's age at child's birth			0.123
<20 years	63 (9.4)	74 (8.0)	
20-24 years	164 (24.5)	178 (19.2)	
25-29 years	167 (24.9)	260 (28.1)	
30-34 years	169 (25.3)	249 (26.9)	
35-39 years	84 (12.6)	127 (13.7)	
≥40 years	19 (2.8)	36 (3.9)	
Missing	3 (0.5)	1 (0.1)	
Mean ± SD	27.7 ± 6.3	28.5 ± 6.1	
Mother's education			0.001
Less than high school	296 (44.3)	334 (36.1)	
High school/some college	190 (28.4)	290 (31.4)	
Bachelor's degree or higher	183 (27.4)	301 (32.5)	
Annual household income			<0.001
<$15,000	109 (16.3)	94 (10.2)	
$15,000-29,999	113 (16.9)	118 (12.8)	
$30,000-44,999	109 (16.3)	119 (12.9)	
$45,000-59,999	104 (15.6)	129 (13.9)	
$60,000-74,999	49 (7.3)	97 (10.5)	
≥$75,000	185 (27.6)	368 (39.8)	
Child's birth weight (grams)			0.046
<2500	39 (5.8)	50 (5.4)	
2500-3999	495 (74.0)	721 (77.9)	
≥4000	121 (18.1)	134 (14.5)	
Missing	14 (2.1)	20 (2.2)	
Mean ± SD	3448.9 ± 644.0	3419.6 ± 579.1	
Parental smoking (mother or father ever smoked)			0.008
No	273 (40.8)	440 (47.6)	
Yes	396 (59.2)	484 (52.3)	
Missing	0	1 (0.1)	

[a]p-values calculated using conditional logistic regression
[b]Matching variables; p-value not calculated for variables used in matching

Table 1B
Demographic and select characteristics of cases and control participating in CCLS (1995-2008) for three ALL subtypes

	c-ALL			ALL (TEL-AML1)			ALL (high-hyperdiploid)		
	Cases n = 312 n (%)	**Controls n = 443 n (%)**	***p value cases vs. controls*[a]**	**Cases n = 120 n (%)**	**Controls n = 169 n (%)**	***p value cases vs. controls*[a]**	**Cases n = 194 n (%)**	**Controls n = 271 n (%)**	***p value cases vs. controls*[a]**
Child's sex			-[b]			-[b]			-[b]
Female	138 (44.2)	191 (43.1)		49 (40.8)	65 (38.5)		94 (48.5)	125 (46.1)	
Male	174 (55.8)	252 (56.9)		71 (59.2)	104 (61.5)		109 (51.5)	146 (53.9)	
Child age at diagnosis			-[b]			-[b]			-[b]
1 year	-	-		9 (7.5)	12 (7.1)		13 (6.7)	18 (6.6)	
2-5 years	310 (99.4)	440 (99.6)		95 (79.2)	134 (79.3)		127 (65.5)	181 (66.8)	
6-10 years	-	-		12 (10.0)	17 (10.1)		40 (20.6)	56 (22.7)	
11-14 years	-	-		4 (3.3)	6 (3.6)		13 (6.7)	15 (5.5)	
Missing	2 (0.6)	3 (0.7)		0 (0)	0 (0)		1 (0.5)	1 (0.4)	
Mean ± SD	3.8 ± 1.1	3.8 ± 1.1		4.3 ± 2.3	4.4 ± 2.5		5.2 ± 3.0	5.1 ± 2.9	
Child's ethnicity			-[b]			-[b]			-[b]
Hispanic	150 (48.1)	205 (46.3)		39 (32.5)	51 (30.2)		93 (47.9)	133 (49.1)	
Non-Hispanic white	109 (34.9)	166 (37.5)		52 (43.3)	82 (48.5)		67 (34.5)	95 (35.1)	
Non-Hispanic other	52 (16.7)	72 (16.3)		29 (24.2)	36 (21.3)		34 (17.5)	43 (15.9)	
Maternal race			-[b]			-[b]			-[b]
White/Caucasian	164 (52.6)	227 (51.2)		69 (57.5)	100 (59.2)		100 (51.5)	132 (48.7)	
African American	6 (1.9)	7 (.6)		3 (2.5)	5 (2.9)		5 (2.6)	6 (2.2)	
Native American	6 (1.9)	2 (0.5)		1 (0.8)	0 (0)		4 (2.06)	4 (1.5)	
Asian/Pacific Islander	33 (10.6)	40 (9.0)		12 (10.0)	20 (11.8)		20 (10.3)	20 (7.4)	
Mixed/others	100 (32.1)	165 (37.3)		34 (28.3)	44 (26.0)		64 (33.0)	108 (39.8)	
Missing	3 (1.0)	2(0.5)		1 (0.8)	0 (0)		1 (0.4)	1 (0.75)	
Mother's age at child's birth			0.639			0.441			0.052
<20 years	28 (9.0)	30 (6.8)		11 (9.2)	10 (5.9)		15 (7.7)	17 (6.3)	
20-24 years	73 (23.4)	92 (20.8)		25 (20.8)	39 (23.1)		53 (27.3)	45 (16.6)	
25-29 years	81 (26.0)	132 (29.8)		36 (30.0)	39 (23.1)		46 (23.7)	79 (29.2)	
30-34 years	82 (26.3)	110 (24.8)		31 (25.8)	50 (29.6)		47 (24.2)	63 (23.3)	
35-39 years	42 (13.5)	61 (13.8)		15 (12.5)	21 (12.4)		29 (14.9)	56 (20.7)	
≥40 years	6 (1.9)	17 (3.8)		2 (1.7)	10 (5.9)		4 (2.1)	10 (3.7)	
Missing	0 (0)	1 (0.2)					1 (0)	1 (0.4)	
Mean ± SD	27.7 ± 6.2	28.5 ± 6.1		27.9 ± 5.9	28.9 ± 6.6		27.8 ± 6.3	29.3 ± 6.1	
Mother's education			0.234			0.781			0.217
Less than high school	135 (43.3)	160 (36.1)		43 (35.8)	52 (30.8)		87 (44.8)	104 (38.4)	
High school/some college	84 (26.9)	140 (31.6)		34 (28.3)	55 (32.5)		55 (28.4)	80 (29.5)	
Bachelor's degree or higher	93 (29.8)	143 (32.3)		43 (35.8)	62 (36.7)		52 (26.8)	87 (32.1)	
Annual household income			<0.001			0.002			0.151
<$15,000	55 (17.6)	43 (9.7)		17 (14.2)	18 (10.6)		27 (13.9.3)	34 (12.6)	
$15,000-29,999	50 (16.0)	55 (12.4)		25 (20.8)	16 (9.5)		30 (15.5)	35 (12.9)	
$30,000-44,999	53 (17.0)	56 (12.6)		12 (10.0)	25 (14.8)		36 (18.7)	36 (13.3)	
$45,000-59,999	49 (15.7)	65 (14.7)		21 (17.5)	16 (9.5)		31 (15.9)	38 (14.0)	
$60,000-74,999	18 (5.8)	40 (9.0)		3 (2.5)	14 (8.3)		16 (8.3)	23 (8.5)	
≥$75,000	87 (27.9)	184 (41.5)		42 (35.0)	80 (47.3)		54 (27.8)	105 (38.7)	
Child's birth weight (grams)			0.234			0.637			0.016
<2500	19 (6.1)	24 (5.4)		2 (1.7)	4 (2.4)		7 (3.6)	23 (8.1)	
2500-3999	234 (75.0)	349 (78.8)		96 (80.0)	139 (82.3)		144 (74.2)	207 (76.4)	
≥4000	56 (17.9)	60 (13.5)		21 (17.5)	23 (13.6)		42 (21.7)	32 (11.8)	
Missing	3 (1.0)	10 (2.3)		1 (0.8)	3 (1.8)		1 (0.5)	10 (3.7)	
Mean ± SD	3469 ± 644	3398 ± 602		3521 ± 557	3459 ± 501		3534 ± 613	3340 ± 628	
Parental smoking (mother or father ever smoked)			0.032			0.421			0.104
No	132 (42.3)	224 (50.6)		46 (38.3)	73 (43.2)		87 (44.9)	140 (51.7)	
Yes	180 (57.7)	219 (49.4)		74 (61.7)	96 (56.8)		107 (55.2)	130 (47.9)	

[a]p-values calculated using conditional logistic regression
[b]Matching variables; p-value not calculated for variables used in matching

Table 2A
Immune-modulating exposures and frequency of infections among cases and controls participating in CCLS (1995-2008) for ALL

	ALL		
	Cases n = 669 n (%)	**Controls n = 925 n (%)**	***p value cases vs. controls*[a]**
Day care by age 6 months (thousand child-hours)			0.106
<2000	629 (94.0)	844 (91.2)	
≥ 2000	39 (5.8)	78 (8.4)	
Missing	1 (0.2)	3 (0.3)	
Mean ± SD	324 ± 1282	469 ± 1477	
Birth order			0.406
Firstborn or only child	493 (73.7)	677 (73.2)	
Not-firstborn	164 (24.5)	234 (25.3)	
Missing	12 (1.8)	14 (1.5)	
Mean ± SD	0.97 ± 1.24	1.03 ± 1.16	
Birth type			0.047
Cesarean Section	150 (22.4)	191 (20.6)	
Vaginal	431 (64.4)	679 (73.4)	
Missing	88 (13.2)	55 (6.0)	
Total times ill in first 6 months			0.146
< 5 times	612 (91.5)	865 (93.5)	
≥ 5 times	57 (8.5)	59 (6.4)	
Mean ± SD	1.55 ± 4.15	1.26 ± 3.13	

[a]p-values calculated using conditional logistic regression

Table 2B
Immune-modulating exposures and frequency of infections among cases and controls participating in CCLS (1995-2008) for subtypes of ALL

	c-ALL			TEL-AML1			High hyperdiploid ALL		
	Cases n = 312 n (%)	**Controls n = 443 n (%)**	***p value cases vs. controls*[a]**	**Cases n = 120 n (%)**	**Controls n = 169 n (%)**	***p value cases vs. controls*[a]**	**Cases n = 194 n (%)**	**Controls n = 271 n (%)**	***p value cases vs. controls*[a]**
Day care by age 6 months (thousand child-hours)			0.350			0.123			0.200
<2000	291 (93.3)	400 (90.3)		115 (95.8)	154 (91.1)		186 (95.9)	248 (91.5)	
≥ 2000	21 (6.7)	40 (9.0)		5 (4.2)	15 (8.9)		8 (4.1)	21 (7.7)	
Missing	0 (0)	3 (0.7)		0 (0)	0 (0)		0 (0)	2 (0.7)	
Continuous (Mean ± SD)	340 ±1227	479 ± 1486		282 ± 889	608 ± 1836		190 ± 843	407 ± 1374	
Birth order			0.854			0.635			0.127
Firstborn or only child	230 (73.7)	329 (74.3)		87 (72.5)	120 (71.0)		148 (76.3)	189 (69.7)	
Not-firstborn	79 (25.3)	111 (25.1)		31 (25.8)	47 (27.8)		45 (23.2)	77 (28.4)	
Missing	3 (1.0)	3 (0.7)		2 (1.7)	2 (1.2)		1 (0.5)	5 (1.9)	
Mean ± SD	0.9 ± 1.1	1.0 ± 1.1		0.9 ± 1.1	1.0 ± 1.04		0.9 ± 1.1	1.1 ± 1.3	
Birth type			0.317			0.383			0.095
Cesarean Section	75 (24.0)	103 (23.3)		32 (26.7)	42 (24.8)		47 (24.2)	55 (20.3)	
Vaginal	209 (67.0)	334 (75.4)		76 (63.3)	123 (72.8)		122 (62.9)	199 (73.4)	
Missing	28 (9.0)	6 (1.4)		12 (10.0)	4 (2.4)		25 (12.9)	17 (6.3)	
Total times ill in first 6 months			0.175			0.572			0.097
< 5 times	286 (91.7)	415 (93.7)		110 (91.7)	158 (93.5)		175 (90.2)	252 (93.0)	
≥ 5 times	26 (8.3)	28 (6.3)		10 (8.3)	11 (6.5)		19 (9.8)	19 (7.0)	
Mean ± SD	1.6 ± 4.7	1.3 ± 3.1		1.7 ± 4.0	1.2 ± 3.1		1.5 ± 3.2	1.3 ± 2.9	

[a]p-values calculated using conditional logistic regression

Table 3A
Breastfeeding among cases and controls participating in CCLS (1995-2008) for ALL and subtypes of ALL

Duration of breastfeeding	ALL		c-ALL		TEL-AML1		High hyperdiploid ALL	
	Cases n = 669 n (%)	Controls n = 925 n (%)	Cases n = 312 n (%)	Controls n = 443 n (%)	Cases n = 120 n (%)	Controls n = 169 n (%)	Cases n = 194 n (%)	Controls n = 271 n (%)
Never	65 (9.7)	63 (6.8)	22 (7.1)	34 (7.7)	9 (7.5)	9 (5.3)	16 (8.3)	18 (6.6)
<6 months	176 (26.3)	188 (20.3)	95 (30.4)	107 (24.2)	32 (26.7)	39 (23.1)	52 (26.8)	44(16.2)
6-12 months	428 (64.0)	674 (72.9)	195 (62.5)	302 (68.2)	79 (65.8)	121 (71.6)	126 (66.0)	209 (77.1)
Mean ± SD (weeks)	28.4 ± 16.2	31.2 ± 14.5	27.9 ± 16.2	29.6 ± 15.2	28.6 ± 16.5	30.7 ± 14.9	28.4 ± 15.8	31.9 ± 13.9

Table 3B
Unadjusted models of breastfeeding patterns on risk of ALL among CCLS participants (1995-2008)

	ALL		c-ALL		ALL (TEL-AML1)		ALL (High Hyperdiploidy)	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Breastfeeding never/ever	0.63 (0.40, 0.99)	0.048	0.91 (0.49, 1.67)	0.754	0.86 (0.30, 2.49)	0.787	0.65 (0.30, 1.44)	0.291
Breastfeeding continuous	0.98 (0.97, 0.99)	0.001	0.99 (0.97, 1.00)	0.121	0.99 (0.96, 1.01)	0.347	0.98 (0.96, 1.00)	0.025
Duration of breastfeeding								
Never			1 (ref)		1 (ref)		1 (ref)	
<6 months	0.83 (0.52, 1.36)	0.477	1.13 (0.59, 2.18)	0.708	1.07 (0.33, 3.42)	0.91	0.98 (0.41, 2.32)	0.961
6-12 months	0.40 (0.23, 0.68)	0.001	0.64 (0.31, 1.29)	0.212	0.65 (0.19, 2.23)	0.497	0.37 (0.14, 0.96)	0.04
Overall test (p-value)	-	0.0004	-	0.1232	-	0.6311	-	0.0304

Table 4A
Breastfeeding and risk of ALL, adjusting for immune-related factors, demographics, and smoking

	ALL	
	OR (95% CI)	**P-value**
Duration of breastfeeding		
Never	1 (ref)	
<6 months	0.89 (0.49, 1.61)	0.695
6-12 months	0.49 (0.26, 0.95)	0.033
Overall test (p-value)	-	**0.0331**
Day care by age 6 months	0.62 (0.37, 1.04)	0.069
Birth order	0.80 (0.57, 1.11)	0.187
Birth type	0.74 (0.55, 0.99)	0.044
Total times ill in first 6 months	1.53 (0.94, 2.48)	0.086
Mother's age at child's birth		
<20 years	1 (ref)	
20-24 years	1.12 (0.68, 1.82)	0.664
25-29 years	0.98 (0.61, 1.60)	0.949
30-34 years	1.15 (0.67, 1.96)	0.609
35-39 years	1.27 (0.69, 2.31)	0.432
≥40 years	0.85 (0.38, 1.94)	0.713
Overall test (p-value)		0.8135
Mother's education		
Less than high school	1 (ref)	
High school/some college	0.80 (0.57, 1.12)	0.195
Bachelor's degree or higher	1.07 (0.69, 1.64)	0.752
Overall test (p-value		0.208
Annual household income		
<$15,000	1 (ref)	
$15,000-29,999	0.81 (0.51, 1.28)	0.368
$30,000-44,999	0.77 (0.48, 1.25)	0.3
$45,000-59,999	0.58 (0.35, 0.95)	0.032
$60,000-74,999	0.32 (0.18, 0.58)	<0.001
≥$75,000	0.28 (0.17, 0.47)	<0.001
Overall test (p-value		<0.001
Child's birth weight (grams)		
<2500	1 (ref)	
2500-3999	0.78 (0.46, 1.32)	0.351
≥4000	1.09 (0.60, 1.95)	0.784
Overall test (p-value		0.126
Parental smoking (mother or father ever smoked)	1.07 (0.83, 1.39)	0.584

Table 4B
Breastfeeding and risk of common ALL subtypes, adjusting for adjusting for immune-related factors, demographics, and smoking

	c-ALL		TEL-AML1		ALL (High Hyperdiploidy)	
	OR (95% CI)	**P-value**	**OR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**
Duration of breastfeeding						
Never	1 (ref)		1 (ref)		1 (ref)	
<6 months	1.41 (0.66, 3.01)	0.371	1.21 (0.26, 5.54)	0.806	0.97 (0.28, 3.31)	0.957
6-12 months	0.66 (0.29, 1.48)	0.316	0.77 (0.15, 4.09)	0.763	0.42 (0.12, 1.50)	0.180
Overall test (p-value)	-	0.077	-	0.823	-	0.177
Day care by age 6 months	0.65 (0.32, 1.34)	0.241	0.20 (0.04, 0.99)	0.049	0.36 (0.11, 1.22)	0.102
Birth order	0.98 (0.61, 1.59)	0.937	0.72 (0.31, 1.68)	0.444	0.50 (0.25, 1.00)	0.051
Birth type	0.79 (0.52, 1.19)	0.263	0.86 (0.44, 1.68)	0.667	0.51 (0.27, 0.96)	0.038
Total times ill in first 6 months	1.65 (0.82, 3.34)	0.163	1.29 (0.37, 4.49)	0.687	2.17 (0.88, 5.39)	0.093
Mother's age at child's birth						
<20 years	1 (ref)		1 (ref)		1 (ref)	
20-24 years	0.61 (0.29, 1.27)	0.188	0.34 (0.09, 1.28)	0.111	2.06 (0.69, 6.11)	0.191
25-29 years	0.75 (0.36 1.54)	0.441	0.59 (0.17, 2.03)	0.406	1.03 (0.36, 2.92)	0.957
30-34 years	0.80 (0.36, 1.79)	0.594	0.52 (0.13, 2.14)	0.368	1.92 (0.59, 6.29)	0.278
35-39 years	1.01 (0.43, 2.39)	0.972	0.92 (0.19, 4.29)	0.919	1.08 (0.32, 3.66)	0.897
≥40 years	0.42 (0.11, 1.56)	0.195	0.22 (0.02, 1.96)	0.176	1.11 (0.18, 6.84)	0.910
Overall test (p-value)		0.454		0.323		0.298
Mother's education						
Less than high school	1 (ref)		1 (ref)		1 (ref)	
High school/some college	0.98 (0.61, 1.58)	0.957	0.88 (0.35, 2.18)	0.782	0.88 (0.46, 1.68)	0.72
Bachelor's degree or higher	1.75 (0.93, 3.28)	0.079	1.84 (0.64, 5.24)	0.251	1.09 (0.46, 2.57)	0.839
Overall test (p-value		0.102		0.260		0.850
Annual household income						
<$15,000	1 (ref)		1 (ref)		1 (ref)	
$15,000-29,999	0.60 (0.30, 1.19)	0.145	1.31 (0.41, 4.12)	0.648	1.26 (0.46, 3.43)	0.645
$30,000-44,999	0.68 (0.33, 1.42)	0.309	0.52 (0.14, 1.98)	0.340	1.72 (0.63, 4.63)	0.285
$45,000-59,999	0.49 (0.24, 0.99)	0.05	0.90 (0.25, 3.17)	0.873	0.91 (0.34, 2.40)	0.855
$60,000-74,999	0.21 (0.08, 0.54)	0.001	0.21 (0.03, 1.25)	0.087	0.55 (0.17, 1.80)	0.329
≥$75,000	0.18 (0.08, 0.38)	<0.001	0.30 (0.08, 1.069)	0.087	0.46 (0.16, 1.30)	0.145
Overall test (p-value		<0.001		0.146		0.151
Child's birth weight (grams)						
<2500	1 (ref)		1 (ref)		1 (ref)	
2500-3999	0.71 (0.33, 1.54)	0.398	1.73 (0.23, 12.86)	0.592	2.27 (0.76, 6.79)	0.139
≥4000	1.11 (0.48, 2.57)	0.797	1.88 (0.22, 15.77)	0.560	3.36 (0.99, 11.33)	0.050
Overall test (p-value)		0.200		0.844		0.140
Parental smoking (mother or father ever smoked) (no vs. yes)	1. 26 (0.85, 1.85)	0.24	0.83 (0.39, 1.79)	0.647	1.19 (0.71, 2.01)	0.505

Table 5A
Breastfeeding (as binary variable) and risk of common ALL subtypes, adjusting for immune-related factors, demographics*, and smoking

	c-ALL		TEL-AML1		ALL (High Hyperdiploidy)	
	OR (95% CI)	**P-value**	**OR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**
Breastfeeding (<6 months vs. ≥ 6 months)	0.56 (0.31, 1.02)	0.059	0.71 (0.20, 2.52)	0.598	0.40 (0.17, 0.96)	0.040
Day care by age 6 months	0.69 (0.35, 1.39)	0.306	0.23 (0.0.5 1.03)	0.054	0.35 (0.11, 1.15)	0.083
Birth order	0.96 (0.62, 1.48)	0.868	0.69 (0.33, 1.44)	0.324	0.45 (0.24, 0.86)	0.015
Birth type	0.77 (0.51, 1.15)	0.205	0.81 (0.43, 1.50)	0.501	0.52 (0.28, 0.97)	0.039
Total times ill in first 6 months	1.77 (0.88, 3.57)	0.110	1.25 (0.38, 4.16)	0.370	2.19 (0.89, 5.37)	0.086

*mother's age at child's birth, maternal education, income, child's birthweight

Additional Tables

Breastfeeding and risk of ALL, adjusting for birth type and income* only

	ALL	
	OR (95% CI)	**P-value**
Duration of breastfeeding		
Never	1 (ref)	
<6 months	0.81 (0.46, 1.43)	0.469
6-12 months	0.51 (0.27, 0.93)	0.029
Overall test (p-value)	-	0.0492
Birth type	0.72 (0.55, 0.95)	0.020

*Adjusting for only significant variables in the model, birth type and income
Income overall test p <0.001

Breastfeeding (as binary variable) and risk of ALL adjusting for immune-related factors, demographics*, and smoking

	ALL	
	OR (95% CI)	**P-value**
Breastfeeding (<6 months vs. ≥ 6 months)	0.55 (0.34, 0.88)	0.012
Day care by age 6 months	0.62 (0.37, 1.03)	0.064
Birth order	0.81 (0.6, 1.09)	0.173
Birth type	0.73 (0.55, 0.98)	0.038
Total times ill in first 6 months	1.51 (0.93, 2.45)	0.093

*mother's age at child's birth, maternal education, income, child's birthweight

Breastfeeding among non-Hispanic white and Hispanic ALL cases and controls participating in CCLS 1995-2008, adjusting for immune-related factors, demographics*, and smoking

	Non-Hispanic White				Hispanic			
	Cases	**Controls**	**OR (95% CI)**	**P-value**	**Cases**	**Controls**	**OR (95% CI)**	**P-value**
	n = 243	n = 352			n = 309	n = 422		
	n (%)	n (%)			n (%)	n (%)		
Duration of breastfeeding								
Never	24 (9.9)	20 (5.7)	1 (ref)		32 (10.4)	33 (7.8)	1 (ref)	
<6 months	70 (28.8)	73 (20.7)	0.63 (0.21, 1.88)	0.415	76 (24.6)	87 (20.6)	0.63 (0.26, 1.57)	0.324
6-12 months	149 (61.3)	259 (73.6)	0.62 (0.20, 1.91)	0.407	201 (65.1)	302 (71.6)	0.34 (0.12, 0.96)	0.042
Overall test				0.6755				0.1056
Day care by age 6 months	222 (91.4)	305 (86.7)	0.41 (0.18, 0.95)	0.037	296 (95.8)	401 (95.0)	1.61 (0.58, 4.48)	0.355
	21 (8.6)	46 (13.1)			12 (3.9)	19 (4.5)		
	0 (0)	1 (0.2)			1 (0.3)	2 (0.5)		
Birth order	196 (80.7)	264 (75.0)	0.41 (0.22, 0.77)	0.006	205 (66.3)	295 (69.9)	1.19 (0.79, 1.80)	0.391
	45 (18.5)	82 (23.3)			100 (32.4)	122 (28.9)		
	2 (0.8)	6 (1.7)			4 (1.3)	5 (1.2)		
Birth type	56 (23.1)	74 (21.0)	0.51 (0.29, 0.89)	0.020	66 (21.4)	77 (18.3)	0.78 (0.49, 1.23)	0.294
	149 (61.3)	252 (71.6)			208 (67.3)	321 (76.1)		
	38 (15.6)	26 (7.4)			35 (11.3)	24 (5.7)		
Total times ill in first 6 months	226 (93.0)	328 (93.2)	1.13 (0.42, 3.08)	0.809	283 (91.6)	389 (92.2)	0.96 (0.49, 1.89)	0.923
	17 (7.0)	24 (6.8)			26 (8.4)	32 (7.6)		
					0 (0)	1 (0.3)		

*mother's age at child's birth, maternal education, income, child's birthweight

Side note:
Binary breastfeeding (less 6 months (0) vs. >6 months breastfeeding (1), controlling for immune-modulating factors and demographics:
Hispanics: OR 0.49; 95% CI: 0.23-1.02; P-value 0.058
Non-Hispanic Whites: OR: 0.87; 95% CI: 0.40 -1.88; P-value 0.725