Evaluation of the Role of Cytomegalovirus Infection in the Development of Acute Lymphoblastic Leukemia and Other Cancers

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Jennifer M. Geris

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Logan G. Spector

May 2022

Acknowledgements

First and foremost, I would like to acknowledge my advisor, Dr. Logan Spector. He has been a truly remarkable mentor and a consistent source of support throughout my doctoral experience. I am grateful for his time, effort, and resources he devoted to my training and career development. I have learned and grown significantly as a scientist while under his mentorship, and I cannot fully express how grateful I am to have been one of his students.

I would also like to acknowledge the other members of my dissertation committee, Drs. Heather Nelson, Shalini Kulasingam, Saonli Basu, and Mark Schleiss. Each has offered considerable time and feedback throughout this process, and through each I have gained a new perspective of epidemiology, cancer, biostatistics, and virology. Thank you for your support and expertise.

There are many other talented scientists outside of my committee I would also like to acknowledge. Without the efforts of AJ Hooten, Erica Langer, and Nelmary Hernandez-Alvarado the SMILES study would not have been possible. Thank you to the other members of the Schleiss lab, Claudia and Mark, you have been enormously kind and helpful. There are also many others in Dr. Spector's group that provided guidance, support, and friendship throughout the years and I feel fortunate to have had the opportunity to work with them. I'd also like to acknowledge Dr. Eric Engels and all the members of the TCM working group. Thank you for your guidance and access to this incredible dataset. Thank you to the Institute for Molecular Virology community, led by Dr. Louis Mansky, for their support and career development opportunities. The last few years have been weird as a virologist, but I am thankful to have had this community to

learn from. Lastly, thank you to Dr. Hank Balfour for your continued mentorship and friendship throughout my career.

I would like to acknowledge the financial support I received from the Institute for Molecular Virology Training Program supported by the National Institutes of Health (T32 AI083196) and the J.B. Hawley Student Research Award.

Finally, I am so grateful for my family and friends who have been with me on this long journey. Thank you to my parents, who have always supported my dreams and for teaching me the value of an education. For my parents-in-law, thank you for always being there and your endless support. Thank you, Dr. Emma Leirdahl for being my forever coffee and homework buddy. I am most grateful for the love and support of my wonderful husband, Ryan, who has been my biggest source of encouragement for the past decade, and my daughter Eleanor, who was there kicking me during my Part A exam to now, when she tells me to get off my computer and play with her. I promise to spend more time with you and less time on my dissertation very, very soon.

Dedication

For Eleanor, my ballerina/princess/astronaut/scientist/superhero and greatest joy.

Abstract

Acute lymphoblastic leukemia (ALL) is the most prevalent pediatric malignancy, and a leading cause of death in children. Understanding the risk factors of pediatric ALL is necessary to enable early detection and prevention. Congenital cytomegalovirus (cCMV) has recently been identified as a possible risk factor of ALL. In *Manuscript 1*, we compared the prevalence of cCMV infection in newborn dried blood spots of ALL cases and cancer-free controls. There was no difference in the odds of cCMV infection comparing ALL cases to controls in our primary analysis. However, cCMV was significantly more prevalent among hyperdiploid ALL cases compared to unmatched controls. These findings offer partial support for the association of cCMV with ALL.

CMV is among the most common viral infections following solid organ transplantation (SOT). CMV disease post-SOT has been associated with an increased risk of subsequent non-Hodgkin lymphoma but has not been well-studied for other hematologic malignancies. In *Manuscript 2*, we aimed to describe CMV infection status pre-transplant as it relates to the incidence of leukemias, lymphomas, and myeloma. We identified that CMV recipient and donor sero-mismatch (R-/D+) was associated with significantly lower risk of diffuse large B-cell lymphoma compared to CMV seronegative R/D pairs, indicating CMV may have a protective role in carcinogenesis.

Lastly, in *Manuscript 3*, we leveraged the same dataset as *Manuscript 2* to examine associations of CMV with solid tumor cancer risk among SOT recipients. Using linked data from the United States SOT registry and 32 cancer registries, we report an inverse association between R-/D+ CMV serostatus and small intestine cancer, and a positive

association between CMV R+ serostatus and lung cancer. CMV status was not associated with risk for other cancers.

Findings from this dissertation may motivate and inform future work to further understand the relationship between this highly prevalent virus and cancer.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Epidemiology of CMV Infection

Cytomegalovirus (CMV) is a member of the herpesvirus family (HHV 5), of which there are 7 other human pathogens that cause disease in humans, including Herpes Simplex viruses types 1 and 2, Epstein-Barr virus, Varicella-Zoster virus, and Human Herpes Viruses 6, 7, and 8. Like all herpesviruses, CMV is able to establish lifelong latency in the host and can periodically undergo reactivation and is then transmissible to susceptible individuals via exchange of bodily fluids.[1–4] In immunocompetent adults, primary CMV infection is typically asymptomatic. However, infection may manifest as CMV mononucleosis. Hallmark symptoms of CMV mononucleosis include fever, fatigue, lymphocytosis, and mild elevation of liver enzymes.

Mode of Acquisition

There are several modes by which CMV can be acquired, however, all require close or intimate contact with persons who are excreting CMV in their urine, saliva, semen, tears, blood, or other bodily fluids. Maternal transmission to the fetus through either a new or reactivated latent infection may occur at any point during gestation, leading to congenital CMV (cCMV) infection.[5] In infancy and toddlerhood, CMV may be transmitted by breastmilk, saliva, or urine.[6–10] Attendance at group daycare centers has been attributed to the spread of CMV infection in young children.[11–14] During adolescence and adulthood, transmission is typically through oral or sexual contact.[15] However, individuals who work with young children in childcare or hospital settings are

also at risk for acquiring CMV through contact with infected urine and saliva.[11,16] Finally, CMV may be acquired through either blood transfusion or organ transplantation. Studies have demonstrated that transfusion-associated CMV infections are associated with receipt of blood from CMV seropositive donors.[17] Likewise, in bone marrow and solid organ transplantation, recipients are at a significant risk for CMV infection through infected tissue from a seropositive donor. In the case of a CMV seropositive donor and seronegative recipient, primary infection develops in 65 – 75% of recipients.[5] Additionally, there is also the risk of reactivated CMV infection due to sustained immunosuppressive therapies. As a result, CMV infection in immunodeficient individuals poses a significant risk of morbidity and mortality.

Risk Factors

The prevalence of CMV varies by demographic characteristics, including age, geographic location, race/ethnicity, sex, and socioeconomic status (SES). In the first two years of life, the prevalence of CMV is roughly 15-20%[18,19] and is often reflective of *in utero* or perinatal infection. CMV seroprevalence then gradually increases with age, from about 36% in 6-11 year old, to nearly 90% in individuals 80 years or greater.[20] An overall global seroprevalence of CMV of approximately 80% has been reported, with higher rates of seropositivity in the Eastern Mediterranean region and lower rates (~65%) in Europe.[21] In Africa and Asia, the seroprevalence is nearly 100%.[22]

Racial and ethnic differences in CMV infection status have also been described in the US. CMV seroprevalence has been shown to be almost 60% higher among non-white groups compared to whites (prevalence ratio (PR): 1.59, 95% CI: 1.57-1.61). Data from

the National Health and Nutrition Examination Survey (NHANES) 1999-2004 revealed that age-adjusted CMV seroprevalence among non-Hispanic white individuals was 39.5% compared to 70.6% in non-Hispanic black persons, and 76.9% in Mexican Americans.[23] Seroprevalence was also higher in females in all racial and ethnic groups, to which the overall seroprevalence was 55.5% compared to 45.2% among males.[23] The trend of seroprevalence by sex and race/ethnicity across age is shown in **The role** of CMV as an etiologic agent for leukemia has only recently emerged and remains largely unexamined. Given the oncogenic potential of CMV, it is warranted to evaluate the association between CMV infection and malignancy. In this dissertation, we aimed to contribute to this knowledge base by first conducting a large, population-based study of ALL cases and cancer-free controls in order to replicate the findings by Francis et al. to describe the association between congenital CMV infection and leukemia. Secondly, we sought to assess the role of CMV infection and hematologic and solid tumor cancers in the setting of solid organ transplantation, as there is a high risk for CMV infection following transplantation. Overall, our findings are poised to elucidate CMV as a modifiable risk factor in the pathogenesis of ALL. Thus, these data could provide rationale for a CMV vaccine for disease reduction in congenital and post-transplant CMV infection and indirectly the burden of ALL.

In Manuscript 1, we aimed to compare CMV prevalence at birth in newborn dried bloodspots of 1,189 ALL cases to that in 4,756 cancer-free controls. Controls were matched 4:1 on year of birth, sex, and mother's reported race/ethnicity. Detection of CMV DNA was performed using quantitative PCR. Odds ratios and 95% confidence intervals were estimated from conditional logistic regression. We conducted a stratified analysis by cases with available immunophenotype and cytogenetic subtype data. We also stratified by demographic and birth characteristics.

Manuscripts 2 and 3 utilized data from the Transplant Cancer Match Study (TCM) by the National Cancer Institute. In Manuscript 2, we aimed to characterize the association between CMV infection and the incidence of leukemia and other hematologic malignancies among solid organ transplant recipients. The TCM uses electronically linked transplant data from the Scientific Registry of Transplant Recipients and 32 cancer

registries with coverage between 1987 -2017. Transplant recipients previously characterized through linkage of the SRTR with known CMV serostatus were included. CMV serostatus was categorized into three risk groups according to SRTR data on pretransplant IgG serostatus of donors and recipients, to reflect risk of active CMV infection and disease post-transplant: high risk (recipient seronegative and donor seropositive [R-/D+]), intermediate risk (recipient seropositive regardless of donor serostatus [R+]), and low risk (recipient and donor seronegative [R-/D-]). Incidence rate ratios and 95% confidence intervals comparing the risk of cancer among CMV exposed individuals to those unexposed were estimated by Poisson regression adjusted for sex, age at transplantation, race/ethnicity, SES quintile, transplanted organ (kidney, liver, or other/multiple), and EBV risk group.

In Manuscript 3, we utilized the same database to ascertain the association between CMV infection and the incidence of solid tumor malignancies among solid organ transplant recipients. We calculated incidence rates for each cancer for recipients in each CMV risk group. To compare cancer risk by CMV risk group, we estimated incidence rate ratios using multivariable Poisson regression models adjusted for sex, age at transplantation, race/ethnicity, SES quintile, transplanted organ (kidney, liver, or other/multiple), and EBV risk group.

Figure 1- *1*.

Other sources for the disparities in CMV seroprevalence are factors associated with SES. Household income has been directly associated with CMV seroprevalence, with lower income having the highest prevalence and the lowest rates among the highest household incomes.[23] An inverse relationship between a higher level of education and CMV seroprevalence has also been reported. Household crowding is also significantly associated with CMV seropositivity, with seroprevalence increasing with the level of crowding.[20,23,24]

Congenital and Perinatal Infection

CMV is the most prevalent congenital infection in the United States occurring in 0.5%-2% of all live births depending on socioeconomic status and race/ethnicity (**Figure 1- 2**).[25,26] Currently it is estimated that 40,000 infants are born with cCMV annually in the US[27,28] and between 15-20% of these infants will have a permanent developmental deficit.[29,30] The impact of cCMV as a cause of disability among newborns is greater than that of fetal alcohol syndrome, Down syndrome, spina bifida, or pediatric HIV/AIDS.[31–33]

Transmission of CMV during pregnancy predominantly occurs either through reactivation of latent infection or primary infection. Risk of congenital infection is higher for seronegative women who have a primary CMV infection during pregnancy than for seropositive women who experience a reactivation or reinfection.[20,34,35] Each year in the US, it is estimated that 2.3% of all pregnant women acquire a primary CMV infection and among these women, 40-60% of their infants will be congenitally infected.[34,36,37] In contrast, women who are seropositive prior to pregnancy transmit CMV to the fetus in only 1-2% of pregnancies.[35]

Among infants with cCMV infection, all are at risk for sequelae, particularly sensorineural hearing loss (SNHL).[38] Those with symptomatic infection at birth, such as thrombocytopenia, petechiae, jaundice, and/or liver disease, have a higher risk of severe long-term sequelae, including mental retardation, microcephaly, developmental delay, and SNHL.[38] The risk of sequelae in infants with asymptomatic cCMV is believed to be limited to SNHL, with a risk of approximately 15%.[38] However, it is becoming increasingly clear that this likely a significant underestimate of the true prevalence of sequelae in infants asymptomatic at birth, and that long-term disabilities may include subtle neurodevelopmental deficits that previously have not been identified as CMV-associated due to the lack of universal cCMV screening at birth. Currently, it is not well understood why some infants with cCMV develop symptomatic infection and subsequent sequelae at greater rates than other children.

Solid Organ Transplantation

CMV is among the most common viral infections following solid organ transplantation (SOT) and is associated with severe symptomatic disease and allograft rejection, leading to rapid organ failure and death.[39–41] The occurrence of disease by CMV infection in SOT recipients varies according to pre-transplant donor/recipient CMV serostatus, intensity of immunosuppression, and the organ transplanted. [42] The greatest risk factor for CMV disease post-SOT occurs when the recipient develops primary CMV infection under sustained immunosuppressive regimens for transplantation.[43,44] Primary infection can occur if the recipient is CMV seronegative at the time of transplant, either following transmission in the community or especially when transmitted from a transplant donor who is CMV seropositive.

CMV disease after SOT can manifest as fever, leukopenia, thrombocytopenia, or systemic syndrome, affecting many organs.[39–41] CMV may also contribute to cancer in SOT recipients. It has been reported that CMV seronegative transplant recipients who receive a CMV positive organ have elevated risk of post-transplant lymphoproliferative disorder (PTLD, a spectrum of conditions that includes lymphoma).[45,46] In a study of liver recipients who seroconverted to Epstein-Barr virus (EBV), CMV disease was reported in 54% of patients who developed PTLD but in only 18% of patients who did not develop PTLD.[47] In addition, hospitalization for CMV disease during the first year post-transplant has been associated with subsequent risk of non-Hodgkin lymphoma (NHL).[48] The relationship between CMV infection and other cancers after SOT is not well-defined.

Pathogenesis

CMV viral genotypes

Variation of CMV viral genes have been proposed as a possible mechanism of cCMV disease severity due to the genetic variability of CMV strains.[49–51] The double-stranded DNA (dsDNA) genome of CMV consists of unique long (UL) and short (US) segments, each flanked by inverted repeat sequences (IRL and IRS)().[52] Within these regions, CMV encodes for proteins responsible for various functions, including immunomodulation and virulence. Several genes have been identified that demonstrate hypervariability in clinical isolates among congenitally infection children, including *UL144, UL146, UL55,* and *UL9*.

UL144 is a truncated tumor necrosis factor-∝-like receptor gene responsible for inhibition of T cell proliferation and induction of chemokines via the NF-kB signaling pathway. [53,54] Three genotypes in *UL144* polymorphisms have been identified through sequencing studies – A, B, and C – of which genotype B is the most prevalent.[53,55–57] Genotype C has been associated with severe cCMV disease in newborns compared to those with asymptomatic infection.[53,58] These differences may relate to differences in NF-kB signaling, attributable to the highly divergent coding sequence (particularly in the COOH-terminus of the proteins), although this has not yet been experimentally verified. Another gene of interest is *UL146*, an α-chemokine gene involved in degranulation of neutrophils, calcium mobilization, and chemotaxis,[59] and also thought to play a role in promoting angiogenesis.[60,61] Like *UL144*, the viral sequence of *UL146* is highly variable and been classified into 14 distinct genotypes to date, of which 2 (G1 and G13) have been associated with elevated CMV IgG and IgM antibodies among congenitally infected

infants.[62] Glycoprotein B (gB), encoded by *UL55*, is a well-recognized as a leading target for subunit vaccine development,[63] but the immunoregulatory aspects of this protein are under-studied. The gB protein mediates viral attachment and cell entry, which activates signaling pathways such as mitogen-activated protein kinase (MAPK)[64] and focal adhesion kinase (FAK)[65], which intriguingly have been shown to be associated with oncogenesis.[66,67] Lastly, variation in *UL9* has been identified in the clinical strain DB-CMV,[68] which is involved in the upregulation of proto-oncogene *Bcl-3*.[69] Therefore, further investigation of the viral sequences of these genes is warranted, especially with their potential role in oncogenesis.

CMV and cancer

CMV is not presently regarded as an oncogenic virus; however, CMV infection has been implicated in malignancy. *In vitro* studies have demonstrated the oncogenic potential of CMV by its ability to transform a variety of mammalian cell lines. In 1973, Albrecht and Rapp revealed that UV-inactivated CMV is able to transform hamster embryo fibroblasts; the resultant tumors were poorly differentiated fibrosarcomas and CMV antigens were detected in the cytoplasm and on the cell surface.[70] Later, this was extended to include transformation of normal human embryonal cells which exhibited enhanced tumorigenicity in nude mice.[71] Recently, it was also demonstrated that a clinical strain of CMV, CMV-DB, can transform primary human mammary epithelial cells *in vitro* and the transformed cells gave rise to fast-growing triple negative breast tumors when injected in immunodeficient mice.[72]

In a clinical setting, CMV DNA, mRNA, and/or antigens have been detected in tumor tissues, suggesting a role for CMV in the etiology in cancer pathogenesis. Specifically, CMV nucleic acids and proteins have been detected in tumor samples of breast, colon, ovarian, and prostate cancers as well as glioblastoma and mesothelioma. [73–76] A serology study of neuroblastoma and Wilm's tumor have also found a greater proportion of neuroblastoma and Wilm's tumor patients (40% and 44%, respectively) had complement-fixing antibodies against CMV compared to cancer-free controls who were all antibody negative. Further, these differences were more pronounced in younger children (ages 0-5 years), of which 53% of neuroblastoma and Wilm's tumor patients were positive.[77] Most recently, congenital CMV infection has been implicated in the development of pediatric acute lymphoblastic leukemia.[78–80]

Epidemiology and Pathophysiology of Acute Lymphoblastic Leukemia

Acute lymphoblastic leukemia is a neoplastic disease of lymphoid progenitor cells in the bone marrow, blood, and extramedullary sites. ALL is the most prevalent pediatric malignancy, accounting for nearly 20% of all childhood cancers[81,82] with an annual incidence of 34 cases per million persons <20 years of age.[82,83] The peak incidence occurs between ages 1 and 4 years, exceeding 100 cases per million in boys and girls (**Figure 1- 4**).

While survivorship has improved greatly due to advances in therapy and cancer screening,[84,85] ALL survivors experience greater morbidities and mortality than the general population.[86,87] Among the long-term adverse effects include secondary neoplasms, obesity, chronic medical conditions, cognitive deficits, congestive heart

failure, and early mortaility.[88] Elucidating risk factors and identifying strategies for prevention is important for improving both pediatric and adult health.

The pathogenesis of ALL involves abnormal proliferation and differentiation of a clonal population of lymphoid cells. Gross chromosomal alterations are a hallmark of ALL, and chromosomal rearrangements creating chimeric fusion genes commonly involve epigenetic modifiers, cytokine receptors, tyrosine kinases, and hematopoietic transcription factors.[89,90] High hyperdiploidy and the translocation (12;21) (encoding *ETV6-RUNX1*) translocation are each present in nearly a third of childhood ALL cases.[91] In pediatric populations, predisposing genetic syndromes, such as Down syndrome or Fanconi anemia, have been identified in a minority of cases.[92–94] Strong environmental exposures, such as ionizing radiation during the prenatal period, have been accepted causes of ALL but are also rare.[95,96] Other modifiable risk factors, including low-frequency magnetic fields, maternal alcohol use, pesticides, and paternal smoking have some evidence of association but remain inconclusive, and there are concerns that the findings of these studies may be due to bias.[97]

Infectious Etiology of ALL

The role of infection in the etiology of leukemia has long been suspected. Three hypotheses have been proposed that describe the nature of this pathogenesis: Greave's 'delayed infection' hypothesis, Kinlen's 'population-mixing' hypothesis, and Smith's hypothesis of 'direct transformation' by an infectious agent. [98–100]

First formulated in 1988, Mel Greave's hypothesized that ALL is triggered by an abnormal immune response to one or more common childhood infections.[101] This abnormality, he postulates, arises due to 1) infectious exposures being delayed beyond

the immunologically anticipated period of infancy; and 2) some degree of inherited genetic susceptibility. Greave's later expanded this hypothesis to include the 'two hit' model, stating that a minimum of two etiologic events are required for the development of ALL, the first being genetic aberration *in utero* and the second genetic event caused indirectly by an aberrant immune response to infection.[98]

Also in 1988, Leo Kinlen postulated that leukemia arises from 'population-mixing,' a term used to describe the interaction between nonimmune and infective children, would lead to greater penetrance of the pathogen in a population. [99] The hypothesis proposes that the immune systems of children who reside in less densely populated communities are less likely to have been exposed to a diverse range of infectious agents as compared to residents of more populated communities. These children are therefore believed to be more likely to develop leukemia once exposed to novel infections from incoming migrants.

Lastly, a third hypothesis by Malcom Smith and colleagues in 1997 posits a leukemogenic virus. [100] In his hypothesis, Smith proposes ALL originates from exposure to infection *in utero*, which could explain the peak in incidence during childhood. Further, an oncogenic virus that is transmitted *in utero* or during the first year of life would be able to infect immature lymphocytes and promote leukemia through mechanisms of direct transformation. However, no specific infection has been causally linked to ALL.

CMV-related ALL

Congenital CMV has recently emerged as a moderate-to-strong risk factor of acute lymphoblastic leukemia. In a report by Francis et al., cCMV infection was assessed by droplet digital polymerase chain reaction (ddPCR) in a population-based sample of newborn dried blood spots (DBS) randomly drawn from the California Childhood Cancer Record Linkage Program of 268 ALL cases and 270 cancer-free controls. In total, cCMV was detected in 26/268 (9.7%) of ALL cases but only 8/270 (3.0%) of controls for a highly significant odds ratio (OR) of 3.71 (95% confidence interval (CI): 1.71 – 8.95). In follow-up to these findings, a second study by Wiemels et al. assessed congenital and early-life, clinically recognized CMV infection and the risk of hematologic malignancy in population-based registries of Sweden. The authors found in the nearly 10 thousand exposed person-years, the hazard ratio (HR) of hematologic malignancy among children exposed to any medically documented cCMV or early-life acquired CMV was 11.2 (95% CI: 5.8 – 21.5).

Despite this small literature base, both studies of ALL and cCMV to date have had robust study designs. The first study used nested samples in a population-based study to detect CMV using state-of-the-art methods. The second study used prospectively collected medical record data from the entire population of Sweden. Moreover, both studies found strong relative risks that are less possible to attribute to unexamined confounding variables. Therefore, it is warranted to further examine this association to gain a better understanding of the etiology and pathophysiology connecting cCMV with ALL.

Dissertation Objectives

The role of CMV as an etiologic agent for leukemia has only recently emerged and remains largely unexamined. Given the oncogenic potential of CMV, it is warranted to evaluate the association between CMV infection and malignancy. In this dissertation, we aimed to contribute to this knowledge base by first conducting a large, population-based study of ALL cases and cancer-free controls in order to replicate the findings by Francis *et al.* to describe the association between congenital CMV infection and leukemia. Secondly, we sought to assess the role of CMV infection and hematologic and solid tumor cancers in the setting of solid organ transplantation, as there is a high risk for CMV infection following transplantation. Overall, our findings are poised to elucidate CMV as a modifiable risk factor in the pathogenesis of ALL. Thus, these data could provide rationale for a CMV vaccine for disease reduction in congenital and post-transplant CMV infection and indirectly the burden of ALL.

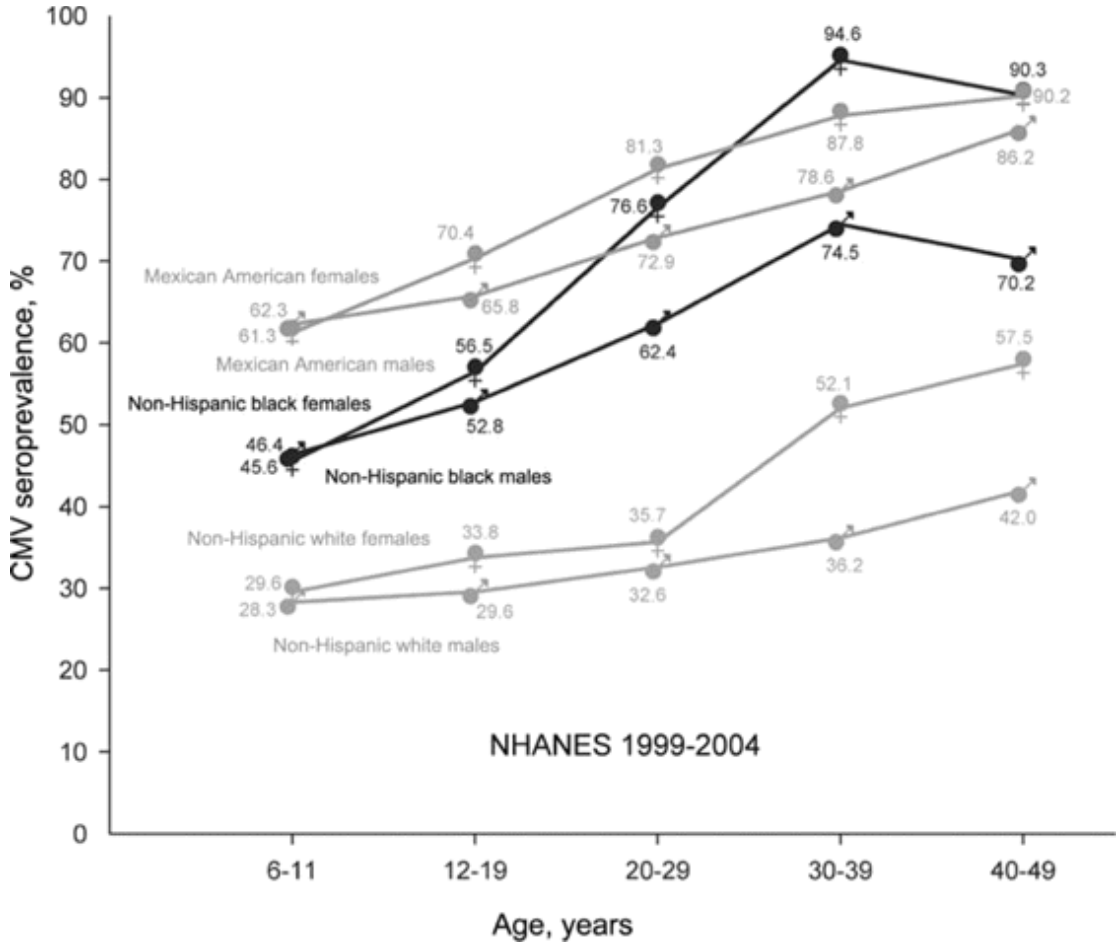
In Manuscript 1, we aimed to compare CMV prevalence at birth in newborn dried bloodspots of 1,189 ALL cases to that in 4,756 cancer-free controls. Controls were matched 4:1 on year of birth, sex, and mother's reported race/ethnicity. Detection of CMV DNA was performed using quantitative PCR. Odds ratios and 95% confidence intervals were estimated from conditional logistic regression. We conducted a stratified analysis by cases with available immunophenotype and cytogenetic subtype data. We also stratified by demographic and birth characteristics.

Manuscripts 2 and 3 utilized data from the Transplant Cancer Match Study (TCM) by the National Cancer Institute. In Manuscript 2, we aimed to characterize the association between CMV infection and the incidence of leukemia and other hematologic malignancies among solid organ transplant recipients. The TCM uses electronically

linked transplant data from the Scientific Registry of Transplant Recipients and 32 cancer registries with coverage between 1987 -2017. Transplant recipients previously characterized through linkage of the SRTR with known CMV serostatus were included. CMV serostatus was categorized into three risk groups according to SRTR data on pretransplant IgG serostatus of donors and recipients, to reflect risk of active CMV infection and disease post-transplant: high risk (recipient seronegative and donor seropositive [R-/D+]), intermediate risk (recipient seropositive regardless of donor serostatus [R+]), and low risk (recipient and donor seronegative [R-/D-]). Incidence rate ratios and 95% confidence intervals comparing the risk of cancer among CMV exposed individuals to those unexposed were estimated by Poisson regression adjusted for sex, age at transplantation, race/ethnicity, SES quintile, transplanted organ (kidney, liver, or other/multiple), and EBV risk group.

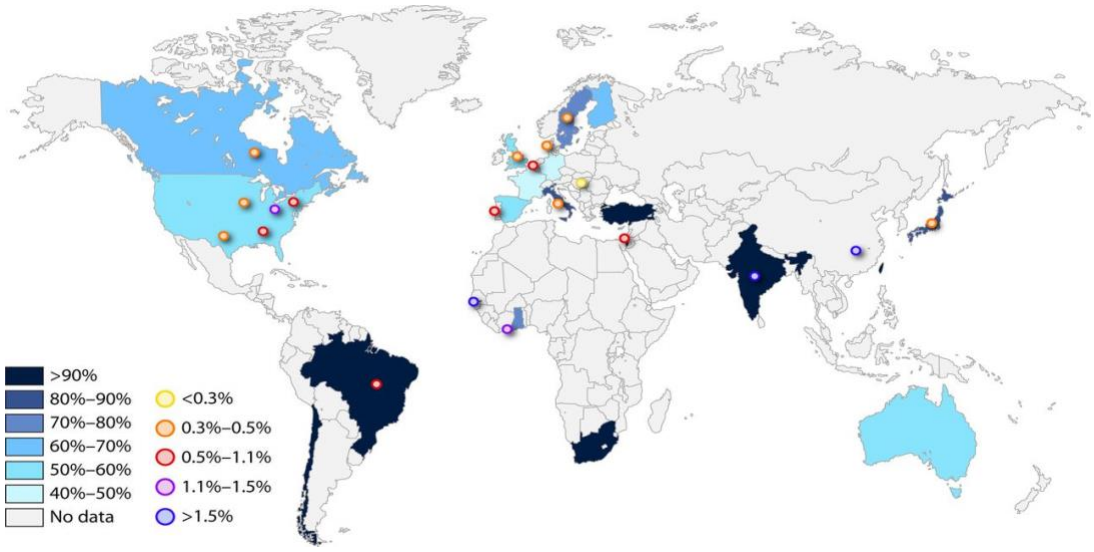
In Manuscript 3, we utilized the same database to ascertain the association between CMV infection and the incidence of solid tumor malignancies among solid organ transplant recipients. We calculated incidence rates for each cancer for recipients in each CMV risk group. To compare cancer risk by CMV risk group, we estimated incidence rate ratios using multivariable Poisson regression models adjusted for sex, age at transplantation, race/ethnicity, SES quintile, transplanted organ (kidney, liver, or other/multiple), and EBV risk group.

Figure 1- 1. CMV seroprevalence in the United States, National Health and Nutrition Examination Survey (NHANES) 1999-2004, stratified by age, sex, and race/ethnicity.



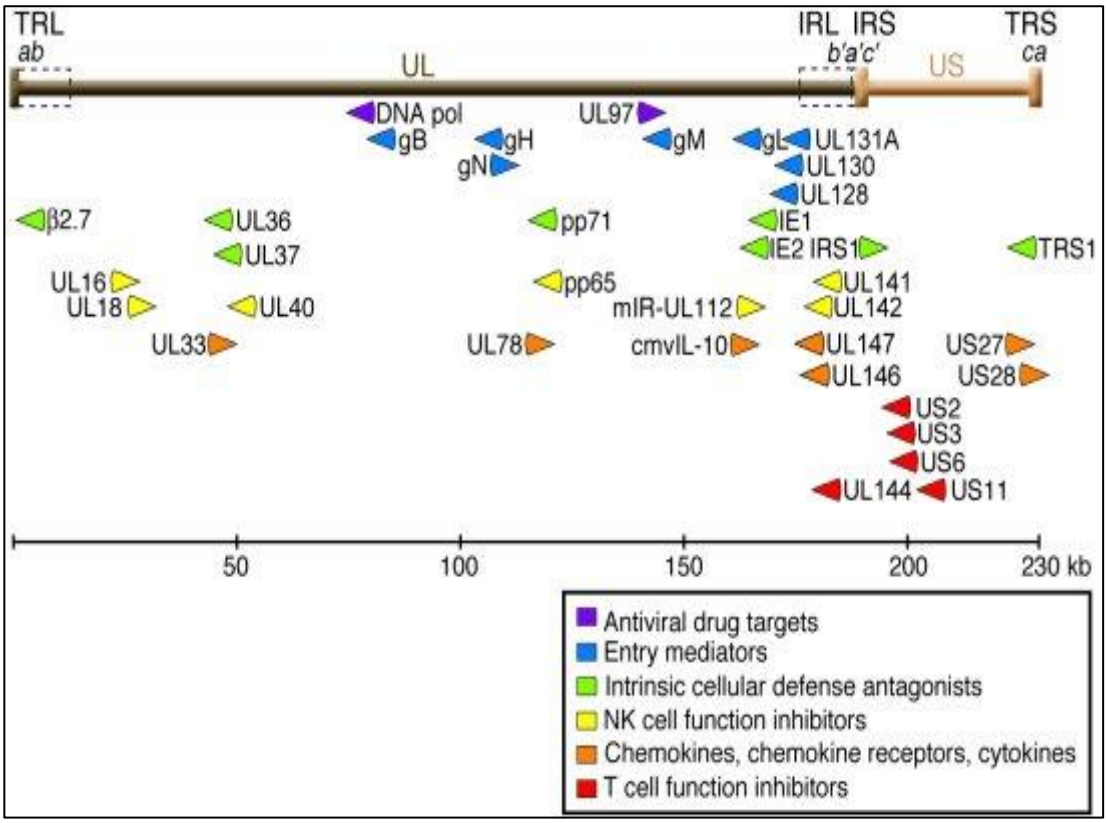
Source: *Clin Infect Dis*, Volume 50, Issue 11, 1 June 2010, Pages 1439–1447, https://doi.org/10.1086/652438 [23]

Figure 1- 2. Worldwide CMV seroprevalence rates among women of childbearing age and birth prevalence of congenital CMV infection.



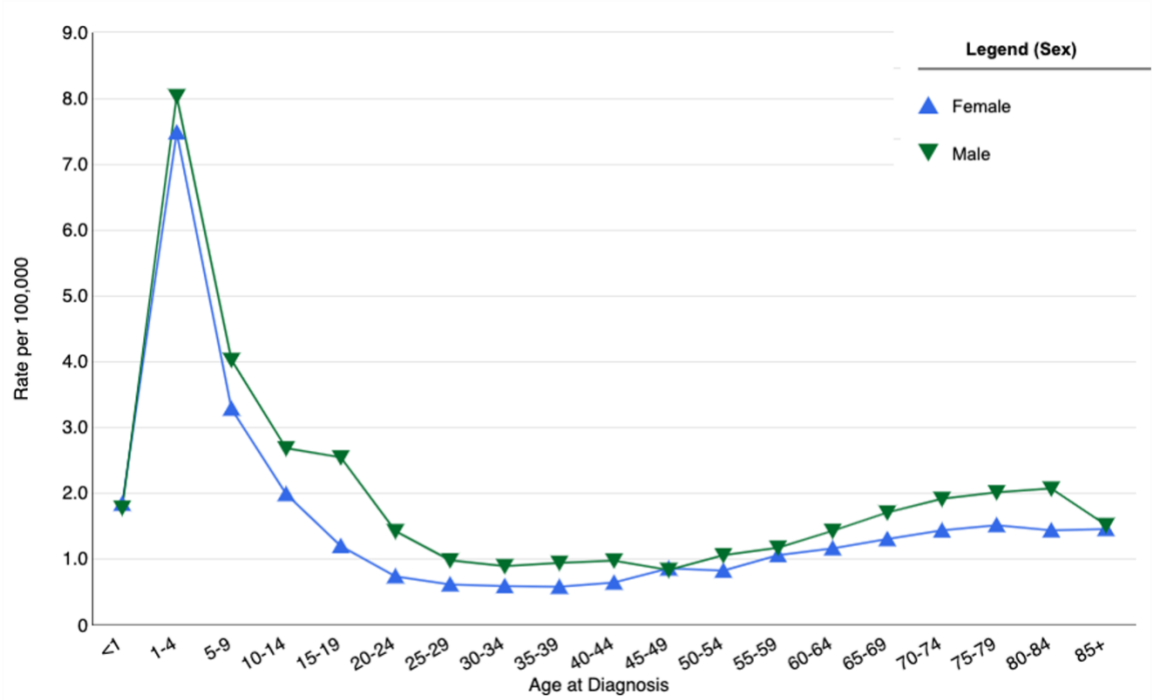
Source: Manicklal et al. Clin Microbiol Rev (2013)[25]

Figure 1- 3. Genome of CMV



Source: *J Clin Invest.* 2011;121(5):1673-1680. https://doi.org/10.1172/JCI45449.[52]

Figure 1- 4. Age-Specific Incidence Rates of ALL by Sex, SEER 2014-2018.



Data source: Surveillance, Epidemiology, and End Results (SEER) program, SEER 21 areas, 2014 – 2018.

Chapter 2: Limited evidence for an association between congenital cytomegalovirus infection and pediatric acute lymphoblastic leukemia in a large, population-based study

---

Introduction

Acute lymphoblastic leukemia (ALL) is the most common form of pediatric cancer, accounting for nearly 20% of all malignancies diagnosed in persons under 20 years of age.[81,82] Highly penetrant genetic predisposition, such as Down syndrome, cause ALL in less than 5% of cases.[102] Environmental risk factors have also been suggested, such as exposure to ionizing radiation, but the evidence is not conclusive.[95,96] Largely, the pathogenesis of most ALL cases remains unknown.

An infectious etiology for ALL has been hypothesized. Mel Greaves[101] was the first to propose the "delayed infection" hypothesis, suggesting that ALL may arise as a consequence of an abnormal immune response to common childhood infections. More recently, this hypothesis has been expanded to include the "two-hit" model, stating a minimum of two etiologic events are required for the development of ALL, the first event being genetic aberration *in utero* and the second genetic event caused indirectly by an atypical immune response to infection.[98] A second hypothesis proposed by Kinlen[99] is the "population mixing" hypothesis, which states that pediatric leukemia may arise from immune naïve individuals' exposure to a common pathogen after interaction with infected populations. Lastly, a third hypothesis by Smith and colleagues[100] postulates that oncogenic viruses transmitted *in utero* or during the first year of life are able to infect immature lymphocytes and promote leukemia through mechanisms of direct transformation. However, no specific infection has been causally linked to ALL.

Recently, cytomegalovirus (CMV) has emerged as a potential moderate-to-strong risk factor of ALL. CMV is a member of the herpesvirus family (HHV 5) and is capable of transplacental infection during pregnancy. In a report by Francis *et al*., congenital CMV (cCMV) infection was assessed by droplet digital polymerase chain reaction (ddPCR) in a population-based sample of newborn dried blood spots (DBS) from 268 ALL cases and 270 controls in California.[78] In total, cCMV was detected in 9.7% of ALL cases but only 3.0% of controls for a highly significant odds ratio (OR) of 3.71 (95% confidence interval (CI): 1.71 – 8.95). Risk was also reportedly higher among Hispanics (OR: 5.90, 95% CI: 1.89 – 25.96). In a second study by Wiemels *et al*. congenital and early-life, clinically-recognized, CMV infection and subsequent ALL was investigated in population-based registries of Sweden[79]. The hazard ratio (HR) of hematologic malignancy among children with any medically documented cCMV or early-life acquired CMV was 11.2 (95% CI: 5.8 – 21.5).

Together, the two prior studies suggest that prenatal CMV infection substantially increases risk of childhood ALL. If true, congenital and early-life acquired CMV infection could represent the first modifiable risk factor for childhood ALL. As universal newborn screening for cCMV is in development, it is particularly important to establish through replication whether CMV infection at birth is a risk factor for ALL. We therefore conducted a large, population-based study of newborn DBS to compare CMV prevalence at birth in ALL cases and controls.

Methods

Selection of Cases and Controls and Data Collection

Congenital CMV infection was assessed in a population-based case-control study through the Michigan BioTrust for Health (MBH). MBH is a Michigan Department of Health and Human Services program that oversees the consent, advisory boards, and research use of residual newborn dried blood spots (DBS) collected shortly after birth. MBH also conducts routine linkage of the DBS repository to the birth and cancer registries and provides these data to approved researchers. Cases consisted of children 0-14 years of age between 1987 - 2014 diagnosed with ALL (ICD-O code morphology 9835, 9836, and 9820) identified through the Michigan Cancer Surveillance Program (MCSP) and born in Michigan on or after October 1, 1987. Cytogenetic and molecular data on cases was abstracted through collaboration with seven major Michigan children's hospitals that treat >95% of ALL in Michigan through an instrument programmed in REDCap. Controls were randomly selected by MBH and matched 4:1 on year of birth, sex, and mother's reported race/ethnicity. Sex and mother's race were matched for efficient control of confounding while year of birth was matched to have equivalent windows of exposure. Birth characteristic data on cases and controls were obtained from linkage to the birth registry and included previously identified risk factors, such as smoking during pregnancy and Cesarean delivery .[103,104] Due to data suppression rules by MBH, any potential identifiable health information, such as birth year, mother's and father's age, were categorized. The study was approved by the institutional review boards of the University of Minnesota and the Michigan BioTrust for Health.

DNA Extraction and CMV Assay

Although a laboratory assay cannot identify the route of acquisition, since DBS were collected 24-48 hours after birth the strong presumption is that the CMV DNA detected in this medium represents congenital infection. Therefore, we use the term "congenital CMV infection (cCMV)" throughout this manuscript. Detection of CMV DNA in DBS of cases and controls was performed as described elsewhere[105], with some modifications. One 6 mm punch from the newborn DBS card was provided by the MBH. DNA was extracted from this punch using the GenTegra GenSolve DNA Complete kit (GenTegra LLC, California, USA). Briefly, the 6 mm punch was placed in a tube, 609 μL lysis solution (GenTegra) and 11 μL of proteinase K (20 mg/mL) was added, and the mixture incubated at 56 °C for 1.5 hours with agitation at 1400 rpm. Samples were then transferred to a spin basket and centrifuged, with a Recovery Solution (GenTegra) added to the sample. This was followed by DNA purification following the manufacturer's protocol (GenTegra). Samples were eluted in 50 μL of elution buffer, and stored as necessary at -20°C.

Quantitative multiplex PCR was performed as described elsewhere.[106] Briefly, 7 μL of eluate was used in a reaction volume of 35 μL using the LightCycle 480 PCR system (Roche). Primers and probes for the CMV IE gene were utilized with NRAS used as a housekeeping gene to confirm recovery of amplifiable DNA from DBS. PCR was run in triplicate, and a sample was considered positive if at least 2 of the 3 replicates were positive or if one well was positive on two separate, independent experimental runs. Four cases (0.3%) and 13 controls (0.3%) did not have sufficient yield of genomic DNA from the DBS and were scored indeterminate for cCMV. Three control DBS (0.1%) had CMV

DNA below the pre-determined limit of detection (LOD) threshold and were therefore scored equivocal.

Statistical Analysis

Pearson's X^2-test was used to assess categorical data differences between cases and controls. Continuous variables (e.g. birth weight, parental age) were examined for a linear association with ALL and were categorized if nonlinear. To assess an association between cCMV infection and ALL, conditional logistic regression was used, thus controlling for matching factors, to construct both univariate and multivariate models. Multivariate models were additionally adjusted for mother's age at birth, maternal diabetes, birth weight, categorical gestational age, and presence of birth defects. Variables in the model for which there was missing data were included as an unknown category. To assess the effect of potential confounding, we conducted a stratified analysis of cCMV infection among cases and controls. Analysis of cases with available subtype data was conducted by both matched and un-matched analysis using exact methods. All comparisons made between cases and controls were conducted using 2-sample *t* tests for continuous variables or Pearson's Chi-Square or Fisher's exact tests for categorical variables. All statistical analyses were performed using Stata/IC Version 15.1 (StataCorp LP, College Station, Texas, USA).

Results

MBH identified 1,199 eligible ALL cases and 4,796 matched controls. Ten (0.83%) of the cases did not have a DBS available for testing, therefore the matched set (n=40

controls) was excluded. The final study population included 1,189 cases and 4,756 controls (**Table 2- 1)**. The mean age at ALL diagnosis among the cases was 4.5 years (SE=0.09). The cases and controls both had a greater proportion of males (57.6%, respectively) (p=1.0) and were predominately born between 1993 – 1998 (53.9%, cases and controls) (p=1.0). Overall, the mean birth weight was higher among the ALL cases (3,448.6 grams, SE=16.7) than controls (3,385.5g , SE=8.3) (p=0.0007). The distribution of mother's categorical age was higher among ALL cases than controls, with 13% of cases having mothers who were ≥35 years of age at the time of birth compared to 11.5% of controls (p=0.02). Despite matching on mother's race/ethnicity, a higher proportion of cases had mothers who identified as White (83%) than compared to controls (83.2%) (p=0.014). There were no significant differences in mother's level of education between cases and controls (p=0.44). Father's categorical age at birth differed somewhat between cases and controls, in which a greater proportion of fathers were ≥35 years of age at the time of birth among cases (25%) than controls (20.9%) (p=0.013).

The mean weight gain during pregnancy was also similar across mothers of cases (31.4 lbs, SE=0.42) and controls (31.0 lbs, SE=0.20) (p=0.39) (**Table 2- 1**). There were no significant differences between mothers of cases and controls in smoking or alcohol use during pregnancy as recorded on birth certificates (p=0.96 and p=0.90, respectively). There was a slightly higher prevalence of pre-pregnancy or gestational diabetes among mothers of cases (4.1%) than controls (3.2%), but this difference was not statistically significant (p=0.12). Among controls, there was a slightly higher prevalence of uterine or vaginal bleeding during pregnancy (1.3%) than cases (0.6%) (p=0.04). Overall, mode of delivery and plurality of birth was similar across cases and controls (p=0.54 and 0.26,

respectively), with most births being single (98%, cases and controls) and spontaneous vaginal delivery (68.3% and 70.6%, respectively). There was a significantly higher prevalence of birth defects among ALL cases (3.4%) compared to controls (1.7%) ($p<0.0001$). However, there were no differences across Kotelchuck indexes ($p=0.11$).

Immunophenotype was available from the Michigan cancer registry for 536 (45.1%) of cases, of those 62 (11.6%) were T-ALL and 474 (88.4%) were B-ALL. Those with immunophenotype were more likely to be born after 1997 (83.2%) than those missing immunophenotype (31.4%) ($p<0.001$). Cytogenetic data was available for 226 (21%) of cases (**Table 2- 2)**. Cases with available subtype data were slightly older at diagnosis (aged 5.4 years) compared to cases who were missing subtype (aged 3.8 years) ($p<0.001$). Additionally, cases with cytogenetic subtype data were more likely to be born after 1997 (84.2%) than those missing (47.8%) ($p<0.001$). However, cases with and without subtype were similar across sex (57.3% and 57.9% male, respectively) and mother's race/ethnicity (83.6% and 82.9% White, respectively). Among those with cytogenetic subtype available, the distribution of B-ALL subtypes is shown in **Figure 2-1**. Hyperdiploid B-ALL was the most prevalent subtype (58%) followed by *ETV6-RUNX1* translocation (24%).

Overall, we detected cCMV DNA in 6/1,189 (0.5%) of ALL cases and 21/4756 (0.4%) of controls (**Table 2- 3)**. The crude odds of cCMV infection were not statistically different between ALL cases and controls (OR: 1.14, 95% CI: 0.46 – 2.83). In the multivariate model, the odds of cCMV exposure appeared slightly elevated among cases compared to controls, however, this point estimate was measured very imprecisely ($OR_{adjusted}$: 1.30, 95% CI: 0.52 – 3.24). When stratified by age at diagnosis, cCMV

exposure appeared elevated among cases diagnosed between ages 1-4 years compared to their matched controls but this difference was not statistically significant (OR: 1.67, 95% CI: 0.59 – 4.73). Among B-ALL cases, the odds of cCMV similarly appeared elevated compared to their matched controls, however, this point estimate was also measured very imprecisely (OR: 4.0, 95% CI: 0.56 – 28.40). There were no cCMV positive cases among recognized T-ALL. There were 2 hyperdiploid cases among those with subtype data that were cCMV positive. When we compared CMV prevalence among the hyperdiploid ALL cases and their matched controls, the model did not converge due to lack of exposure among the matched controls. However, compared to all controls (n=4,756) in an unmatched analysis, hyperdiploid ALL cases were 6.26 times more likely to be CMV positive (95% CI: 1.44 – 27.19). We also assessed this effect in an unmatched analysis including all matched controls of cases who had subtype data available (n=2,144) and found the odds of cCMV exposure was 13.4 times greater among hyperdiploid cases compared to controls (95% CI: 1.25 – 83.21).

Among those positive for cCMV DNA, the mean viral copies of CMV per mL blood was not significantly different across cases (35,966.7 copies/mL, SE:18,263) and controls (34,368.1 copies/mL, SE: 19,277) (p=0.97) (**Figure 2- 2**). When considering the mean copies CMV per microgram of genomic DNA, cases had a higher viral load (3,301.3 copies/μg, SE: 2,684.7) than controls (840.1 copies/μg, SE: 274.1), but this difference was not statistically significant (p=0.1)

In the stratified analysis of CMV infection among cases and controls, we did not detect any statistically significant associations (**Table 2- 4)**. Among infants whose mothers had some post-high school education, the overall prevalence of CMV was

significantly different across cases and controls (Fisher's exact p-value = 0.041), however the crude odds ratio was not statistically significant (OR: 4.97, 95% CI: 0.84 – 34.07). We were not able to detect any differences in the odds of CMV infection by mother's race/ethnicity, as all cases (n=6) had White mothers.

Discussion

Congenital CMV infection has recently emerged as a potential modifiable risk factor of pediatric acute lymphoblastic leukemia. In response to the two extant studies of the topic by Francis et al[78]. and Wiemels et al.[79], we conducted a large, population-based case-control study of cCMV infection and pediatric ALL. In the entire study of 1,189 ALL cases and 4,756 controls we did not detect an association between leukemia and exposure to cCMV infection in the main analysis. However, among cCMV positive individuals CMV DNA levels trended higher among individuals who went on to develop ALL. In addition, among hyperdiploid ALL cases, the odds of being cCMV positive were six times greater than unmatched controls. Below we discuss in brief the hypothetical basis for the investigation of cCMV and ALL and compare our findings with those of Francis et al. and Wiemels et al.

There is a growing body of evidence of an infectious etiology for pediatric ALL, and there are three main hypotheses on the nature of this pathogenesis: Greave's 'delayed infection' hypothesis, Kinlen's 'population mixing' hypothesis', and Smith's hypothesis of 'direct transformation' by an infectious agent.[98–100] Of these, CMV best fits the criteria of Smith's hypothesis, which states that the infectious agent causing ALL should possess: 1) the ability to induce genomic instability; 2) specific effects on B lymphocytes,

specifically in the context of B-precursor ALL; 3) higher rates of infection in regions with lower socioeconomic status; 4) limited general oncogenic potential; 5) minimal symptoms associated with the primary infection; and 6) the ability to cross the placenta and infect the fetus, but not cause severe fetal abnormalities.[100]

CMV is capable of causing direct chromosomal breakage in congenital infection, which is suspected to be related to its teratogenic properties.[107] In addition, CMV encodes several proteins that modulate both cell cycle control and the host DNA damage response.[108] The link between cCMV and B lymphocytes are CD34+ cells, which are early hematopoietic progenitor cells in bone marrow,[109,110] and the cell type in which CMV establishes latency. A study by Albano et al. investigated the impact of cCMV on hematopoietic progenitor cell concentrations in cord blood and found that among infants with cCMV infection, CD34+ cell populations were roughly 2.6 times greater than those of matched controls.[111] This finding suggests a mechanism by which cCMV increases the risk of ALL by encouraging proliferation of cells vulnerable to transformation. Collectively these observations support the plausibility of CMV being involved in the etiology of ALL. Aside from this plausibility, two independent studies have suggested a moderately strong association between cCMV infection and ALL.

The initial study by Francis et al. had a study design most similar to ours[78]. The authors conducted a case-control study of newborn dried blood spots from 268 ALL cases and 270 cancer-free controls randomly drawn from the Childhood Cancer Record Linkage Program (CCRLP), and matched on date of birth, race, and sex. In comparison, our case-control study was nearly ten times larger with a similarly strong, population-based study design that used newborn DBS to capture prenatal CMV exposure. The rates

of CMV DNA positivity were very different in our two studies, with the positive prevalence 7.5 times higher in the controls from Francis et al. compared to our controls. This raises the possibility of technical differences contributing to our differing results. One major difference between our studies was the use by Francis et al. of droplet digital PCR (ddPCR) to detect CMV DNA in the newborn DBS, while we used a quantitative PCR method that has been optimized for detection of CMV DNA in newborn DBS.[106] While ddPCR has shown to have increased precision over qPCR in certain applications, the methods have similar sensitivity.[112,113] Therefore it is unlikely that the platform for CMV detection explains the differing results in our two studies. Differences in the starting material may have contributed to differences in the detection of CMV. Francis et al. reported using a quarter of a Guthrie spot, equivalent to about 33.2 mm^2 area, which was more than the single 6 mm punch area of 28.27^2 mm used in our study.[114] Therefore, it could be expected that the likelihood of detecting cCMV would increase with the amount of material sampled. While we cannot be certain what is driving the differences in CMV DNA prevalence at birth in the two studies, cCMV infection is detected in 0.45% of newborn dried blood spots, which is consistent with our results.[106] Further, universal screening in high income countries have shown cCMV prevalence is consistently about 0.6%, even when using urine and saliva.[115,116] It is also possible that the method used by Francis et al. had a lower limit of detection of quantification, and further development of methods is warranted.

An outstanding question is whether the severity of cCMV is associated with future ALL risk. In Wiemels et al.[79] both congenital and early-life acquired CMV infection was evaluated for future risk of hematologic malignancy in population-based

registries of Sweden.[79] CMV infection was identified through linkage of the cohort and their mothers to the nationwide patient registry. Through their passive screening, the prevalence of *clinically recognized* CMV infection at birth was 0.0066% in non-cases and 0.088% in children who later were diagnosed with a hematologic malignancy (HR: 14.8, 95% CI:4.8 – 45.9). Congenital CMV infection is clinically recognized in 10-15% of all infected infants, as the majority of babies are asymptomatic.[29,117] Therefore, the CMV infections investigated in Wiemels et al. were likely severe. However, we observed higher CMV levels in cases versus controls. Newborn screening for cCMV has recently revealed viral load at birth may be higher with asymptomatic infections.[118] We do not know who in our study had a clinically recognized cCMV infection; further investigation into this relationship is necessary.

Our findings suggest a CMV-ALL association may be specific to hyperdiploid ALL, consistent with recent reports from diagnostic specimens.[80] Cytogenetic subtype was available for 21% of cases. When stratified by ALL subtype, we found hyperdiploid ALL cases had 6.3 the odds of cCMV compared to all controls in our study, albeit based on only 2 exposed cases. Further, hyperdiploid cases had 13.4 the odds of cCMV when compared to all controls who had cases with subtype data. High hyperdiploid ALL (generally defined as 50-67 chromosomes) was very recently shown to be associated with CMV by Gallant et al. in case-only analyses.[80] Half of all bone marrow biopsies were CMV DNA positive, and bone marrow biopsies from B-cell ALL cases were more likely to be CMV positive compared to T-cell ALL (OR 1.63 (95% CI 0.88 – 3.06). Considering just B-cell ALL, the biopsies from high hyperdiploid ALL cases were 1.7 times more likely to be CMV positive than *ETV6-RUNX1* ALL, and 2.71 times more

likely to be in the upper tertile of CMV-load (95% CI: 1.34 – 4.73).[80] Our results, in combination with those from Gallant et al, strongly suggest that CMV contributes specifically to hyperdiploid ALL.

Our study had several strengths. First, this study was a very large, population-based case-control study of 1,189 ALL cases and 4,756 controls, all with dried blood spots obtained immediately after birth. Compared to the study by Francis et al., our study is nearly ten times larger. Second, data linkage through the Michigan BioTrust for Health enabled us to examine potential associations with birth characteristics and parental demographics and leukemia. Lastly, the use of nested samples taken shortly after birth make the temporality between measurement of cCMV and development of ALL clear. Since all samples were collected immediately following birth any detectable DNA would have by necessity been passed *in utero*. Furthermore, this is an advantage over the study by Wiemels et al., since they used only medical records which reported only clinically recognized CMV disease. Cases and controls were selected by MBH and the laboratory was blinded to case and control status when assaying DBS for CMV DNA. Thus, any misclassification of cCMV exposure would be nondifferential with respect to case-control status and would bias our estimates toward the null. We also controlled for known confounders by matching on year of birth, sex, and mother's race/ethnicity, and adjusted for other risk factors in the multivariate analysis. Lastly, when we consider the width of the confidence interval around our point estimate in adjusted analysis, it leaves room for the possibility of an effect as large as 3.24, which is of the same magnitude of the estimate by Francis et al.

A limitation of our study is that this was a largely White population which limited our ability to look in subgroups defined by demographics. Francis et al. showed that risk was more pronounced in Hispanics (OR: 5.90, 95% CI: 1.89 – 25.96) than in non-Hispanic Whites (2.10, 0.69 – 7.13).[78] Another limitation was that Michigan DBS were stored at ambient temperature while all DBS specimens collected after 2009 in the study by Francis et al. were at -20°C. This may explain why we had a lower prevalence of CMV DNA in our samples than in Francis et al, due to degradation of material. However, DBS stored at room temperature in other studies of CMV DNA has not been associated with diminished quality[119] and our estimated prevalence of CMV DNA was in line with other estimates of cCMV infection including using gold standard method of detection like urine.[120] Thus, it seems more likely that the assay by Francis et al. had a lower limit of detection. Lastly, we had immunophenotype and cytogenetic subtype data available for only a small portion of cases which contributed to imprecision of our point estimates.

Overall, while we did not find evidence of an association between cCMV infection and pediatric ALL in our primary analysis, we did detect an association among hyperdiploid ALL cases based, however, on small numbers. The partial support given here to the association of cCMV with ALL, previously seen in two epidemiologic studies, and its accumulating biologic plausibility argue for continued research. Two jurisdictions have recently initiated universal cCMV screening -Minnesota and Ontario -which may present the best locales in which to examine this emerging association further.[121,122] Future work may also consider incorporating additional biomarkers, such as DNA methylation, that may help establish a mechanistic link between cCMV and ALL. Finally, work that considers not just CMV infection *in utero*, but also early childhood

infections, is needed to fully understand the relationship of this highly prevalent virus and childhood ALL.

Table 2- 1. Demographic characteristics of ALL cases and matched controls.

	ALL cases		Controls		
Characteristics	(n=1,189)	%	(n=4,756)	%	p-value
Mean age at diagnosis (SE)	4.5 (0.09)		N/A	-	-
Birth year*					1.0
1988-1992	212	17.8 %	848	17.8%	
1993-1997	326	27.4 %	1,304	27.4%	
1998-2002	315	26.5 %	1,260	26.5%	
2003-2007	237	19.9 %	957	20.1%	
2007-2012	99	8.3%	387	8.1%	
Sex*					1.0
Female	504	42.4 %	2,016	42.4%	
Male	685	57.6 %	2,740	57.6%	
Mother's age at birth, years					<0.001
<25	346	29.1 %	1,592	33.5%	
25-34	679	57.1 %	2,617	55.0%	
35+	153	12.9 %	547	11.5%	
Unknown	11	0.9%	0	0%	
Mother's race/ethnicity*					0.014
White	987	83.0 %	3,959	83.2%	
Black	103	8.7%	420	8.8%	
Other	26	2.2%	122	2.6%	
Hispanic	56	4.7%	231	4.9%	
Unknown	17	1.4%	24	0.5%	
Mother's level of education					0.44
High school	556	46.8 %	2,332	49.0%	
Some post-high school	326	27.4 %	1,205	25.3%	
College	293	24.6 %	1,168	24.6%	
Unknown	14	1.2%	51	1.1%	

Father's age at birth, years					0.013
<25	158	13.30 %	732	15.4%	
25-34	590	49.60 %	2,426	51.0%	
35+	297	25.00 %	994	20.9%	
Unknown	144	12.10 %	604	12.7%	
Father's race/ethnicity					0.52
White	899	75.6 %	3,500	73.6%	
Black	56	4.7%	264	5.6%	
Other	25	2.1%	114	2.4%	
Hispanic	45	3.8%	214	4.5%	
Unknown	164	13.8 %	664	14.0%	
Father's level of education					0.92
High school	477	40.1 %	1,900	40.0%	
Some post-high school	247	20.8 %	954	20.1%	
College	298	25.1 %	1,206	25.4%	
Unknown	167	14.1 %	696	14.6%	
Pregnancy and Birth Characteristics					
Pregnancy					
Mean weight gain during pregnancy, lbs (SE)	31.4 (0.42)		31.0 (0.20)		0.39
Smoking before or during pregnancy					0.96
Yes	198	16.7 %	804	16.9%	
No	968	81.4 %	3,864	81.2%	
Missing	23	1.9%	88	1.9%	
Alcohol use					0.90
Yes	17	1.4%	60	1.3%	
No	1,147	96.5 %	4,595	96.6%	
Missing	25	2.1%	101	2.1%	

Pre-pregnancy or gestational diabetes					0.12
Yes	49	4.1%	153	3.2%	
No	1,140	95.9 %	4,603	96.8%	
Chronic hypertension					0.3
Yes	12	1.0%	34	0.7%	
No	1,177	99.0 %	4.722	99.3%	
Gestational hypertension					0.34
Yes	9	0.8%	25	0.5%	
No	1,180	99.2 %	4,731	99.5%	
Uterine/vaginal bleeding					0.04
Yes	7	0.6%	61	1.3%	
No	1,182	99.4 %	4,695	98.7%	
Previous Caesarean section					0.49
Yes	143	12.0 %	538	11.3%	
No	1,046	88.0 %	4,218	88.7%	
Birth					
Mean birth weight, grams (SE)	3448.6 (16.7)	-	3385.5 (8.3)	-	0.0007
Gestational age					0.32
<37 weeks	90	7.6%	368	7.7%	
≥ 37 weeks	1,099	92.4 %	4,379	92.1%	
Unknown	0	0%	9	0.2%	
Method of delivery					0.54
Vaginal, spontaneous	812	68.3 %	3,357	70.6%	
Vaginal, forceps	25	2.1%	99	2.1%	
Vaginal, vacuum	43	3.6%	149	3.1%	
Cesarean	300	25.2 %	1,125	23.7%	
Unknown	9	0.8%	26	0.6%	

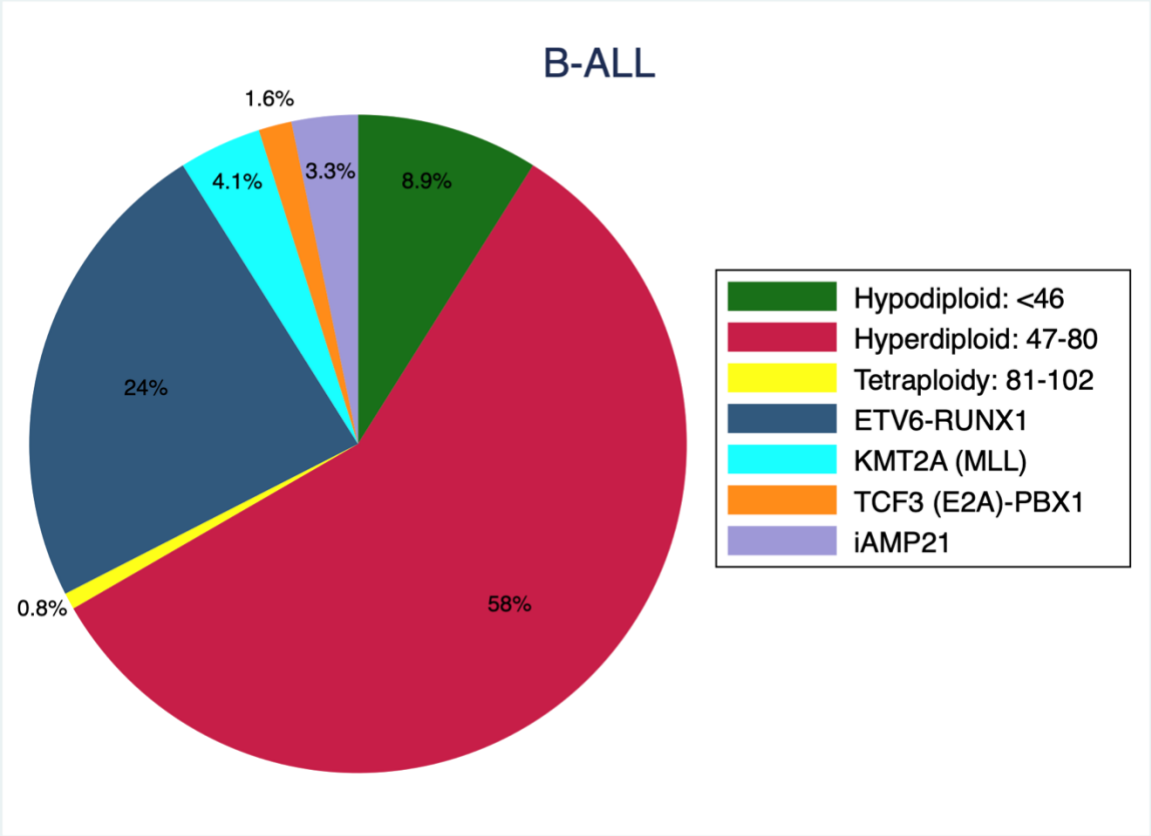
Plurality of birth					0.26
Single	1,165	98.0 %	4,663	98.0%	
Twin	22	1.8%	88	1.9%	
Triplet	1	0.1%	5	0.1%	
Quadruplet	1	0.1%	0	0.0%	
Birth injury					0.51
Yes	3	0.3%	18	0.4%	
No	1,186	99.7 %	4,738	99.6%	
Birth defects					<0.0001
Yes	40	3.4%	81	1.7%	
No	1,116	93.9 %	4,611	97.0%	
Missing	33	2.8%	64	1.3%	
Kessner index					0.54
Not collected	71	6.0%	275	5.8%	
Adequate	887	74.6 %	3,449	72.5%	
Intermediate	153	12.9 %	672	14.1%	
Inadequate	73	6.1%	334	7.0%	
Unknown	5	0.4%	26	0.6%	
Kotelchuck index					0.11
Not collected	51	4.3%	204	4.3%	
Adequate plus	395	33.2 %	1,379	29.0%	
Adequate	495	41.6 %	2,091	44.0%	
Intermediate	112	9.4%	488	10.3%	
Inadequate	89	7.5%	407	8.6%	
Unknown	47	3.9%	187	3.9%	

*Indicates matching factor. P-values calculated by Pearson's Chi-Square statistic for categorical variables or by two-sided t-test for continuous variables. Abbreviations: SE – standard error;

Table 2- 2. Comparison of cases with available leukemia subtype to cases without, by demographic characteristics

	Cases with subtype		**Cases without subtype**		
Characteristics	(n=536)	%	(n=653)	%	p-value
Mean age at diagnosis (SD)	5.4 (3.6)		3.8 (2.8)		<0.001
Birth year					<0.001
1988-1992	16	3.0%	196	30.0%	
1993-1997	74	13.8%	252	38.6%	
1998-2002	168	31.3%	147	22.5%	
2003-2007	182	34.0%	55	8.4%	
2007-2012	96	17.9%	3	0.5%	
Sex					
Female	229	42.7%	275	42.1%	
Male	307	57.3%	378	57.9%	
Mean birth weight, grams (SD)	3418.3 (571.0)		3473.4 (580.7)		
Gestational Age					0.063
<37 weeks	49	9.1%	41	6.3%	
37+ weeks	487	90.9%	612	93.7%	
Mother's age at birth, years					0.43
<25	148	27.6%	198	30.3%	
25-34	306	57.1%	373	57.1%	
35+	75	14.0%	78	11.9%	
Unknown	7	1.3%	4	0.6%	
Mother's race/ethnicity					0.23
White	435	81.2%	552	84.5%	
Black	52	9.7%	51	7.8%	
Other	9	1.7%	17	2.6%	
Hispanic	31	5.8%	25	3.8%	
Unknown	9	1.7%	8	1.2%	

Figure 2- 1. Distribution of B-ALL subtypes.



Cytogenetic data was available for 226 (21%) of cases. The chart above describes the distribution of B-ALL subtypes diagnosed.

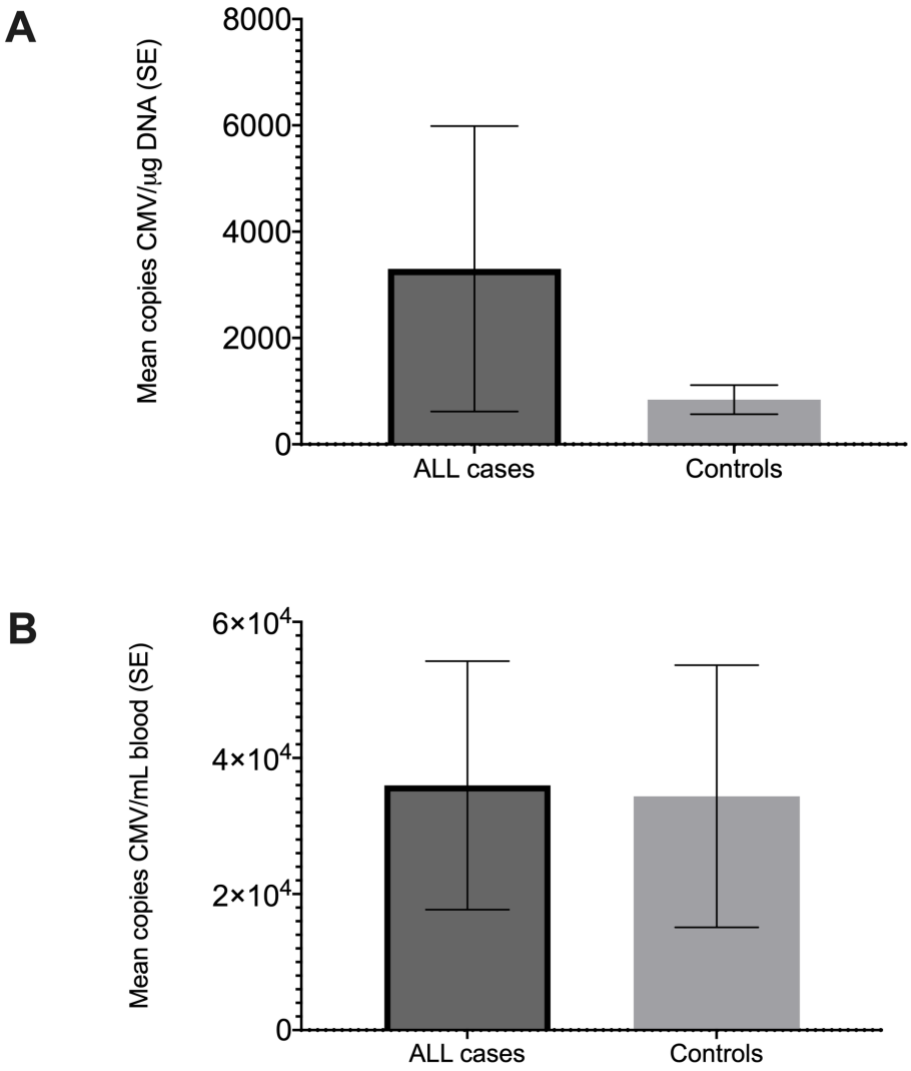
Table 2- 3. Prevalence of cCMV in dried blood spots of ALL cases compared to matched controls, by diagnostic factors.

	Proportion CMV positive (%)						
	ALL cases	**Controls**	**P-value**	**OR (95% CI)**	**OR adjusted (95% CI)**	**OR unmatched[1] (95% CI)**	**OR unmatched[2] (95% CI)**
Overall	6/1,189 (0.5%)	21/4,756 (0.4%)	0.77	1.14 (0.46 - 2.83)	1.30 (0.52 - 3.24)		
Age at diagnosis							
0	0/51 (0%)	1/204 (0.5%)	1.0	-			
1-4	5/692 (0.7%)	12/2,768 (0.4%)	0.49	1.67 (0.59 - 4.73)			
5-9	0/322 (0%)	8/1,288 (0.6%)	0.37	-			
10-14	1/124 (0.8%)	0/494 (0%)	0.23	-			
B-ALL	2/474 (0.4%)	2/1,896 (0.1%)	0.17	4.0 (0.56 - 28.40)			
T-ALL	0/62 (0%)	3/248 (1.2%)	1.0				
Ploidy							
Hypoploidy (<46)	0/11 (0%)	0/44 (0%)	-	-			
Hyperploidy (47-80)	2/74 (2.7%)	0/296 (0%)	0.04	-	-	6.26 (1.44 - 27.19)	13.37 (1.25 – 83.21)
Tetraploidy (81-102)	0/1 (0%)	0/4 (0%)	-	-			

Odds ratio and 95% confidence interval calculated by conditional logistic regression. Multivariate model adjusted for mother's categorical age, mother's education level, father's race/ethnicity, father's education level, father's categorical age, maternal diabetes,

delivery method, plurality, birth weight, and birth defects. Odds ratios and 95% confidence intervals were calculated first by conditional logistic regression among hyperdiploid ALL cases and their matched controls, but the model did not converge. 1) An odds ratio calculated for hyperdiploid cases compared to all controls (n=4,756) using exact methods. 2) Odds ratio calculated in an unmatched analysis for hyperdiploid cases compared to all controls who had cases with available subtype data (n=2,144) using exact methods. Abbreviations: CI – confidence interval; OR – odds ratio;

Figure 2- 2. Mean CMV viral load among ALL cases and controls.



A). Viral load expressed as mean viral copies CMV/µg genomic DNA. B). Viral load expressed as copies CMV/mL blood. Abbreviations: SE – standard error;

Table 2- 4. Congenital CMV status among ALL cases and controls, stratified by demographic characteristics.

	All Cases		Controls			
Characteristic	**CMV +**	**CMV -**	**CMV+**	**CMV-**	**p-value***	**OR (95% CI)**
Sex						
Female	3	501	9	2,007	0.72	1.34 (0.23 - 5.38)
Male	3	693	12	2,725	1	0.98 (0.18 - 3.66)
Birth weight (grams)						
Low (<2500) (n=336)	1	70	3	262	1	1.25 (0.02 - 15.80)
Normal (2500 - 4000)	5	934	16	3,927	0.58	1.31 (0.38 - 3.76)
High (>4000)	0	179	2	543	1	-
Gestational Age						
<37 weeks	1	89	2	366	0.48	2.06 (0.03 - 39.83)
37+ weeks	5	1,094	19	4,357	1	1.05 (0.31 - 2.92)
Mother's age at birth, years						
<25	4	342	12	1,578	0.51	1.54 (0.36 - 5.11)
25-34	2	677	6	2,610	0.67	1.29 (0.13 - 7.21)
35+	0	153	3	544	1	-
Mother's race/ethnicity						
White	6	981	13	3943	0.24	1.86 (0.58 - 5.24)
Black	0	103	5	415	0.59	-
Other	0	26	1	121	1	-
Hispanic	0	56	1	230	1	-
Mother's level of education						
High school	2	554	17	2314	0.56	0.49 (0.05 - 2.08)
Some post-high school	4	322	3	1201	0.041	4.97 (0.84 - 34.07)
College	0	293	1	1166	1	-
Father's age at birth, years						
<25	1	157	3	729	0.54	1.55 (0.03 - 19.41)
25-34	2	588	7	2,416	0.69	1.17 (0.12 - 6.17)

35+	2	295	3	991	0.33	2.24 (0.19 - 19.63)
Father's race/ethnicity						
White	4	895	9	3,488	0.32	1.73 (0.39 - 6.22)
Black	0	56	1	263	1	-
Other	0	25	0	114	-	-
Hispanic	0	45	1	213	1	-
Father's level of education						
High school	2	475	7	1,892	1	1.14 (0.11 - 5.99)
Some post-high school	2	245	3	950	0.27	2.59 (0.21 - 22.67)
College	0	298	1	1,204	1	-

P-values were calculated by Fisher's exact test. Abbreviations: CI – confidence interval; OR – odds ratio

Chapter 3: Cytomegalovirus infection and the risk of hematologic malignancy among solid organ transplant recipients

Introduction

Cytomegalovirus (CMV) is a ubiquitous human herpesvirus (HHV-5) infecting roughly 60% of the U.S adult population.[123] CMV typically only causes asymptomatic or mild flulike illness in otherwise healthy adults, but in transplant recipients CMV poses a significant risk of morbidity and mortality. CMV infection following solid organ transplantation (SOT) can manifest as undifferentiated fever; CMV pneumonia, hepatitis, and enteritis; and CMV retinitis. CMV is also a significant co-factor in allograft rejection and death. [39–41,124] The greatest risk for CMV disease post-SOT occurs when the recipient develops primary CMV infection under sustained immunosuppressive regimens used for SOT. [43,44] Primary infection can occur if the recipient is CMV seronegative at the time of transplant and is either exposed in the community or, more commonly, when receiving an organ transplant from a donor who is CMV seropositive.

Though typically associated with Epstein-Barr virus (EBV) infections, post-transplant lymphoproliferative disorder (PTLD) has also been shown to develop in CMV infected transplant recipients.[45,46] In a retrospective cohort study of primary EBV infection among liver transplant recipients, CMV disease post-transplant was significantly associated with PTLD in the presence of EBV (RR: 7.3, 95% CI: 2.36 – 22.6).[47] CMV has also been shown as an independent predictor of PTLD; a prospective study of risk factors for PTLD among recipients of solid organs found CMV seromismatch (recipient CMV negative, donor CMV positive) was associated with a 3-

fold risk of PTLD compared to no CMV seromismatch (IRR: 3.02, 95% CI: 1.11, 8.27).[46] Similarly, CMV disease during the first year post-transplant was also associated with subsequent non-Hodgkin lymphoma (NHL).[48] CMV infection as it relates to other hematologic malignancies post-transplantation is not well-defined.

Recently, congenital CMV (cCMV) has emerged as a potential risk factor for acute lymphoblastic leukemia (ALL) the most common pediatric malignancy.[81] In a population-based sample, CMV from newborn dried blood spots (DBS) was detected in 9.7% of ALL cases but in only 3.0% of healthy controls (OR: 3.71, 95% CI: 1.71-8.95).[78] A second study conducted in population-based registries of Sweden found that medically documented CMV acquired in early childhood was associated with a dramatic increase in risk of hematological malignancies, including ALL (HR: 11.2; 95% CI: 5.8-21.5).[79]

Transplant recipients have an increased risk of cancer following transplantation and prior donor and recipient CMV infections are common. We therefore examined the potential association between CMV infection post-transplant and the risk of hematologic malignancy among solid organ transplant recipients. Using data from the Transplant Cancer Match Study (TCM), which uses registry-based ascertainment of cancer in transplant recipients, this represents the largest prospective study of CMV and hematologic cancer incidence to date.

Methods

Transplant Cancer Match Study

The study cohort consisted of solid organ transplant recipients enrolled in the TCM study. The TCM study (http://transplantmatch.cancer.gov) has been previously

described in detail.[125] Briefly, computer-based linkages were made between the Scientific Registry of Transplant Recipients (SRTR) and 32 U.S state and regional cancer registries. Transplant recipients are included in TCM if, based on their address at the time of listing or transplantation, they resided in a region covered by one of the participating cancer registries. The SRTR includes data regarding all US solid organ transplants in participating registries since 1987.

Since 2015, the United Network of Organ Sharing (UNOS) has required reporting of recipient and donor CMV serologies at the time of transplant, and these data are available in the SRTR. [126] The primary exclusion criterion in our present study was missing CMV IgG status at the time of transplant for the recipient or for the donor when the recipient was CMV IgG negative. Data on CMV IgG serostatus at the time of transplant was unavailable for most recipients (69%) between 1987 – 1999, therefore, we restricted analysis to the years 2000 and onward. Of the 672,603 individuals in the US who received a first transplant during 1987-2017, we excluded 211,973 (31.5%) because they were transplanted outside of participating cancer registry regions, 117,404 (17.5%) because they were transplanted before the year 2000 or lacked follow-up, 29,875 (4.4%) because they had a cancer diagnosis before transplantation, and 1,021 (0.2%) because they had human immunodeficiency virus infection, and 106 (0.02%) if they had more than one donor. Finally, we excluded an additional 19,139 (2.8%) transplant recipients who were missing CMV IgG serostatus pre-transplant, or if they were seronegative, their donor was missing CMV serostatus. After these exclusions, 247,318 (36.8%) transplants were included in the final analysis.

The study is considered non-human subjects research by the National Cancer Institute and was approved, as required, by participating cancer registries.

Cancer Ascertainment

Incident cases of Hodgkin lymphoma, NHL, leukemia, and myeloma were identified from the 32 linked population-based cancer registries using the Surveillance, Epidemiology, and End Results (SEER) program "site recode" based on the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3) histology codes, updated for the Hematopoietic codes based on World Health Organization (WHO) 2008 *Classification of Tumours of Haematopoetic and Lymphoid Tissues.*[127,128] Lymphoma subtypes were classified using ICD-O-3 site and histology codes according to current International Lymphoma Epidemiology (InterLymph) Consortium consensus guidelines.[129]

CMV Risk Groups and Risk Factor Data

CMV serostatus was categorized into three risk groups according to pre-transplant IgG serostatus obtained from SRTR data of donors and recipients, to reflect risk of active CMV infection and disease post-transplant: high risk (recipient negative and donor positive [R-/D+]), intermediate risk (recipient seropositive regardless of donor serostatus [R+]), and low risk (recipient and donor seronegative [R-/D-]).[43]

Data on covariate risk factors were obtained from linked SRTR records, including recipient characteristics (age at transplantation, sex, race/ethnicity), transplant characteristics (organ, number, calendar year), and immunosuppression medications

(induction and baseline immunosuppression maintenance therapies). EBV serostatus for recipients was defined by three risk groups as follows: high risk: recipient negative and donor positive (EBV R-/D+), intermediate risk: recipient positive and donor positive or negative (EBV R+), and low risk: recipient negative and donor negative (EBV R-/D-). EBV status was available for 74% of the cohort. Socioeconomic status (SES) quintiles were constructed by recipients' ZIP code at the time of transplantation using methods developed by Yost et al.[130] Briefly, SES index was assessed for every ZIP Code Tabulation Area (ZCTA) using American Community Survey data on seven measures of SES: percent working class and unemployed, education index, percent living below 150% of the poverty line, median household income, median home value, and median rent. ZCTAs were categorized into quintiles and merged on recipient ZIP code.

Statistical Analysis

Follow-up for cancer started at the time of transplantation and ended at the earliest of death, graft failure, retransplantation, loss to follow-up by SRTR, or end of cancer registry coverage. Crude incidence rates (IRs) for each cancer, defined as the number of outcomes per 100,000 person-years of follow-up, were calculated in recipients by CMV risk group. To compare cancer risk by CMV risk group, we estimated incidence rate ratios (IRRs) with 95% confidence intervals using multivariable Poisson regression models adjusted for recipient sex, age at transplantation, race/ethnicity, SES quintile, transplanted organ (kidney, liver, or other/multiple), and EBV risk group. Included variables in the models were evaluated by backward elimination and likelihood ratio test of the nested models. An offset term of the log of the follow-up time was included in the

models to account for differences in exposure period. Goodness of fit and overdispersion of the models were assessed by deviance goodness-of-fit Chi-square and kurtosis score, respectively. For diffuse large B-cell lymphoma (DLBCL), the most common subtype of NHL, we assessed risk by time post-transplantation by stratifying on follow-up. The interaction of CMV and EBV risk groups on DLBCL risk was assessed by including an interaction term in the multivariable models. Reported *P* values are two-sided. Because our analysis was exploratory, we did not correct for multiple testing and p-values less than 0.05 are considered statistically significant. Stata/MP version 16.1 (StataCorp LP, College Station, Texas) was used for all statistical analyses.

Results

We evaluated 247,318 solid organ transplant recipients with 1,245,369 total person-years of follow-up. **Table 3- 1** describes the demographic and transplant characteristics of these individuals by CMV risk group. Overall, 62.9% of the cohort was CMV seropositive pre-transplant (R+), while 20.3% was CMV seronegative with a seropositive donor (R-/D+) and 16.8% was seronegative with a seronegative donor (R-/D-) The R+ group was more likely to be female (42.2%) than the R-/D- (33.2%) and the R-/D+ groups (32.7%) ($p<0.001$) and was also older (mean age 50.0 years) than seronegative recipient groups (R-/D-: 43.4 years; R-/D+: 44.4 years) ($p<0.001$). The greatest racial/ethnic diversity was among the R+ group, among which 51.5% identified as a racial/ethnic group other than non-Hispanic white. Recipients who were CMV seropositive tended to also be EBV seropositive pre-transplant (67.4%), and this proportion was slightly higher than among R-/D- (62.2%) and R-/D+ (61.7%) groups

(p<0.001). The CMV seronegative groups had a higher proportion of individuals in the two highest SES quintiles (R-/D-: 44.0%; R-/D+: 41.9%) than the R+ group (35.8%) (p<0.001). Kidneys were the most commonly transplanted organ across all three groups (61.4% of transplants overall).

Through the 32 linked population-based cancer registries, a total of 2,339 primary incident hematologic malignancies were identified: 61 Hodgkin lymphoma, 1,786 NHL, 276 leukemias, and 216 myeloma diagnoses. Among NHL histological subtypes, diffuse large B-cell lymphoma (DLBCL) was the most prevalent (66.8%). Sixty-two percent of the leukemia cases were of myeloid lineage, with acute myeloid leukemia (AML) having the greatest incidence (9.0 cases per 100,000 p-y).

Table 3- 2 describes the risk of hematologic malignancy by CMV recipient/donor status post-transplantation. Overall, the risk of Hodgkin lymphoma was 67% lower among the R+ group compared to the lowest-risk group (R-/D-) (IRR: 0.33, 95% CI: 0.18 – 0.62). However, after adjusting potential confounders, which included age at transplant, sex, race/ethnicity, transplanted organ, EBV recipient/donor status, and SES quintile, the association was no longer significant (adjusted IRR [aIRR]: 0.47, 95% CI: 0.19 – 1.18). Risk was also inversely associated among the R-/D+ group compared to the R-/D- group, but was not statistically significant (aIRR: 0.75, 95% CI: 0.29 – 1.96). The risk of DLBCL was significantly lower in both the R+ and R-/D+ groups compared to the R-/D- group. After adjusting for covariates, the R+ group had a 17% reduction in risk (aIRR: 0.83, 95% CI: 0.69 – 1.00) while the R-/D+ group had a 26% reduction in risk (aIRR: 0.74, 95% CI: 0.59 – 0.91). There were no other significant associations within NHL by other specific histologic subtypes.

The risk of lymphoid leukemias as a single entity was not significantly different than that of the low-risk group; however, there was a non-significant inverse association for the R-/D+ group compared with the R-/D- group (aIRR 0.65, 95% CI: 0.23 – 1.89). In contrast, the risk of myeloid leukemias appeared elevated among the R+ (aIRR: 1.39, 95% CI: 0.79 – 2.43) and R-/D+ (aIRR: 1.08, 95% CI: 0.56 – 2.07) groups as compared to the R-/D- group but was not statistically significant.

Associations of CMV serostatus with DLBCL differed by time post-transplant (**Figure 3- 1)**. The greatest reduction in risk was seen immediately following transplantation (0-1.99 years) and 10+ years post-transplant in both the R+ and R-/D+ groups compared to the R-/D- group. For the R+ group, this corresponded to a 37% reduction in risk 0-1.99 years following transplant (aIRR: 0.63, 95%CI: 0.42-0.94) and 34% reduction at 10+ years (0.66, 0.45-0.96). For the R-/D+ group, there was a nonsignificant reduction in risk 0-1.99 years post-transplant (aIRR: 0.73, 95%CI: 0.48-1.12) and a significant 51% reduction 10+ years post-transplant (0.49, 0.29-0.81).

To assess the effect of EBV on the association between CMV and risk of DLBCL, we stratified on EBV recipient and donor serostatus pre-transplant using the same risk groupings we created for CMV recipient and donor status (**Figure 3- 2 and Table 3- 3)**. EBV serostatus was an effect modifier between CMV serostatus and DLBCL (p-value for interaction =0.0006). In the absence of prior CMV infection (CMV R-/D-), there was a significantly increased risk of DLBCL among EBV R-/D+ recipients compared to EBV R-/D- (aIRR: 3.46, 95% CI: 1.50 – 7.95) (p-value =0.004). In contrast, DLBCL risk was not significantly different among the EBV R-/D+ recipients in the CMV R+ group or R-/D+ groups when compared to the recipients who were R-/D- for both CMV and EBV.

We also assessed the potential interaction between induction or maintenance medications and CMV on the risk of DLBCL. **Table 3- 4** describes the frequency of therapies used post-transplant by CMV risk group. The use of induction medication was similar across risk groups, with an average of 81.2% of all recipients undergoing some type of induction therapy. Steroids were the most frequently used (64.6%), followed by polyclonal antibodies (32.1%) and IL2 receptor antagonists (28.0%). For maintenance immunosuppression, tacrolimus or mycophenolate mofetil (MMF) were the most commonly used (82.8%) and typically in conjunction with steroids (78.1%). Among those who were given an IL2 receptor antagonist for induction immunosuppression, the risk of DLBCL was significantly lower, however, the interaction was not significant (p-value for interaction =0.69) **(Table 3- 5)**; Compared to the R-/D- group, there were 31% (aIRR: 0.69, 95% CI: 0.53 – 0.90) and 45% (aIRR: 0.55, 95% CI: 0.38 – 0.79) reductions in risk among the R+ and R-/D+ groups, respectively. In contrast, the risks of DLBCL were significantly elevated among those who received monoclonal antibody therapy in the R+ (IRR: 2.32, 95% CI: 1.02 – 5.24) and R-/D+ (aIRR: 3.28, 95% CI: 1.05 – 10.30) groups (p-value for interaction =0.45). Maintenance medication, in contrast, was an effect modifier between CMV serostatus and DLBCL (p-value for interaction=0.009). Those in the R+ group who were given cyclosporine and/or azathioprine for maintenance immunosuppression also had an elevated risk of DLBCL compared to the R-/D- group (IRR: 1.94, 95% CI: 1.33 - 2.84). However, this effect was not significant among the R-/D+ group (IRR: 1.09, 95% CI: 0.56 – 2.14).

Discussion

Recipients of a solid organ transplant have an elevate risk of cancer, especially for malignancies caused by viral infections.[125,131,132] Virus-associated cancers included NHL and Hodgkin lymphoma (both due to EBV) and anogenital cancers (human papillomavirus). CMV is among the most common viral infections following SOT and has been implicated in the development of PTLD and NHL. However, CMV in relation to other hematologic malignancies have not been studied. Here we present the largest investigation of CMV infection status pre-transplant as it relates to the risk of incident Hodgkin lymphoma, NHL, leukemia, and myeloma following solid organ transplantation in the US We identified that both positive CMV recipient serostatus and CMV sero-mismatch (R-/D+) were associated with significantly lower risks of DLBCL compared to CMV seronegative recipient/donor pairs. Our results suggest the risk of DLBCL differs according to recipient and donor EBV serostatus. Moreover, risk was modified by receipt of maintenance immunosuppression, in which cyclosporine or azathioprine were associated with significantly elevated risk among the R+ group.

There have been few comparable studies of hematologic malignancy and CMV infection in the context of solid organ transplantation. A retrospective cohort study by Opelz et al.[48] of 23,340 kidney, heart, and liver transplants was undertaken to assess the risk of NHL and recipient seropositivity for EBV or CMV. The authors found hospitalization for CMV disease during the first year post-transplant was associated with subsequent NHL (hazard ratio: 6.1, 95% CI: 2.0-18.4). However, there were no significant differences in lymphoma rates according to CMV positive or negative serostatus. The risk of PTLD, which includes NHL, and its association with CMV

infection post-transplant was assessed by Mañez et al.[47] among recipients who EBV seroconverted post-transplant. Of the 35 patients who seroconverted to EBV, CMV disease was observed in 54% (7/13) of patients who developed PTLD but in only 18% (4/22) of patients who did not develop PTLD (RR: 7.30, 95% CI: 2.36 - 22.61). However, neither study assessed lymphoma subtypes or other hematologic cancers.

We observed a significant inverse association between R-/D+ CMV serostatus and risk of DLBCL, which was contrary to our initial hypothesis that cancer incidence would be greatest among those at the highest risk of CMV infection or reactivation post-transplant. Two types of immune cells that are expanded during CMV infection have been implicated in CMV-induced protection against cancer: γδ T cells and a subset of natural killer (NK) cells.[133] During primary CMV infection, γδ T cells are significantly expanded.[134,135] *In vivo*, this expansion of γδ T cells has been associated with the viral clearance of CMV and reduced cancer occurrence or leukemia relapse risk in kidney transplant patients and allogenic stem cell transplant (ASCT) recipients, respectively.[136–141] In the setting of ASCT, a prospective study of 153 ALL and AML patients undergoing partially mismatched related ASCT showed improved leukemia-free survival and overall survival in patients with increased γδ T cell numbers in peripheral blood.[142] Like γδ T cells, NK cells have also exhibited anti-leukemic properties, specifically in the context of ASCT, in which CMV reactivation after transplantation is associated with decreased relapse risk.[143–148] It has been hypothesized that the effect is due to a shift in the composition of NK cell subsets, specifically driven by an increased proportion of $NKG2C^{pos}/NKG2A^{neg}$ NK cells, the "adaptive NK cell" subset. Reactivated CMV infections trigger expansions of $NKG2C^{pos}/NKG2A^{neg}$ NK cells, which enable

elimination and containment of CMV-infected cells that up-regulate HLA-E.[149–153] This CMV-induced accumulation of NKG2C^{pos}/NKG2A^{neg} NK cells is associated with a strong anti-leukemic and anti-myeloma effect that is proportionate with the magnitude of tumor cell HLA-E expression.[147,150] We thus hypothesize that it is plausible that CMV could reduce the risk of DLBCL but it remains unclear why we don't see this effect in the other lymphoma subtypes.

EBV infection is the most important predictor of PTLD, including DLBCL, among transplant recipients. We found when stratified by CMV and EBV risk groups, those who were CMV R-/D- and were EBV R-/D+ were significantly more likely to develop DLBCL compared to those who were EBV and CMV R-/D-. This finding is consistent with our expectation that the incidence of DLBCL would be elevated among those EBV R-/D+ post-transplant. However, when the recipient had R+ or R-/D+ CMV status, the association of EBV and DLBCL was no longer present.

Immunosuppression therapies appeared to modify the association between CMV risk post-transplant and subsequent DLBCL, particularly for maintenance immunosuppression regimens. We observed a significant interaction between maintenance therapy combinations used and CMV on risk of DLBCL. Recipients who were CMV positive at baseline who were taking cyclosporine and/or azathioprine had 94% greater risk of DLBCL compared to the low-risk CMV group. Azathioprine, an inhibitor of purine synthesis, has been associated with 3 times greater odds of cancer compared to MMF therapy. However in the same study, there was no significant difference between cyclosporine and tacrolimus on the odds of cancer occurrence.[154]

We did not observe any significant associations with CMV risk and the development of HL, leukemia, or myeloma. However, in the unadjusted model, the risk of HL was significantly lower among the R+ group compared to the R-/D- group, and risk remained lower after adjusting for confounders but was no longer significant. Considering the number of HL cases was low (n=61), this may have been an issue of power. Similarly, there were very few cases of leukemia (lymphoid n = 67; myeloid n=171) which may have contributed to our null finding. Specifically, our findings were not concordant with the small literature on the association between CMV and ALL.[78,79] However, as the confidence intervals around our estimate were wide and we saw only 16 cases of ALL, this also may have been an issue of power.

Strengths of this study include access to data from the SRTR and linked cancer registries for the ascertainment of transplants and cancers, resulting in this being the largest investigation of CMV and risk of hematologic malignancies to our knowledge. The large sample size allowed us to restrict our analysis to a clinically relevant time period for ascertainment of CMV donor and recipient serostatus. The use of transplant registry data linked to the cancer registry provided a unique cohort to examine cancer outcomes and exposures that would be otherwise costly or difficult to obtain data on, such as CMV infection. Another strength of this study is our findings are generalizable to solid organ transplants performed in the US Limitations of our study are primarily based on missing serology data for CMV post-transplantation. Also, follow-up CMV IgG antibody data was available for less than 10% of recipients in 24 months of follow up. Therefore, we had limited information on CMV seroconversion rates and how they corresponded to CMV risk groups in regard to distinguishing between primary and

reactivated infection. Additionally, we did not have data on CMV viral load during follow-up which could allow for better characterization of the role of CMV replication in the risk of DLBCL. EBV serostatus was also missing for 22.3% of recipients at baseline and therefore residual confounding may be possible.

Cancer and viral infections are two of the leading causes of morbidity among solid organ transplant recipients. CMV in particular has been implicated in increased risk of subsequent NHL post-transplant and ALL in *in utero* infection, and yet has a protective effect against AML relapse in HSCT and SOT with kidney transplantation. Given the findings of our study, it appears CMV may have both a pathogenic and protective role in DLBCL carcinogenesis, depending upon the context, specifically, the adjunctive use of immunomodulatory therapy. While this finding is unexpected, both *in vitro* and epidemiologic data have pointed to the immunomodulatory effects of CMV as a possible mechanism. Future work is needed to characterize the role of CMV infection and reactivation post-transplant and to further investigate the impact of maintenance immunotherapy regimens on the risk of DLBCL.

Table 3- 1. Characteristics of US solid organ transplant recipients, according to recipient and donor CMV serostatus

	CMV Status Pre-TX (Recipient/Donor Pairs)			
Recipient Characteristic	**R-/D-**	**R+**	**R-/D+**	**Total**
Total	41,518 (16.8%)	155,666 (62.9%)	50,134 (20.3%)	247,318 (100%)
Gender				
Male	27,743 (66.8%)	89,912 (57.8%)	33,763 (67.4%)	151,418 (61.2%)
Female	13,775 (33.2%)	65,754 (42.2%)	16,371 (32.7%)	95,900 (38.8%)
Age at Transplant, years				
0-17	4,937 (11.9%)	6,536 (4.2%)	5,458 (10.9%)	16,931 (6.9%)
18-34	6,561 (15.8%)	17,237 (11.1%)	7,417 (14.8%)	31,215 (12.6%)
35-49	11,503 (27.7%)	40,006 (25.7%)	13,694 (27.3%)	65,203 (26.4%)
50-64	14,646 (35.3%)	67,639 (43.4%)	18,405 (36.7%)	100,690 (40.7%)
65+	3,871 (9.3%)	24,248 (15.6%)	5,160 (10.3%)	33,279 (13.5%)
Median age, years	47	53	48	51
Race/Ethnicity				
NH White	31,264 (75.3%)	73,556 (47.3%)	34,842 (69.5%)	139,662 (56.5%)
NH Black	5,328 (12.8%)	35,704 (22.9%)	7,668 (15.3%)	48,700 (19.7%)
Hispanic	3,870 (9.3%)	31,373 (20.2%)	6,032 (12.0%)	41,275 (16.7%)
Asian/Pacific Islander	738 (1.8%)	13,104 (8.4%)	1,117 (2.2%)	14,959 (6.0%)
Other or Unknown	318 (0.8%)	1,929 (1.2%)	475 (1.0%)	2,722 (1.1%)
Transplanted Organ				
Kidney	25,959 (62.5%)	97,980 (62.9%)	27,842 (55.5%)	151,781 (61.4%)
Other/Multiple	9,611 (23.2%)	31,248 (20.1%)	13,651 (27.2%)	54,510 (22.0%)
Liver	5,948 (14.3%)	26,438 (17.0%)	8,641 (17.2%)	41,027 (16.6%)
Calendar Year of Transplant				
2000-2004	11,817 (28.5%)	43,301 (27.8%)	13,402 (26.7%)	68,520 (27.1%)
2005-2009	12,239 (29.5%)	48,330 (31.1%)	15,748 (31.4%)	76,317 (30.9%)
2010-2014	11,235 (27.1%)	42,053 (27.0%)	13,642 (27.2%)	66,930 (27.1%)
2015-2017	6,227 (15.0%)	21,982 (14.1%)	7,342 (14.6%)	35,551 (14.4%)
EBV Status Pre-Transplant				
Positive	25,814 (62.2%)	104,973 (67.4%)	30,928 (61.7%)	161,715 (65.4%)

Negative	7,709 (18.6%)	14,013 (9.0%)	8,839 (17.6%)	30,561 (12.4%)
Unknown	7,995 (19.3%)	36,680 (23.6%)	10,367 (20.7%)	55,042 (22.3%)
Education Status (for recipients >21 years)				
None	47 (0.1%)	864 (0.6%)	97 (0.2%)	1,008 (0.4%)
Grade school	621 (1.7%)	10,141 (6.9%)	1,005 (2.3%)	11,767 (5.2%)
High school/GED	12,235 (34.2%)	57,267 (38.9%)	15,523 (35.5%)	85,025 (37.5%)
Attended college/technical	8,855 (24.7%)	31,869 (21.6%)	10,909 (24.9%)	51,633 (22.8%)
Associate/Bachelors	7,211 (20.1%)	20,633 (14.0%)	8,060 (18.4%)	35,904 (15.8%)
Post-graduate	3,181 (8.9%)	8,497 (5.8%)	3,410 (7.8%)	15,088 (6.6%)
Unknown	3,644 (10.2%)	18,066 (12.3%)	4,730 (10.8%)	26,440 (11.7%)
Yost SES Quintile				
1: Lowest	5,711 (13.8%)	32,573 (20.9%)	7,952 (15.9%)	46,236 (18.7%)
2: Low	7,192 (17.3%)	30,145 (19.4%)	8,976 (17.9%)	46,313 (18.7%)
3: Mid	7,819 (18.3%)	30,676 (19.7%)	9,652 (19.3%)	48,147 (19.5%)
4: High	8,885 (21.4%)	29,127 (18.7%)	10,480 (20.9%)	48,492 (19.6%)
5: Highest	9,361 (22.6%)	26,570 (17.1%)	10,527 (21.0%)	46,458 (18.8%)
Unknown	2,550 (6.1%)	6,575 (4.2%)	2,547 (5.1%)	11,672 (4.7%)

All entries are N (%) unless otherwise noted. All percentages are column percentages except for totals (row percentages).

Table 3- 2. Risk of hematologic malignancies by CMV recipient/donor serostatus pre-transplantation.

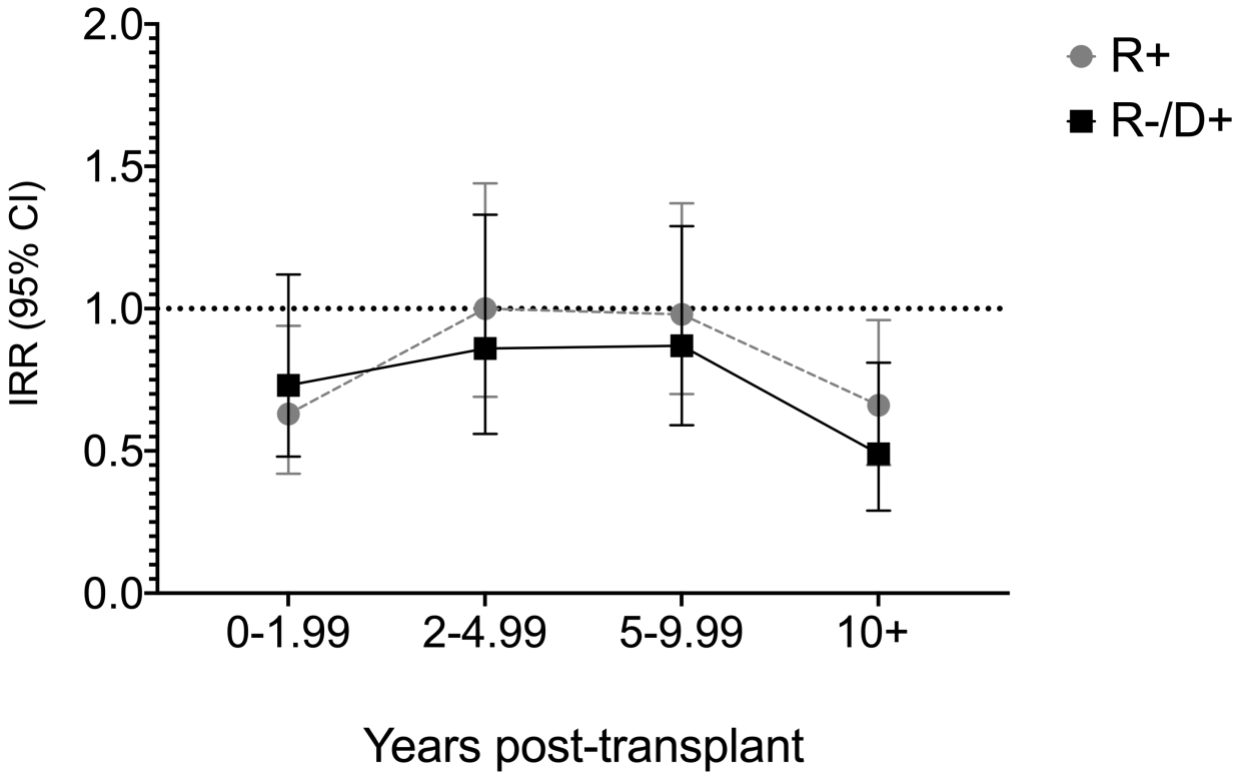
	CMV Recipient/Donor Status Post-Transplant									
	Total		**R-/D- (Referent)**		**R+**			**R-/D+**		
Cancer Group	N	IR	N	IR	N	IR	Adj. IRR (95% CI)	N	IR	Adj. IRR (95% CI)
Hodgkin lymphoma	61	4.9	19	8.8	21	2.69	0.47 (0.19 - 1.18)	21	8.49	0.75 (0.29 - 1.96)
Non-Hodgkin lymphoma, NOS	270	21.7	48	22.2	164	21	1.36 (0.66 - 2.80)	58	23.5	1.56 (0.74 - 3.30)
Diffuse large B-cell lymphoma	1193	95.8	267	123.6	680	86.9	0.83 (0.69 - 1.00)	246	99.5	0.74 (0.59 - 0.91)
Burkitt lymphoma	101	8.1	22	10.2	53	6.8	1.36 (0.66 - 2.80)	26	10.5	1.56 (0.74 - 3.30)
Follicular Lymphoma	39	3.1	9	4.2	21	2.7	0.94 (0.29 - 3.01)	9	3.6	1.48 (0.43 - 5.09)
Lymphoplasmacytic Lymphoma	12	1.0	1	0.5	11	1.4	-	0	-	-
Mantle Cell	10	0.8	2	0.9	7	0.9	-	1	0.4	-
Marginal Zone	61	4.9	11	5.1	39	5.0	0.79 (0.34 - 1.82)	11	4.4	0.53 (0.17 - 1.64)
Peripheral T-cell Lymphoma	49	3.9	10	4.6	26	3.3	0.71 (0.28 - 1.80)	13	5.3	0.98 (0.35 - 2.73)
ALCL	23	1.8	5	2.3	10	1.3	0.70 (0.13 - 3.73)	8	3.2	1.62 (0.30 - 8.90)
Mycosis Fungoides/Sézary's Syndrome	16	1.3	3	1.4	9	1.2	0.92 (0.17 - 4.88)	4	1.6	0.82 (0.11 - 5.87)
Pre-cursor B- or T-Cell Lymphoblastic Leukemia/Lymphoma	12	1.0	4	1.9	6	0.8	0.60 (0.13 - 2.71)	2	0.8	0.23 (0.02 - 2.19)
Lymphoid leukemia, total	67	5.4	14	6.5	40	5.1	0.97 (0.42 - 2.24)	13	5.3	0.65 (0.23 - 1.89)
Acute Lymphocytic	15	1.2	4	1.9	11	1.4	0.95 (0.23 - 3.85)	0	-	-

Chronic Lymphocytic/Small Lymphocytic Lymphoma	52	4.2	10	4.6	29	3.7	0.98 (0.35 0 2.79)	13	5.3	1.06 (0.32 - 3.49)
Myeloid leukemia, total	171	13.7	25	11.6	116	14.8	1.39 (0.79 - 2.43)	30	12.1	1.08 (0.56 - 2.07)
Acute Myeloid	112	9.0	13	6.0	76	9.7	1.83 (0.84 - 3.95)	23	9.3	1.58 (0.67 - 3.70)
Chronic Myeloid	59	4.7	12	5.6	40	5.1	0.93 (0.41 - 2.14)	7	2.8	0.54 (0.18 - 1.66)
Other leukemia	18	1.4	3	1.4	11	1.4	3.02 (0.37 - 24.93)	4	1.6	1.72 (0.16 - 19.03)
Aleukemic, subleukemic and NOS	20	1.6	4	1.9	10	1.3	0.55 (0.16 - 1.92)	6	2.4	0.76 (0.19 - 3.06)
Myeloma	216	17.3	32	14.8	145	18.5	0.94 (0.59 - 1.51)	39	15.8	0.82 (0.46 - 1.46)

Incidence rates are per 100,000 person-years. Adjusted IRR models include adjustment for recipient sex, age (0-17, 18-34, 35-49, 50-64, 65+ years), race/ethnicity, organ type (kidney, liver, other/multiple), EBV recipient/donor status (EBV R-/D-, EBV R+, EBV R-/D+), and SES quintile. Significant associations are underlined.

Abbreviations: ALCL – anaplastic large cell lymphoma; ALL – acute lymphoblastic leukemia; AML – acute myeloid leukemia; CI - confidence interval; CLL/SLL - chronic lymphocytic leukemia / small lymphocytic lymphoma; CML – chronic myeloid leukemia ; DLBCL – diffuse B cell lymphoma; IR- incidence rate; IRR – incidence rate ratio; NOS- not otherwise specified.

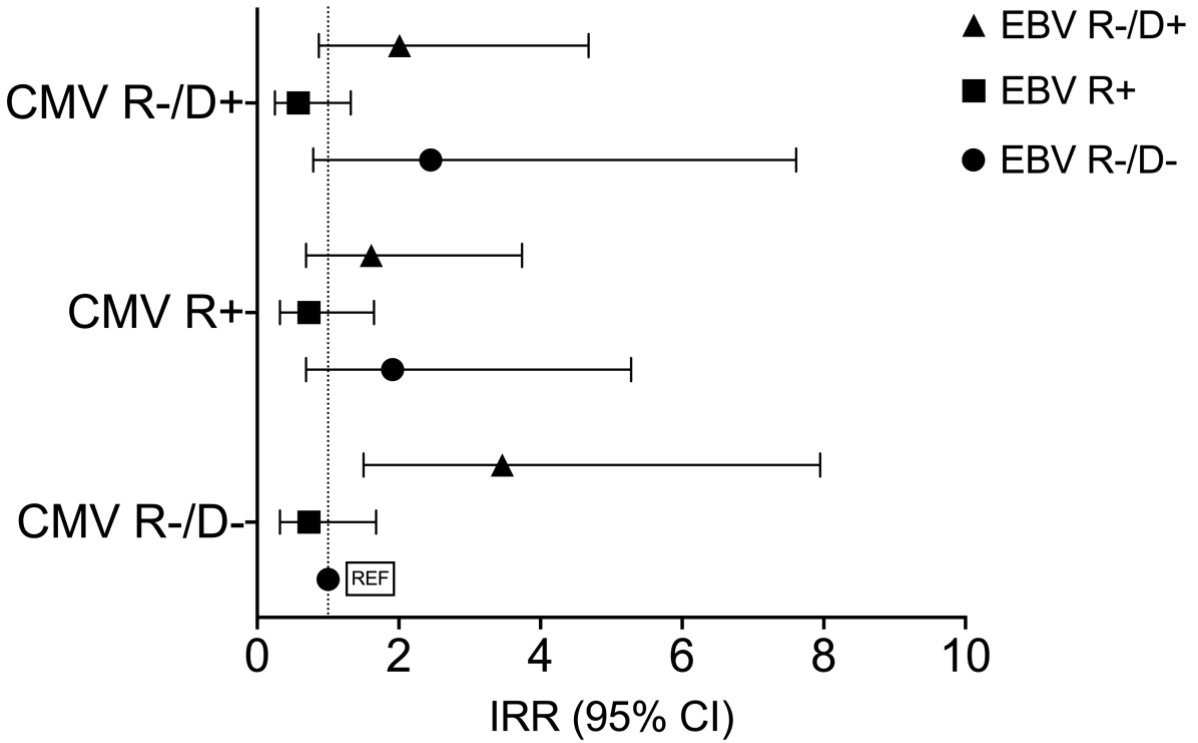
Figure 3- 1. Associations of diffuse large B-cell lymphoma with CMV risk group as a function of time since transplantation.



Associations of diffuse large B-cell lymphoma with CMV risk group as a function of time since transplantation.

Incidence rate ratios and 95% confidence intervals of CMV risk groups compared to R-/D- group. Abbreviations: CI – confidence interval; IRR – incidence rate ratio.

Figure 3- 2. Interaction between CMV and EBV risk groups and the risk of DLBCL.



The model is adjusted for recipient sex, age (0-17, 18-34, 35-49, 50-64, 65+ years), race/ethnicity, organ type (kidney, liver, other/multiple), and SES quintile. Abbreviations:CI – confidence interval; EBV – Epstein-Barr virus; IRR – incidence rate ratio; REF – reference;

Table 3- 3. Interaction between CMV and EBV on the risk of DLBCL post-transplant.

CMV status	EBV status	N	IR	Adjusted IRR (95% CI)
R-/D-	R-/D-	7	124.9	1 (Reference)
	R+	103	82.6	0.73 (0.32-1.68)
	R-/D+	77	419.5	3.46 (1.50-7.95)
R+	R-/D-	10	190.0	1.91 (0.69-5.28)
	R+	379	77.0	0.73 (0.32-1.65)
	R-/D+	68	201.1	1.61 (0.69-3.74)
R-/D+	R-/D-	6	249.9	2.45 (0.79-7.61)
	R+	100	70.9	0.58 (0.25-1.32)
	R-/D+	58	268.0	2.01 (0.87-4.68)

Likelihood ratio: 19.61; p=0.0006. IRs per 100,000 person-years. Adjusted IRR models include adjustment for recipient sex, age (0-17, 18-34, 35-49, 50-64, 65+), race/ethnicity, organ type (kidney, liver, other/multiple), and SES quintile.

Table 3- 4. Distribution of induction and maintenance immunosuppression medications by CMV recipient/donor status pre-transplant.

Medication	CMV Status Pre-TX (Recipient/Donor Pairs)			
Induction regiment, n (%)	**R-/D-**	**R+**	**R-/D+**	**Total**
Any induction therapy	33,553 (80.8%)	127,342 (81.8%)	39,975 (79.7%)	200,870 (81.2%)
Polyclonal antibody	12,998 (31.3%)	51,057 (32.8%)	15,117 (30.2%)	79,172 (32.1%)
Monoclonal antibody	223 (0.5%)	832 (0.5%)	274 (0.6%)	1,329 (0.5%)
IL2 receptor antagonists	11,681 (28.1%)	43,297 (27.8%)	14,358 (28.6%)	69,336 (28.0%)
Campath	3,034 (7.3%)	10,130 (6.5%)	3,156 (6.3%)	16,320 (6.6%)
Rituximab	161 (0.4%)	817 (0.5%)	191 (0.4%)	1,169 (0.5%)
Steroids*	26,287 (63.3%)	101,768 (65.4%)	31,659 (63.2%)	159,714 (64.6%)
Maintenance immunosuppression, n (%)				
Only Tacrolimus or MMF	34,023 (82.0%)	129,657 (83.3%)	40,993 (81.8%)	204,673 (82.8%)
Only Cyclosporine or Azathioprine	1,645 (4.0%)	5,131 (3.3%)	2,064 (4.1%)	8,840 (3.6%)
Other combination	5,850 (14.1%)	20,878 (13.4%)	7,077 (14.1%)	33,805 (13.7%)
mTOR inhibitor	2,998 (7.2%)	9,190 (5.9%)	3,265 (6.5%)	15,453 (6.3%)
Steroids*	31,444 (74.7%)	122,479 (78.7%)	39,228 (78.3%)	193,151 (78.1%)

All entries are N (%) unless otherwise noted. All percentages are column percentages. Steroid use was in combination with other induction or maintenance regiments. Abbreviations: MMF – mycophenolate mofetil.

Table 3- 5. Interaction between induction or maintenance medications and CMV on the risk of DLBCL post-transplantation.

Medications		**CMV Status Pre-TX (Recipient/Donor Pairs)**		
Induction Medications				
Any induction		R-/D- (REF)	R+	R-/D+
LR: 0.48; p=0.79	Y	1	1.03 (0.71 - 1.50)	0.94 (0.63 - 1.39)
	N	1	0.87 (0.57 - 1.33)	0.67 (0.40 - 1.12)
Polyclonal				
LR: 0.25; p=0.88	Y	1	0.86 (0.66 - 1.11)	0.80 (0.57 - 1.12)
	N	1	0.86 (0.69 - 1.06)	0.74 (0.57 - 0.96)
Monoclonal				
LR: 1.61; p=0.45	Y	1	2.32 (1.02 - 5.24)	3.28 (1.05 - 10.30)
	N	1	0.83 (0.69 - 0.99)	0.73 (0.58 - 0.90)
IL2 receptor antagonists				
LR: 0.73; p=0.69	Y	1	0.69 (0.53 - 0.90)	0.55 (0.38 - 0.79)
	N	1	0.85 (0.69 - 1.05)	0.78 (0.61 - 1.00)
Campath				
LR: 0.21; p=0.90	Y	1	1.06 (0.71 - 1.56)	0.97 (0.53 - 1.79)
	N	1	0.83 (0.68 - 1.00)	0.73 (0.58 - 0.91)
Rituximab				
LR: 1.89; p=0.39	Y	1	1.02 (0.38 - 2.75)	-
	N	1	0.83 (0.69 - 1.00)	0.74 (0.60 - 0.92)
Steroids				
LR: 0.86; p=0.65	Y	1	0.87 (0.66 - 1.15)	0.82 (0.60 - 1.11)
	N	1	0.77 (0.57 - 1.05)	0.64 (0.44 - 0.93)
Maintenance Medications				
Combination of Medications	Only Tacrolimus or MMF	1	0.72 (0.59 - 0.88)	0.69 (0.55 - 0.87)
LR: 13.64 ; p=0.009	Only Cyclosporine or Azathioprine	1	1.94 (1.33 - 2.84)	1.09 (0.56 - 2.14)
	Other combination	1	0.98 (0.73 - 1.30)	0.65 (0.40 - 1.04)
mTOR inhibitor				
LR: 0.18 ; p=0.92	Y	1	1.01 (0.67 - 1.53)	1.02 (0.55 - 1.88)

	N	1	0.84 (0.69 - 1.01)	0.73 (0.58 - 0.91)
Steroids				
LR: 3.64 ; p=0.16	Y	1	0.83 (0.60 - 1.14)	0.67 (0.47 - 0.95)
	N	1	0.91 (0.64 - 1.31)	1.03 (0.68 - 1.55)

Models adjusted for recipient's sex, categorical age, race/ethnicity, organ type, EBV recipient/donor status, and SES quintile. Abbreviations: LR – likelihood ratio;

Chapter 4: Solid tumor malignancies associated with cytomegalovirus infection among solid organ transplant recipients in the United States

Introduction

The field of solid organ transplantation (SOT) has made significant strides since the 1980s as a treatment for end-stage organ disease. However, long-term use of immunosuppressive medications to prevent rejection remains a source of morbidity among transplant recipients. In particular, opportunistic viral infections following transplantation are a cause of post-transplant malignancy and a threat to long-term graft survival.[125,131,132]

Cytomegalovirus (CMV) is among the most common viral infections following solid organ transplantation. CMV is a member of the herpesvirus family and establishes lifelong latency and can undergo periodic reactivation. In immunocompetent individuals, CMV infection is typically asymptomatic.[1–4] However, in transplant recipients, CMV infection can cause severe disease and allograft rejection. The occurrence of CMV infection varies according to the organ transplanted, but it is estimated that the incidence is up to 32% after kidney transplantation and 22-29% following liver transplantation.[42,155,156] Primary infection is most likely to occur by transmission from a CMV seropositive donor to a seronegative recipient at the time of transplant, but community and household transmission following transplantation has also been reported.[157–160] Alternatively, immunosuppressive therapy required for transplantation may cause CMV reactivation due to depressed immune surveillance.[43,44]

The long-term use of immunosuppressive agents to prevent allograft rejection increases the risk of cancer. Among US solid organ transplant recipients, the risk of any type of cancer is nearly 2 times greater than that of the general population.[125,161] CMV may also contribute to cancer in SOT recipients. In a small, retrospective cohort study of liver transplant recipients who seroconverted to Epstein-Barr virus (EBV), CMV disease was a significant predictor of the development of PTLD (RR: 7.3, 95% CI: 2.36-22.6) .[47] In another study, hospitalization for CMV disease during the first year post-transplant has also been associated with six-fold higher (HR: 6.1, 95% CI: 1.1-11.3) subsequent risk of non-Hodgkin lymphoma (NHL).[48] However, there was no significant difference in lymphoma rates according to CMV serostatus.

CMV infection as it relates to other cancers post-transplantation is not well-defined. Among patients in the general population, CMV nucleic acids and proteins have been detected in tumor samples of breast, colon, and prostate cancer as well as glioblastoma.[73–75] An investigation of breast cancer patients found 100% of primary breast cancer samples were CMV positive and that virus positivity was restricted to tumor cells in established tumors and metaseses.[162] Similarly, the role for CMV in glioblastoma pathogenesis has been proposed in several studies, the most notable being by Cobbs *et al*., in which investigators detected CMV nucleic acids and proteins in 100% (n=27) of low- and high-grade malignant gliomas.[163] However, temporality is of issue in these studies since the timing of CMV infection cannot be established. A role for CMV has also been hypothesized in the development of pediatric neuroblastoma and Wilm's tumor; in a small study of pediatric brain tumor cases, 53% of Wilm's tumor patients and 53% of neuroblastoma patients were found to have complement-fixing antibodies against

CMV.[77] Overall, these findings support the hypothesis that CMV may exhibit oncomodulatory effects that interfere with cellular properties, thereby promoting tumor progression.[61,74,164]

CMV infection is common among SOT recipients, and cancer is a major adverse outcome. [132] We therefore examined the potential association between CMV infection post-transplant and the risk of malignancy among solid organ transplant recipients using data from the Transplant Cancer Match Study (TCM).

Methods

The Transplant Cancer Match Study (http://transplantmatch.cancer.gov) has been previously described in detail.[125] Briefly, the study links data from the Scientific Registry of Transplant Recipients (SRTR), which includes all US transplants since 1987, with 32 US state and regional cancer registries. We included recipients of a first organ transplant who resided in a region covered by one of the participating cancer registries at the time of transplantation. Data on CMV IgG serostatus at the time of transplant were unavailable for most recipients (69%) before 2000, therefore, we restricted analysis to transplants in 2000 and in subsequent years. The study was considered non-human subjects research by the National Cancer Institute and was approved, as required, by participating cancer registries.

Of the 672,603 individuals in the US who received a first transplant during 1987 – 2017, we excluded 211,973 (31.5%) because they were transplanted outside of participating cancer registry regions, 117,404 (17.5%) because they were transplanted before the year 2000 or lacked follow-up, 29,875 (4.4%) because they had a cancer

diagnosis before transplantation, and 1,021 (0.2%) because they had human immunodeficiency virus infection, and 106 (0.02%) if they had more than one donor. Finally, we excluded an additional 19,139 (2.8%) transplant recipients who were missing CMV IgG serostatus pre-transplant, or if they were seronegative, their donor was missing CMV serostatus. After these exclusions, 247,318 (36.8%) transplants were included in the final analysis.

Incident cancers after transplantation were identified from the 32 linked population-based cancer registries and classified using a modified version of the Surveillance, Epidemiology, and End Results (SEER) program "site recode." [127,128] Cancers with similar site (e.g. ureters and urinary bladder) were grouped when the specific subtype had fewer than 60 cases. Tumors with fewer than 60 cases that could not be grouped (e.g. bones and joints, mesothelioma, eye and orbit) were included in a miscellaneous group.

CMV status was divided into three risk groups according to SRTR data on pretransplant serostatus of donors and recipients, to reflect the risk of active CMV infection and disease post-transplant: high risk (recipient seronegative and donor seropositive [R-/D+]), intermediate risk (recipient seropositive and donor seropositive or seronegative [R+]), and low risk (recipient seronegative and donor seronegative [R-/D-]).[39,165] Other data were obtained from the SRTR regarding recipient characteristics (age at transplantation, sex, race/ethnicity), transplant characteristics (organ, calendar year), and immunosuppression medications (induction and baseline immunosuppression maintenance therapies). EBV serostatus was categorized into three risk groups as described for CMV above.

Socioeconomic status (SES) quintiles were constructed utilizing recipients' ZIP code at the time of transplantation using methods developed by Yost et al.[130] Briefly, SES index was assessed for every ZIP Code Tabulation Area (ZCTA) using American Community Survey data on seven measures of SES: percent working class and unemployed, education index, percent living below 150% of the poverty line, median household income, median home value, and median rent. ZCTAs were categorized into quintiles.

Follow-up for cancer started at the time of transplantation and ended at the earliest of death, graft failure, retransplantation, loss to follow-up by SRTR, or end of cancer registry coverage. We calculated crude incidence rates, defined as the number of outcomes per 100,000 person-years of follow-up, for each cancer for recipients in each CMV risk group. To compare the risk by CMV risk group, we estimated incidence rate ratios (IRRs) with 95% confidence intervals using multivariable Poisson regression models adjusted for sex, age at transplantation, race/ethnicity, SES quintile, organ (kidney, liver or other), and EBV risk group. Included variables in the models were evaluated by backward elimination and likelihood ratio test of the nested models. An offset term of the log of the follow-up time was included in the models to account for differences in exposure period. Goodness of fit and overdispersion of the models were assessed by deviance goodness-of-fit Chi-square and kurtosis score, respectively. Reported *P* values were two-sided. Because our analysis was exploratory, we did not correct for multiple testing and p-values less than 0.05 are considered statistically significant. Stata/MP version 16.1 (StataCorp LP, College Station, Texas) was used for all statistical analyses.

Results

A total of 247,318 solid organ transplant recipients with 1,245,369 person-years at risk were identified. **Table 4- 1** describes the demographic and transplant characteristics of these individuals by CMV risk group. Overall, 62.9% of the cohort was CMV seropositive pretransplant, while 20.3% was CMV seronegative with a seropositive donor (R-/D+), and 16.8% was seronegative with a seronegative donor (R-/D-). The R+ group was more likely to be female (42.2%) than the R-/D- (33.2%) and R-/D+ (32.7%) groups ($p<0.001$) and was also older (median age: 53 years) than the seronegative recipient groups (R-/D-: 47 years; R-/D+: 48 years) ($p<0.001$). The greatest racial/ethnic diversity was among the R+ group, of which more than half (51.5%) identified as a racial/ethnic group other than non-Hispanic white. Recipients who were CMV seropositive were also likely to be EBV seropositive (67.4%), and this proportion was slightly higher than among R-/D- (62.2%) and R-/D+ (61.7%) groups ($p<0.001$). Kidneys were the most prevalent organ transplanted across all three groups, representing 61.4% of transplants, overall. Highest attained education status different somewhat by CMV group, with a higher proportion of individuals in the R-/D- (29.0%) and R-/D+ (26.2%) groups having received either an associate's or bachelor's degree compared to the R+ group (19.8%). The CMV seronegative groups had a higher proportion of individuals in the high and highest SES quintiles (R-/D-: 44.0%; R-/D+: 41.9%) than the R+ group (35.8%) ($p<0.001$). Overall, 81.2% of recipients received some form of induction therapy. For maintenance immunosuppressive therapies, 81.2% of recipients received tacrolimus

and/or mycophenolate mofetil (MMF), 3.6% received cyclosporine or azathioprine, and 13.7% were given some other combination of these medications.

Overall, we identified 11,831 incident malignant solid tumors. The most common were cancers of the lung (16.3%), prostate (13.3%), and kidney (11.1%). By and large, incidence was not significantly elevated in the CMV R+ and R-/D+ groups as compared to the R-/D- group for most cancers (**Table 4- 2)**. However, incidence of lung cancer was found to be 24% higher among the R+ group compared to the R-/D- group (adjusted IRR [aIRR]: 1.24, 95% CI: 1.05-1.46), but there was no statistical difference among the R-/D+ group (0.94, 0.77-1.14). In contrast, the R-/D+ group had significantly lower incidence of small intestine cancer (aIRR: 0.23, 95% CI: 0.09-0.63) compared to the low-risk group. The R+ group shared this inverse association, but this association was not significant (aIRR: 0.65, 95% CI: 0.37 – 1.16).

Table 4- 3 presents the stratified analysis by recipient age at transplant and sex for select cancers of interest. For both lung and small intestine cancers, there was not sufficient data to examine risk in recipients ≤34 years of age. However, one notable finding was among recipients aged 50-64 years, the risk of small intestine cancer was 86% lower in the R-/D+ group compared to the R-/D- group (aIRR: 0.14, 95% CI: 0.03 – 0.64). The risk of lung cancer was elevated in this age group within the R+ group compared to the R-/D- group (aIRR: 1.26, 95% CI: 1.03 – 1.55). When stratified by sex, CMV R+ females had a 68% reduction of risk of small intestine cancer compared to females in the R-/D- group (aIRR: 0.32, 95% CI: 0.11 – 0.87), while within the R-/D+ group males experienced 78% reduction of risk compared to the R-/D- group (0.22, 0.06 – 0.78). The risk of lung cancer, however, was elevated within R+ males (aIRR: 1.21,

95% CI: 1.01 – 1.47). The risk of bladder cancer was also elevated within both R+ and R-/D+ groups (aIRR: 1.81, 95% CI: 1.05 – 3.12; 2.05, 1.14 – 3.68, respectively) but this difference was not seen in females. We did not find any significant differences across age or sex for brain, breast, or prostate cancers.

Discussion

Cytomegalovirus is among the most common viral infections following solid organ transplantation. While CMV is generally not regarded to be an oncogenic virus, CMV infection has been implicated in cancer. However, CMV in relation to cancers that occur post-transplantation is not well-defined. Here we present the largest investigation of CMV infection as it relates to the risk of solid tumor malignancies in the US, specifically following solid organ transplantation. Overall, the incidence of solid tumor cancers did not differ by CMV R+ or R-/D+ serostatus compared to CMV R-/D-. However, we identified R-/D+ CMV was inversely associated with small intestine cancer compared to the R-/D- CMV group. Additionally, we found among the R+ group an elevated risk of lung cancer compared to the R-/D- group. For both cancers, these differences were most pronounced among males and recipients aged 50-64 years. We also identified an elevated risk of bladder cancer among males in both the R+ and R-/D+ CMV groups compared to the R-/D- group.

The role of CMV in carcinogenesis is understudied, however, several studies exploring CMV and cancer in the setting of transplantation have been conducted yielding inconclusive results. In a small, retrospective cohort study assessing CMV exposure and cancer occurrence among 455 kidney transplant recipients, patients who had positive

CMV serology pre-transplant (HR: 1.83; 95% CI: 1.17 – 2.88) and post-transplant (HR: 2.17; 95% CI: 1.02 – 4.59) had significantly greater risk of any cancer compared to CMV unexposed recipients. However, no cases of small intestine cancer were observed. Lung cancer was marginally more frequent in CMV-exposed patients (3.3% vs 0.5%), but was not statistically significant (p=0.092).[166] A study by Desai *et al* [167]. of nearly 23,000 solid organ recipients in the U.K. Transplant Registry found no statistical difference of overall cancer incidence by CMV risk groups but found a two-fold higher risk of kidney cancer among kidney transplant recipients (HR: 2.33, p=0.036). In contrast, a case-control study by Couzi *et al.*[154] reported the odds of post-transplant cancer was 5.28 times greater among CMV-negative recipients compared to recipients exposed to CMV before or after transplantation (p=0.006).

From this large cohort of solid organ transplant recipients, we demonstrated that CMV has no independent association with the risk of cancer post-transplant for the majority of the 26 types of solid tumors we assessed. In the univariate analysis, the incidence of cancer post-transplant was generally higher among those CMV seropositive or CMV seronegative with a seropositive donor compared to seronegative donor and recipient group, but this difference was not statistically significant in the multivariate analysis after adjusting for known confounders, particularly EBV recipient and donor status and SES.

In our study, we observed a 24% greater risk of lung cancer among the R+ group that was not seen among the R-/D+ group. In the general population, CMV has not been studied in lung cancer before, however, correlation has been found between lung cancer relapse time and high CMV DNA concentration in lung cancer tissue.[168] Interestingly, we

also observed a moderately elevated risk of bladder cancer among men in both moderate- and high-risk CMV groups that was not seen in the unstratified analysis. As smoking is the most important risk factor for both bladder and lung cancers[169], the associations may be due to residual confounding since we did not have data on smoking status or tobacco use.

The inverse association between CMV risk and small intestine cancer was unexpected. Meta-analysis of several smaller studies investigating the prevalence of CMV in tumor tissues of colorectal cancer samples have suggested a significantly higher prevalence of virus in these tissues than normal tissues (OR: 6.59, 95% CI: 4.48 -9.69).[170] Similarly, CMV infection has also shown to be significantly associated with increased risk of gastrointestinal cancer, which includes colorectal cancer.[171,172] However, CMV has not been specifically associated with small intestine cancer. In the general population, small intestine cancer is rare; the annual incidence is only 2.4 cases per 100,000 men and women per year.[173] Transplant recipients have an elevated risk of small intestine cancer (SIR: 2.43, 95% CI: 1.80 – 3.20)[125] but has previously been attributed to long-term immunosuppression. However, our finding is most consistent with Couzi *et al.*[154] who demonstrated that reduced incidence of cancer was associated with elevated blood γδ T cells following CMV infection. *In vitro,* γδ T cells are capable of killing myeloma and carcinoma cell lines.[174] Thus, it is plausible that γδ T cells may play a role in conferring protection against small intestine cancer following CMV infection, although it is unclear why we did not see this effect for other cancers.

There were several strengths of our study compared to others on CMV infection and cancer risk post-transplantation. First, the large size of the Transplant Cancer Match

Study enabled us to examine CMV and cancer risk in nearly 250,000 solid organ transplant recipients, an order of magnitude larger than the largest similar study. This also makes ours the largest study of CMV and cancer risk in any setting to our knowledge. Our study was nearly ten times larger than the U.K. study by Desai and colleagues, which allowed for more precise estimates of risk for less common cancers. Second, our study population is a representative sample of the US transplant population and thus, our results are generalizable. Another strength of our study is the use of the transplant registry data linked to the cancer registry provided a unique cohort to examine exposures and cancer outcomes that would be otherwise costly or difficult to obtain, such as CMV infection. Lastly, the addition of Yost index data to our models resulted in significant adjustment of confounding, by which we found SES was independently associated with both CMV status and cancer risk. Limitations of our study are primarily based on missing serology data for CMV post-transplantation. CMV IgG antibody data was available for less than 10% of recipients in the 24 months of follow-up and therefore not included in our analysis. Therefore, we had limited information on how CMV risk groups corresponded to seroconversion rates during follow-up. Another limitation of our study, as previously mentioned, was the lack of data on smoking and tobacco use or other major carcinogenic exposures. We did not have any surrogate variables in our dataset to be able to control for smoking (apart from SES), therefore it is possible there may be residual confounding.

Overall, while this large study of risk of cancer by CMV infection risk following solid organ transplantation did not yield significant findings for the majority of solid tumors, we did observe CMV was associated with an elevated risk of lung cancer and a

decreased risk of small intestine cancer. However, the mechanisms associated with these findings remain unclear. Further research is needed to understand the immunologic profile following CMV infection post-transplantation.

Table 4- 1. Characteristics of US solid organ transplant recipients, by pretransplant donor and recipient CMV serostatus.

	CMV status (recipient/donor)			
Recipient *c*haracteristic	**R-/D-**	**R+**	**R-/D+**	**Total**
Total	41,518 (16.8%)	155,666 (62.9%)	50,134 (20.3%)	247,318 (100%)
Sex				
Male	27,743 (66.8%)	89,912 (57.8%)	33,763 (67.4%)	151,418 (61.2%)
Female	13,775 (33.2%)	65,754 (42.2%)	16,371 (32.7%)	95,900 (38.8%)
Age at transplant, years				
0-17	4,937 (11.9%)	6,536 (4.2%)	5,458 (10.9%)	16,931 (6.9%)
18-34	6,561 (15.8%)	17,237 (11.1%)	7,417 (14.8%)	31,215 (12.6%)
35-49	11,503 (27.7%)	40,006 (25.7%)	13,694 (27.3%)	65,203 (26.4%)
50-64	14,646 (35.3%)	67,639 (43.4%)	18,405 (36.7%)	100,690 (40.7%)
65+	3,871 (9.3%)	24,248 (15.6%)	5,160 (10.3%)	33,279 (13.5%)
Median age, years	47	53	48	51
Race/Ethnicity				
NH White	31,264 (75.3%)	73,556 (47.3%)	34,842 (69.5%)	139,662 (56.5%)
NH Black	5,328 (12.8%)	35,704 (22.9%)	7,668 (15.3%)	48,700 (19.7%)
Hispanic	3,870 (9.3%)	31,373 (20.2%)	6,032 (12.0%)	41,275 (16.7%)
Asian/Pacific Islander	738 (1.8%)	13,104 (8.4%)	1,117 (2.2%)	14,959 (6.0%)
Other or Unknown	318 (0.8%)	1,929 (1.2%)	475 (1.0%)	2,722 (1.1%)
Transplanted *o*rgan				
Kidney	25,959 (62.5%)	97,980 (62.9%)	27,842 (55.5%)	151,781 (61.4%)
Liver	5,948 (14.3%)	26,438 (17.0%)	8,641 (17.2%)	41,027 (16.6%)
Other/multiple	9,611 (23.2%)	31,248 (20.1%)	13,651 (27.2%)	54,510 (22.0%)
Calendar *y*ear of transplant				

2000-2004	11,817 (28.5%)	43,301 (27.8%)	13,402 (26.7%)	68,520 (27.1%)
2005-2009	12,239 (29.5%)	48,330 (31.1%)	15,748 (31.4%)	76,317 (30.9%)
2010-2014	11,235 (27.1%)	42,053 (27.0%)	13,642 (27.2%)	66,930 (27.1%)
2015-2017	6,227 (15.0%)	21,982 (14.1%)	7,342 (14.6%)	35,551 (14.4%)
EBV status pretransplant				
Positive	25,814 (62.2%)	104,973 (67.4%)	30,928 (61.7%)	161,715 (65.4%)
Negative	7,709 (18.6%)	14,013 (9.0%)	8,839 (17.6%)	30,561 (12.4%)
Unknown	7,995 (19.3%)	36,680 (23.6%)	10,367 (20.7%)	55,042 (22.3%)
Education status (for recipients >21 years old)				
None	47 (0.1%)	864 (0.6%)	97 (0.2%)	1,008 (0.4%)
Grade school	621 (1.7%)	10,141 (6.9%)	1,005 (2.3%)	11,767 (5.2%)
High school/GED	12,235 (34.2%)	57,267 (38.9%)	15,523 (35.5%)	85,025 (37.5%)
Attended college / technical	8,855 (24.7%)	31,869 (21.6%)	10,909 (24.9%)	51,633 (22.8%)
Associate/Bachelors	7,211 (20.1%)	20,633 (14.0%)	8,060 (18.4%)	35,904 (15.8%)
Post-graduate	3,181 (8.9%)	8,497 (5.8%)	3,410 (7.8%)	15,088 (6.6%)
Unknown	3,644 (10.2%)	18,066 (12.3%)	4,730 (10.8%)	26,440 (11.7%)
Yost SES Quintile				
1: Lowest	5,711 (13.8%)	32,573 (20.9%)	7,952 (15.9%)	46,236 (18.7%)
2: Low	7,192 (17.3%)	30,145 (19.4%)	8,976 (17.9%)	46,313 (18.7%)
3: Mid	7,819 (18.3%)	30,676 (19.7%)	9,652 (19.3%)	48,147 (19.5%)
4: High	8,885 (21.4%)	29,127 (18.7%)	10,480 (20.9%)	48,492 (19.6%)
5: Highest	9,361 (22.6%)	26,570 (17.1%)	10,527 (21.0%)	46,458 (18.8%)
Unknown	2,550 (6.1%)	6,575 (4.2%)	2,547 (5.1%)	11,672 (4.7%)
Induction regiment				
Any induction therapy	33,553 (80.8%)	127,342 (81.8%)	39,975 (79.7%)	200,870 (81.2%)

Polyclonal antibody	12,998 (31.3%)	51,057 (32.8%)	15,117 (30.2%)	79,172 (32.1%)
Monoclonal antibody	223 (0.5%)	832 (0.5%)	274 (0.6%)	1,329 (0.5%)
IL2 receptor antagonist	11,681 (28.1%)	43,297 (27.8%)	14,358 (28.6%)	69,336 (28.0%)
Campath	3,034 (7.3%)	10,130 (6.5%)	3,156 (6.3%)	16,320 (6.6%)
Rituximab	161 (0.4%)	817 (0.5%)	191 (0.4%)	1,169 (0.5%)
Steroids	26,287 (63.3%)	101,768 (65.4%)	31,659 (63.2%)	159,714 (64.6%)
Maintenance immunosuppression				
Tacrolimus or MMF	34,023 (82.0%)	129,657 (83.3%)	40,993 (81.8%)	204,673 (82.8%)
Cyclosporine or azathioprine	1,645 (4.0%)	5,131 (3.3%)	2,064 (4.1%)	8,840 (3.6%)
Other combination	5,850 (14.1%)	20,878 (13.4%)	7,077 (14.1%)	33,805 (13.7%)
mTOR inhibitor	2,998 (7.2%)	9,190 (5.9%)	3,265 (6.5%)	15,453 (6.3%)
Steroids	31,444 (74.7%)	122,479 (78.7%)	39,228 (78.3%)	193,151 (78.1%)

All entries are N (%) unless otherwise noted. All percentages are column percentages except for totals (row percentages). Abbreviations: SES – socioeconomic status; MMF - mycophenolate mofetil;

Table 4- 2. Risk of solid tumors by CMV recipient and donor serostatus pretransplant.

			CMV status (recipient/donor)							
	Total		R-/D-		R+			R-/D+		
Cancer group	N	IR	N	IR	N	IR	Adj. IRR (95% CI)	N	IR	Adj. IRR (95% CI)
Lip	104	8.35	25	11.57	45	5.75	0.66 (0.37 - 1.18)	34	13.75	1.03 (0.56 - 1.92)
Tongue	140	11.24	27	12.5	77	9.85	1.03 (0.59 - 1.80)	36	14.56	0.90 (0.47 - 1.74)
Salivary gland	60	4.82	12	5.56	39	4.99	0.87 (0.37 - 2.02)	9	3.64	0.74 (0.27 - 2.05)
Other oral cavity and pharynx	205	16.46	33	15.28	124	15.85	1.11 (0.67 - 1.84)	48	19.41	1.17 (0.66 - 2.06)
Esophagus	151	12.12	34	15.74	94	12.02	0.78 (0.45 - 1.34)	23	9.3	0.78 (0.41 - 1.50)
Stomach	223	17.91	21	9.72	167	21.35	1.73 (0.96 - 3.12)	35	14.16	1.13 (0.56 - 2.30)
Small intestine	103	8.27	19	8.8	72	9.21	0.65 (0.37 - 1.16)	12	4.85	0.23 (0.09 - 0.63)
Colorectum	674	54.12	102	47.22	434	55.49	1.17 (0.88 - 1.56)	138	55.81	1.23 (0.89 - 1.70)
Anus	119	9.56	12	5.56	86	11	1.24 (0.60 - 2.59)	21	8.49	1.05 (0.44 - 2.50)
Liver	296	23.77	31	14.35	217	27.75	1.37 (0.86 - 2.17)	48	19.41	0.81 (0.46 - 1.45)
Intrahepatic Bile Duct	75	6.02	9	4.17	43	5.5	1.38 (0.52 - 3.67)	23	9.3	2.13 (0.76 - 6.00)
Pancreas	300	24.09	47	21.76	202	25.83	1.06 (0.70 - 1.62)	51	20.63	0.98 (0.59 - 1.61)
Nose, middle ear, & larynx	164	13.17	19	8.8	110	14.06	1.62 (0.84 - 3.09)	35	14.16	1.22 (0.57 - 2.61)
Lung	1932	155.13	274	126.84	1347	172.23	1.24 (1.05 - 1.46)	311	125.78	0.94 (0.77 - 1.14)
Soft tissue and heart	109	8.75	13	6.02	68	8.69	1.24 (0.63 - 2.44)	28	11.32	1.50 (0.72 - 3.15)
Melanoma	557	44.73	125	57.87	308	39.38	0.86 (0.67 - 1.11)	124	50.15	0.83 (0.61 - 1.12)
Skin (non-melanoma, non-epithelial)	254	20.4	51	23.61	157	20.07	0.94 (0.61 - 1.43)	46	18.6	0.89 (0.54 - 1.47)
Breast	663	53.24	103	47.68	464	59.33	0.86 (0.65 - 1.14)	96	38.83	0.75 (0.52 - 1.07)
Genital	136	10.92	19	8.8	98	12.53	1.19 (0.64 - 2.19)	19	7.68	0.90 (0.42 - 1.92)
Prostate	1578	126.71	262	121.29	1044	133.49	1.12 (0.93 - 1.34)	272	110.01	0.86 (0.69 - 1.08)
Bladder	314	25.21	42	19.44	194	24.8	1.43 (0.92 - 2.22)	78	31.55	1.58 (0.97 - 2.57)
Kidney	1316	105.67	196	90.73	895	114.44	1.01 (0.83 - 1.24)	225	91	0.89 (0.70 - 1.13)
Brain and nervous system	89	7.15	17	7.87	58	7.42	0.89 (0.46 - 1.69)	14	5.66	0.60 (0.25 - 1.40)

Thyroid	387	31.08	69	31.94	248	31.71	0.95 (0.66 - 1.37)	70	28.31	1.07 (0.70 - 1.62)
Kaposi sarcoma	93	7.47	5	2.31	80	10.23	2.13 (0.84 - 5.44)	8	3.24	0.80 (0.23 - 2.76)
Miscellaneous	1789	143.65	295	136.56	1109	141.8	0.98 (0.83 - 1.15)	385	155.71	0.98 (0.81 - 1.18)

IRs per 100,000 person-years. Total person-time at risk: 1,245,369. Adjusted IRR models include recipient sex, age (0-17, 18-34, 35-49, 50-64, 65+), race/ethnicity, organ type (kidney, liver, other/multiple), EBV recipient/donor status, and SES quintile. Underlined results are considered statistically significant.

Table 4- 3. Risk of select cancer stratified by recipient age and sex within CMV risk group.

	Age at transplant				Sex	
Cancer Site	18-34	35-49	50-64	65+	Male	Female
Small intestine						
R+	-	2.67 (0.33 - 21.97)	0.63 (0.31 - 1.28)	0.37 (0.07 - 2.11)	0.88 (0.43 - 1.78)	0.32 (0.11 - 0.87)
R-/D+	-	1.60 (0.14 - 17.72)	0.14 (0.03 - 0.64)	0.39 (0.04 - 4.34)	0.22 (0.06 - 0.78)	0.25 (0.05 - 1.26)
Lung						
R+	-	1.12 (0.72 - 1.73)	1.26 (1.03 - 1.55)	1.20 (0.86 - 1.65)	1.21 (1.01 - 1.47)	1.31 (0.96 - 1.79)
R-/D+	-	1.05 (0.64 - 1.73)	0.96 (0.75 - 1.23)	0.77 (0.51 - 1.17)	0.92 (0.73 - 1.15)	1.00 (0.68 - 1.47)
Brain						
R+	-	0.47 (0.13 - 1.64)	0.86 (0.38 - 1.97)	-	0.79 (0.37 - 1.69)	1.20 (0.34 - 4.26)
R-/D+	-	-	0.45 (0.14 - 1.51)	-	0.34 (0.11 - 1.10)	1.42 (0.34 - 5.99)
Bladder						
R+	1.10 (0.10 - 12.63)	1.22 (0.39 - 3.88)	1.33 (0.75 - 2.36)	1.88 (0.73 - 4.78)	1.81 (1.05 - 3.12)	0.78 (0.36 - 1.66)
R-/D+	3.55 (0.37 - 34.38)	1.46 (0.43 - 5.01)	1.60 (0.85 - 3.02)	1.44 (0.48 - 4.31)	2.05 (1.14 - 3.68)	0.74 (0.29 - 1.94)
Breast						
R+	2.43 (0.52 - 11.28)	0.65 (0.39 - 1.09)	0.88 (0.59 - 1.31)	0.94 (0.44 - 2.01)		
R-/D+	0.92 (0.13 - 6.58)	0.79 (0.43 - 1.46)	0.73 (0.43 - 1.22)	0.74 (0.28 - 1.97)		
Prostate						
R+	-	1.27 (0.74 - 2.17)	1.20 (0.96 - 1.50)	0.89 (0.62 - 1.28)		
R-/D+	-	0.82 (0.43 - 1.58)	0.79 (0.60 - 1.05)	1.070.70 - 1.64)		

Adjusted IRRs and 95% confidence intervals calculated compared to reference R-/D-. Adjusted models include recipient sex or age (0-17, 18-34, 35-49, 50-64, 65+), race/ethnicity, organ type (kidney, liver, other/multiple), EBV recipient/donor status, and SES quintile. Underlined results are considered statistically significant

Chapter 5: Summary

Key findings

The primary research aims of this dissertation were as follows. First, we sought to describe the association between congenital CMV infection and acute lymphoblastic leukemia by estimating the prevalence of cCMV at birth in ALL cases and cancer-free controls. Second, we aimed to assess the role of CMV infection and hematologic malignancies in the setting of solid organ transplantation, as there is a high risk for CMV infection following transplantation. Finally, we sought to further elucidate the role of CMV in carcinogenesis by estimating the association between CMV risk and incidence of solid tumor cancers.

In Manuscript 1, among the 1,189 ALL cases and 4,756 matched controls, we detected cCMV in 0.5% of ALL cases and 0.4% of controls. There was no difference in the odds of cCMV infection comparing ALL cases to controls in our primary analysis. We also did not detect any differences in our stratified analysis by demographic characteristics. However, among those with available cytogenetic data, hyperdiploid ALL was associated with six time greater odds of cCMV infection compared to unmatched controls, albeit on the basis of 2 exposed cases. These findings offer partial support for the association of cCMV with ALL, specifically hyperdiploid disease, and should encourage continued research into this possible association.

In Manuscript 2, we conducted the largest investigation of CMV infection status pre-transplant as it relates to the risk of incident Hodgkin lymphoma, NHL, leukemia,

and myeloma following solid organ transplantation in the US. We found no significant differences in the risk for leukemia across CMV risk groups. We identified that CMV sero-mismatch (R-/D+) was associated with significantly lower risks of DLBCL compared to CMV seronegative recipient/donor pairs. Our results suggest the risk of DLBCL differs according to recipient and donor EBV serostatus. Moreover, risk was modified by receipt of maintenance immunosuppression, in which cyclosporine or azathioprine were associated with significantly elevated risk among the R+ group. Given the findings of our study, it appears CMV may have both a pathogenic and protective role in DLBCL carcinogenesis, depending upon the context, specifically, the adjunctive use of immunomodulatory therapy.

Finally, in Manuscript 3, we found that the incidence of solid tumor cancers did not differ by CMV R+ or R-/D+ serostatus compared to CMV R-/D-. However, we identified R-/D+ CMV was inversely associated with small intestine cancer compared to the recipient/donor CMV seronegative group. Additionally, we found among the R+ group an elevated risk of lung cancer compared to the R-/D- group. For both cancers, these differences were most pronounced among males and recipients aged 50-64 years. We also identified an elevated risk of bladder cancer among males in both the R+ and R-/D+ CMV groups compared to the R-/D- group. Based on the findings of our study, it appears likely that CMV plays little if any role in carcinogenesis after transplantation.

Strengths and limitations

The key strength of Manuscript 1 that this was a very large, population-based case-control study with newborn dried blood spots available to us for biologic testing. Additionally, the data linkage through the Michigan BioTrust for Health enabled us to

examine potential associations with birth characteristics and parental demographics and leukemia. We also note that since all samples were collected immediately following birth, the temporality between measurement of cCMV and development of ALL was clear.

Limitations of Manuscript 1 include the lack of immunophenotype and cytogenetic data for most cases which contributed to imprecision of our point estimates in our stratified analysis. Another limitation was that our study population was largely White which limited our ability to examine subgroup differences that have been identified in the previous studies. Lastly, we detected a lower prevalence of CMV DNA in our samples compared to Francis et al, which may have been due to either a higher limit of detection in our assay or degradation of starting material due to differences in storage conditions.

The primary strength of Manuscripts 2 and 3 was the large size of the Transplant Cancer Match Study enabled us to examine CMV and cancer risk in nearly 250,000 recipients, which is the largest study of CMV and cancer to our knowledge. Further, the use of this data provided a unique cohort to examine cancer outcomes and exposures that would be otherwise difficult or costly to obtain data on, such as CMV infection. Another strength was that our study population was a representative sample of the US transplant population, so our results are generalizable to future investigation of cancer and solid organ transplantation.

Limitations of Manuscripts 2 and 3 are primarily based on missing serology data for CMV post-transplantation. CMV IgG antibody data was available for less than 10% of recipients in the 24 months of follow-up and therefore not included in our analysis. Therefore, we had limited information on how CMV risk groups corresponded to

seroconversion rates during follow-up. Additionally, we restricted our analysis to transplantations occurring in the year 2000 and beyond because CMV serostatus was not widely recorded in the years prior. Another limitation that was specific to Manuscript 3 was the lack of data on smoking and tobacco use. We did not have any surrogate variables in our dataset to be able to control for smoking, apart from SES, and therefore it is possible the association of lung cancer and CMV we detected may be due to residual confounding.

Future directions

Overall, our findings provide insight into the role of cytomegalovirus infection in the development of pediatric ALL and other cancers, with direction for future research opportunities. In Manuscript 1, we provide evidence that congenital CMV may play a role in the development of pediatric ALL, specifically in hyperdiploid subtype. Obtaining cytogenetic data for a greater proportion of cases would further expound this finding. Additionally, this potential association raises questions as to the underlying mechanism of how cCMV contributes to ALL. Future work could include epigenetic investigations that focus on establishing a biomarker of cCMV infection in ALL. This could be facilitated by universal screening for cCMV which has recently been adopted by Minnesota.

Our findings from Manuscripts 2 and 3 indicate there is little evidence for an association between risk for CMV infection and incident cancer. Although there was evidence for a greater incidence of lung cancer among CMV seropositive recipients, this was likely due to confounding. The inverse association of CMV risk group with DLBCL may be explained by the effect of antiviral treatment and CMV prophylaxis on EBV

infection following transplantation. Future analyses should incorporate additional immunologic data following CMV infection and reactivation in transplant recipients to further characterize this association.

Conclusions

Overall, we provide partial evidence that the development of pediatric acute lymphoblastic leukemia is influenced by congenital cytomegalovirus infection, specifically hyperdiploid disease. However, as it relates to other cancers in the setting of solid organ transplantation, there is an inverse relationship between CMV risk and diffuse large B-cell lymphoma, but this may be mediated by other factors including antiviral therapies given prophylactically. Altogether, these findings encourage continued research into CMV to fully understand the relationship between this highly prevalent virus and pediatric leukemia, as well as its role in malignancy in general.

References

1 Griffiths P, Baraniak I, Reeves M. The pathogenesis of human cytomegalovirus. *J Pathol* 2015;**235**:288–97. https://doi.org/10.1002/path.4437.

2 Jackson SE, Mason GM, Wills MR. Human cytomegalovirus immunity and immune evasion. *Virus Res* 2011:151–60. https://doi.org/10.1016/j.virusres.2010.10.031.

3 Dupont L, Reeves MB. Cytomegalovirus latency and reactivation: recent insights into an age old problem. *Rev Med Virol* 2016:75–89. https://doi.org/10.1002/rmv.1862.

4 Bruggeman CA. Cytomegalovirus and latency: an overview. *Virchows Arch B Cell Pathol Incl Mol Pathol* 1993:325–33. https://doi.org/10.1007/BF02915131.

5 Forbes BA. Acquisition of cytomegalovirus infection: an update. *Clin Microbiol Rev* 1989;**2**:204. https://doi.org/10.1128/CMR.2.2.204.

6 Amin MM, Bialek SR, Dollard SC, Wang C. Urinary cytomegalovirus shedding in the United States: The National Health and Nutrition Examination Surveys, 1999-2004. *Clin Infect Dis* 2018. https://doi.org/10.1093/cid/ciy143.

7 Hamprecht K, Goelz R. Transmission of human cytomegalovirus via breastmilk and potential risks to very preterm infants. *Microbiol Aust* 2015. https://doi.org/10.1071/ma15066.

8 Schleiss MR. Acquisition of human cytomegalovirus infection in infants via breast milk: Natural immunization or cause for concern? *Rev Med Virol* 2006. https://doi.org/10.1002/rmv.484.

9 Mayer BT, Matrajt L, Casper C, Krantz EM, Corey L, Wald A, *et al.* Dynamics of persistent oral cytomegalovirus shedding during primary infection in ugandan infants. *J Infect Dis* 2016. https://doi.org/10.1093/infdis/jiw442.

10 Schleiss MR. Cytomegalovirus. *Matern. Immun.* Elsevier; 2019. p. 253–88.

11 Adler SP. Molecular epidemiology of cytomegalovirus: Viral transmission among children attending a day care center, their parents, and caretakers. *J Pediatr* 1988. https://doi.org/10.1016/S0022-3476(88)80314-7.

12 Cannon MJ, Hyde TB, Schmid DS. Review of cytomegalovirus shedding in bodily fluids and relevance to congenital cytomegalovirus infection. *Rev Med Virol*

2011:240–55. https://doi.org/10.1002/rmv.695.

13 Hutto C, Ricks R, Garvie M, Pass RF. Epidemiology of cytomegalovirus infections in young children: day care vs. home care 1985;**4**:149–52. https://doi.org/10.1097/00006454-198503000-00008.

14 Hutlo C, Little EAE, Ricks R, Lee JJD, Pass RF, Hutto C, *et al.* Isolation of cytomegalovirus from toys and hands in a day care center. *J Infect Dis* 1986;**154**:527–30. https://doi.org/10.1093/infdis/154.3.527.

15 Handsfield HH, Chandler SH, Caine VA, Meyers JD, Corey L, Medeiros E, *et al.* Cytomegalovirus infection in sex partners: evidence for sexual transmission. *J Infect Dis* 1985;**151**:344–8. https://doi.org/10.1093/INFDIS/151.2.344.

16 Adler SP. The molecular epidemiology of cytomegalovirus transmission among children attending a day care center. *J Infect Dis* 1985;**152**:760–8. https://doi.org/10.1093/infdis/152.4.760.

17 Adler SP. Transfusion-associated cytomegalovirus infections. *Rev Infect Dis* 1983;**5**:977–93. https://doi.org/10.1093/CLINIDS/5.6.977.

18 Sinnott IV JT, Cancio MR. Cytomegalovirus. *Infect Control* 1987;**8**:79–82. https://doi.org/10.1017/S0195941700067138.

19 Lanzieri TM, Kruszon-Moran D, Amin MM, Bialek SR, Cannon MJ, Carroll MD, *et al.* Seroprevalence of cytomegalovirus among children 1 to 5 years of age in the United States from the national health and nutrition examination survey of 2011 to 2012. *Clin Vaccine Immunol* 2015. https://doi.org/10.1128/CVI.00697-14.

20 Staras SAS, Dollard SC, Radford KW, Flanders WD, Pass RF, Cannon MJ. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. *Clin Infect Dis* 2006;**43**:1143–51. https://doi.org/10.1086/508173/2/43-9-1143-TBL003.GIF.

21 Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleesschauwer B, *et al.* Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. *Rev Med Virol* 2019. https://doi.org/10.1002/rmv.2034.

22 Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol* 2010;**20**:202–13. https://doi.org/10.1002/rmv.655.

23 Bate SL, Dollard SC, Cannon MJ. Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988-2004. *Clin Infect Dis* 2010;**50**:1439–47. https://doi.org/10.1086/652438.

24 Dowd JB, Aiello AE, Alley DE. Socioeconomic disparities in the seroprevalence of cytomegalovirus infection in the US population: NHANES III. *Epidemiol Infect* 2009;**137**:58–65. https://doi.org/10.1017/S0950268808000551.

25 Manicklal S, Emery VC, Lazzarotto T, Boppana SB, Gupta RK. The 'silent' global burden of congenital cytomegalovirus. 2013;**26**:86–102. https://doi.org/10.1128/CMR.00062-12.

26 Kenneson A, Cannon MJ. Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev Med Virol* 2007:253–76. https://doi.org/10.1002/rmv.535.

27 Demmler GJ. Infectious Diseases Society of America and Centers for Disease Control: Summary of a workshop on surveillance for congenital cytomegalovirus disease. *Rev Infect Dis* 1991:315–29. https://doi.org/10.1093/clinids/13.2.315.

28 Schleiss MR. Congenital cytomegalovirus infection: Update on management strategies. *Curr Treat Options Neurol* 2008:186–92. https://doi.org/10.1007/s11940-008-0020-2.

29 Boppana SB, Pass RF, Britt WJ, Stagno S, Alford CA. Symptomatic congenital cytomegalovirus infection: Neonatal morbidity and mortality. *Pediatr Infect Dis J* 1992;**11**:93–9. https://doi.org/10.1097/00006454-199202000-00007.

30 Bale JF, Petheram SJ, Robertson M, Murph JR, Demmler G. Human cytomegalovirusa sequence and UL144 variability in strains from infected children. *J Med Virol* 2001;**65**:90–6. https://doi.org/10.1002/jmv.2006.

31 Cannon MJ, Davis KF. Washing our hands of the congenital cytomegalovirus disease epidemic. *BMC Public Health* 2005. https://doi.org/10.1186/1471-2458-5-70.

32 Ross DS, Dollard SC, Victor M, Sumartojo E, Cannon MJ. The epidemiology and prevention of congenital cytomegalovirus infection and disease: activities of the Centers for Disease Control and Prevention Workgroup. *J Womens Health (Larchmt)* 2006;**15**:224–9. https://doi.org/10.1089/jwh.2006.15.224.

33 Cannon MJ, Westbrook K, Levis D, Schleiss MR, Thackeray R, Pass RF. Awareness of and behaviors related to child-to-mother transmission of cytomegalovirus. *Prev Med (Baltim)* 2012;**54**:351–7. https://doi.org/10.1016/j.ypmed.2012.03.009.

34 Fowler KB, Stagno S, Pass RF, Britt WJ, Boll TJ, Alford CA. The outcome of congenital cytomegalovirus infection in relation to maternal antibody status. *N Engl J Med* 1992;**326**:663–7. https://doi.org/10.1056/NEJM199203053261003.

35 Fowler KB, Stagno S, Pass RF. Maternal immunity and prevention of congenital cytomegalovirus infection. *JAMA* 2003;**289**:1008–11. https://doi.org/10.1001/jama.289.8.1008.

36 Hyde TB, Schmid DS, Cannon MJ. Cytomegalovirus seroconversion rates and risk factors: Implications for congenital CMV. *Rev Med Virol* 2010:311–26. https://doi.org/10.1002/rmv.659.

37 Adler SP. Screening for Cytomegalovirus during Pregnancy. *Infect Dis Obstet Gynecol* 2011;**2011**:. https://doi.org/10.1155/2011/942937.

38 Schleiss MR. Congenital cytomegalovirus: Impact on child health. *Contemp Pediatr* 2018.

39 Azevedo LS, Pierrotti LC, Abdala E, Costa SF, Strabelli TMV, Campos SV, *et al. Cytomegalovirus infection in transplant recipients*. vol. 70. 2015.

40 Stratta RJ, Shaefer MS, Markin RS, Wood RP, Kennedy EM, Langnas AN, *et al.* Clinical Patterns of Cytomegalovirus Disease After Liver Transplantation. *Arch Surg* 1989;**124**:1443–50. https://doi.org/10.1001/archsurg.1989.01410120093018.

41 Fryd DS, Peterson PK, Ferguson RM, Simmons RL, Balfour HH, Najarian JS. Cytomegalovirus as a risk factor in renal transplantation. *Transplantation* 1980;**30**:436–9. https://doi.org/10.1097/00007890-198012000-00010.

42 Simon DM, Levin S. Infectious complications of solid organ transplantations. *Infect Dis Clin North Am* 2001;**15**:521–49. https://doi.org/10.1016/S0891-5520(05)70158-6.

43 Humar A, Snydman D. *Cytomegalovirus in solid organ transplant recipients*. vol. 9. Am J Transplant; 2009.

44 Fishman JA, Emery V, Freeman R, Pascual M, Rostaing L, Schlitt HJ, *et al.*

Cytomegalovirus in transplantation - Challenging the status quo. *Clin Transplant* 2007:149–58. https://doi.org/10.1111/j.1399-0012.2006.00618.x.

45 Aucejo F, Rofaiel G, Miller C. Who is at risk for post-transplant lymphoproliferative disorders (PTLD) after liver transplantation? *J Hepatol* 2006;**44**:19–23. https://doi.org/10.1016/j.jhep.2005.10.008.

46 Walker R, Marshall W, Strickler J, Wiesner R, Velosa J, Habermann T, *et al.* Pretransplantation assessment of the risk of lymphoproliferative disorder. *Clin Infect Dis* 1995;**20**:1346–53. https://doi.org/10.1093/CLINIDS/20.5.1346.

47 Mañez R, Breinig MC, Linden P, Wilson J, Torre-Cisneros J, Kusne S, *et al.* Posttransplant Lymphoproliferative Disease in Primary Epstein-Barr Virus Infection after Liver Transplantation: The Role of Cytomegalovirus Disease. *J Infect Dis* 1997;**176**:1462–7. https://doi.org/10.1086/514142.

48 Opelz G, Daniel V, Naujokat C, Döhler B. Epidemiology of Pretransplant EBV and CMV Serostatus in Relation to Posttransplant Non-Hodgkin Lymphoma. *Transplantation* 2009;**88**:962–7. https://doi.org/10.1097/TP.0b013e3181b9692d.

49 Arav-Boger R. Strain Variation and Disease Severity in Congenital Cytomegalovirus Infection: In Search of a Viral Marker. *Infect Dis Clin North Am* 2015;**29**:401–14. https://doi.org/10.1016/j.idc.2015.05.009.

50 Cha TA, Tom E, Kemble GW, Duke GM, Mocarski ES, Spaete RR. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J Virol* 1996;**70**:78–83.

51 Brown JM, Kaneshima H, Subianto B, Mocarski ES. Dramatic interstrain differences in the replication of human cytomegalovirus in scid-hu mice. *J Infect Dis* 1995;**171**:1599–603. https://doi.org/10.1093/infdis/171.6.1599.

52 Boeckh M, Geballe AP. Cytomegalovirus: Pathogen, paradigm, and puzzle. *J Clin Invest* 2011:1673–80. https://doi.org/10.1172/JCI45449.

53 Arav-Boger R, Willoughby RE, Pass RF, Zong J, Jang W, Alcendor D, *et al.* Polymorphisms of the Cytomegalovirus (CMV)–Encoded Tumor Necrosis Factor–α and β-Chemokine Receptors in Congenital CMV Disease. *J Infect Dis* 2002;**186**:1057–64. https://doi.org/10.1086/344238.

54 Poole E, King CA, Sinclair JH, Alcami A. The UL144 gene product of human

cytomegalovirus activates NFκB via a TRAF6-dependent mechanism. *EMBO J* 2006;**25**:4390–9. https://doi.org/10.1038/sj.emboj.7601287.

55 Lurain NS, Kapell KS, Huang DD, Short JA, Paintsil J, Winkfield E, *et al.* Human Cytomegalovirus UL144 Open Reading Frame: Sequence Hypervariability in Low-Passage Clinical Isolates. *J Virol* 1999;**73**:10040–50. https://doi.org/10.1128/jvi.73.12.10040-10050.1999.

56 Bale JF, Petheram SJ, Robertson M, Murph JR, Demmler G. Human cytomegalovirus a sequence and UL144 variability in strains from infected children. *J Med Virol* 2001;**65**:90–6.

57 Picone O, Costa JM, Chaix ML, Ville Y, Rouzioux C, Leruez-Ville M. Human cytomegalovirus UL144 gene polymorphisms in congenital infections. *J Clin Microbiol* 2005;**43**:25–9. https://doi.org/10.1128/JCM.43.1.25-29.2005.

58 Arav-Boger R, Battaglia CA, Lazzarotto T, Gabrielli L, Zong JC, Hayward GS, *et al.* Cytomegalovirus (CMV)–Encoded *UL144* (Truncated Tumor Necrosis Factor Receptor) and Outcome of Congenital CMV Infection. *J Infect Dis* 2006;**194**:464–73. https://doi.org/10.1086/505427.

59 Penfold MET, Dairaghi DJ, Duke GM, Saederup N, Mocarski ES, Kemble GW, *et al.* Cytomegalovirus encodes a potent α chemokine. *Proc Natl Acad Sci U S A* 1999;**96**:9839–44. https://doi.org/10.1073/pnas.96.17.9839.

60 Caposio P, Orloff SL, Streblow DN. The role of cytomegalovirus in angiogenesis. *Virus Res* 2011;**157**:204–11. https://doi.org/10.1016/j.virusres.2010.09.011.

61 Cinatl J, Vogel JJ-U, Kotchetkov R, Doerr HW, Wilhelm Doerr H. Oncomodulatory signals by regulatory proteins encoded by human cytomegalovirus: a novel role for viral infection in tumor progression. *FEMS Microbiol Rev* 2004;**28**:59–77.

62 Guo G, Zhang L, Ye S, Hu Y, Li B, Sun X, *et al.* Polymorphisms and features of cytomegalovirus UL144 and UL146 in congenitally infected neonates with hepatic involvement. *PLoS One* 2017;**12**:e0171959. https://doi.org/10.1371/journal.pone.0171959.

63 Nelson CS, Herold BC, Permar SR. A new era in cytomegalovirus vaccinology: considerations for rational design of next-generation vaccines to prevent

congenital cytomegalovirus infection. *Npj Vaccines* 2018:1–9. https://doi.org/10.1038/s41541-018-0074-4.

64 Johnson RA, Huong S-M, Huang E-S. Activation of the Mitogen-Activated Protein Kinase p38 by Human Cytomegalovirus Infection through Two Distinct Pathways: a Novel Mechanism for Activation of p38. *J Virol* 2000;**74**:1158–67. https://doi.org/10.1128/jvi.74.3.1158-1167.2000.

65 Feire AL, Koss H, Compton T. Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proc Natl Acad Sci U S A* 2004;**101**:15470–5. https://doi.org/10.1073/pnas.0406821101.

66 Luo M, Guan JL. Focal adhesion kinase: A prominent determinant in breast cancer initiation, progression and metastasis. *Cancer Lett* 2010:127–39. https://doi.org/10.1016/j.canlet.2009.07.005.

67 Adeyinka A, Nui Y, Cherlet T, Snell L, Watson PH, Murphy LC. Activated mitogen-activated protein kinase expression during human breast tumorigenesis and breast cancer progression. *Clin Cancer Res* 2002;**8**:1747–53.

68 Kumar A, Tripathy MK, Pasquereau S, Al Moussawi F, Abbas W, Coquard L, *et al.* The Human Cytomegalovirus Strain DB Activates Oncogenic Pathways in Mammary Epithelial Cells. *EBioMedicine* 2018;**30**:167–83. https://doi.org/10.1016/j.ebiom.2018.03.015.

69 Khan KA, Coaquette A, Davrinche C, Herbein G. Bcl-3-Regulated Transcription from Major Immediate-Early Promoter of Human Cytomegalovirus in Monocyte-Derived Macrophages. *J Immunol* 2009;**182**:7784–94. https://doi.org/10.4049/jimmunol.0803800.

70 Albrecht T, Rapp F. Malignant transformation of hamster embryo fibroblasts following exposure to ultraviolet-irradiated human cytomegalovirus. *Virology* 1973;**55**:53–61. https://doi.org/10.1016/S0042-6822(73)81007-4.

71 Geder L, Kreider J, Rapp F. Human cells transformed in vitro by human cytomegalovirus: tumorigenicity in athymic nude mice. *J Natl Cancer Inst* 1977;**58**:1003–9. https://doi.org/10.1093/JNCI/58.4.1003.

72 Kumar A, Tripathy MK, Pasquereau S, Al Moussawi F, Abbas W, Coquard L, *et al.* The Human Cytomegalovirus Strain DB Activates Oncogenic Pathways in

Mammary Epithelial Cells. *EBioMedicine* 2018;**30**:167–83. https://doi.org/10.1016/J.EBIOM.2018.03.015.

73 Cinatl J, Cinatl J, Vogel J-U, Rabenau H, Kornhuber B, Doerr HW. Modulatory Effects of Human Cytomegalovirus Infection on Malignant Properties of Cancer Cells. *Intervirology* 1996;**39**:259–69. https://doi.org/10.1159/000150527.

74 Cinatl J, Scholz M, Kotchetkov R, Vogel JU, Doerr HW. Molecular mechanisms of the modulatory effects of HCMV infection in tumor cell biology. *Trends Mol Med* 2004;**10**:19–23. https://doi.org/10.1016/J.MOLMED.2003.11.002.

75 Harkins L, Matlaf L, Soroceanu L, Klemm K, Britt W, Wang W, *et al.* Detection of human cytomegalovirus in normal and neoplastic breast epithelium. *Herpesviridae* 2010;**1**:. https://doi.org/10.1186/2042-4280-1-8.

76 Ingerslev K, Høgdall E, Skovrider-Ruminski W, Schnack TH, Lidang M, Høgdall C, *et al.* The prevalence of EBV and CMV DNA in epithelial ovarian cancer. *Infect Agent Cancer* 2019;**14**:1–6. https://doi.org/10.1186/S13027-019-0223-Z/TABLES/3.

77 Wertheim P, Voute P. Neuroblastoma, Wilms' tumor, and cytomegalovirus. *J Natl Cancer Inst* 1976;**57**:701–7. https://doi.org/10.1093/JNCI/57.3.701.

78 Francis SS, Wallace AD, Wendt GA, Li L, Liu F, Riley LW, *et al.* In utero cytomegalovirus infection and development of childhood acute lymphoblastic leukemia. *Blood* 2017;**129**:1680–4. https://doi.org/10.1182/blood-2016-07-723148.

79 Wiemels JL, Talbäck M, Francis S, Feychting M. Early Infection with Cytomegalovirus and Risk of Childhood Hematologic Malignancies. *Cancer Epidemiol Biomarkers Prev* 2019;**28**:1024–7. https://doi.org/10.1158/1055-9965.EPI-19-0044.

80 Gallant RE, Arroyo K, Bracci PM, Li S, Metayer C, Kogan SC, *et al.* Clinical characteristics of cytomegalovirus-positive pediatric acute lymphoblastic leukemia at diagnosis. *Am J Hematol* 2022. https://doi.org/10.1002/AJH.26528.

81 Pui C-H. Childhood Leukemias. *N Engl J Med* 1995;**332**:1618–30. https://doi.org/10.1056/NEJM199506153322407.

82 Siegel DA, Henley SJ, Li J, Pollack LA, Van Dyne EA, White A. Rates and Trends of Pediatric Acute Lymphoblastic Leukemia — United States, 2001–2014.

MMWR Morb Mortal Wkly Rep 2017;**66**:950–4. https://doi.org/10.15585/mmwr.mm6636a3.

83 Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *CA Cancer J Clin* 2014;**64**:83–103. https://doi.org/10.3322/caac.21219.

84 Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, *et al.* Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin* 2019;**69**:363–85. https://doi.org/10.3322/CAAC.21565@10.3322/(ISSN)1542-4863.STATISTICS.

85 Carroll WL, Hunger SP. Therapies on the horizon for childhood acute lymphoblastic leukemia. *Curr Opin Pediatr* 2016:12–8. https://doi.org/10.1097/MOP.0000000000000293.

86 Kızılocak H, Okcu F. Late Effects of Therapy in Childhood Acute Lymphoblastic Leukemia Survivors. *Turkish J Haematol Off J Turkish Soc Haematol* 2019:1–11. https://doi.org/10.4274/tjh.galenos.2018.2018.0150.

87 Mody R, Li S, Dover DC, Sallan S, Leisenring W, Oeffinger KC, *et al.* Twenty-five-year follow-up among survivors of childhood acute lymphoblastic leukemia: a report from the Childhood Cancer Survivor Study 2008. https://doi.org/10.1182/blood.

88 Robison LL. *Late effects of acute lymphoblastic leukemia therapy in patients diagnosed at 0-20 years of age.* vol. 2011. 2011.

89 Hunger SP, Mullighan CG. Redefining ALL classification: Toward detecting high-risk ALL and implementing precision medicine. *Blood* 2015:3977–87. https://doi.org/10.1182/blood-2015-02-580043.

90 Harrison CJ. Cytogenetics of paediatric and adolescent acute lymphoblastic leukaemia. *Br J Haematol* 2009:147–56. https://doi.org/10.1111/j.1365-2141.2008.07417.x.

91 Moorman A V. The clinical relevance of chromosomal and genomic abnormalities in B-cell precursor acute lymphoblastic leukaemia. *Blood Rev* 2012;**26**:123–35. https://doi.org/10.1016/j.blre.2012.01.001.

92 Seif AE. Pediatric leukemia predisposition syndromes: Clues to understanding

leukemogenesis. *Cancer Genet* 2011:227–44. https://doi.org/10.1016/j.cancergen.2011.04.005.

93 Chessells JM, Harrison G, Richards SM, Bailey CC, Hill FGH, Gibson BE, *et al.* Down's syndrome and acute lymphoblastic leukaemia: Clinical features and response to treatment. *Arch Dis Child* 2001;**85**:321–5. https://doi.org/10.1136/adc.85.4.321.

94 Alter BP. Cancer in Fanconi anemia, 1927-2001. *Cancer* 2003;**97**:425–40. https://doi.org/10.1002/cncr.11046.

95 Stewart A, Webb J, Giles D, Hewitt D. Malignant Disease in Childhood and Diagnostic Irradiation In Utero. *Lancet* 1956;**268**:447. https://doi.org/10.1016/S0140-6736(56)91923-7.

96 Stewart A, Kneale GW. Radiation dose effects in relation to obstetric x-rays and childhood cancers. *Lancet (London, England)* 1970;**1**:1185–8. https://doi.org/10.1016/s0140-6736(70)91782-4.

97 Spector LG, Robison LL, S. B. Epidemiology and Etiology Childhood Leukemias. 2nd ed. Cambridge University Press; n.d. p. 48–66.

98 Greaves M. Childhood leukaemia. *BMJ Br Med J* 2002;**324**:283. https://doi.org/10.1136/BMJ.324.7332.283.

99 Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer* 1995;**71**:1. https://doi.org/10.1038/BJC.1995.1.

100 Smith M. Considerations on a possible viral etiology for B-precursor acute lymphoblastic leukemia of childhood. *J Immunother* 1997;**20**:89–100. https://doi.org/10.1097/00002371-199703000-00001.

101 Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* 1988;**2**:120–5.

102 Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet* 2008;**371**:1030–43. https://doi.org/10.1016/S0140-6736(08)60457-2.

103 Williams LA, Yang JJ, Hirsch BA, Marcotte EL, Spector LG. Is There Etiologic Heterogeneity between Subtypes of Childhood Acute Lymphoblastic Leukemia? A Review of Variation in Risk by Subtype. *Cancer Epidemiol Biomarkers Prev* 2019;**28**:846–56. https://doi.org/10.1158/1055-9965.EPI-18-0801.

104 Lupo PJ, Spector LG. Cancer Progress and Priorities: Childhood Cancer. *Cancer Epidemiol Biomarkers Prev* 2020;**29**:1081–94. https://doi.org/10.1158/1055-9965.EPI-19-0941.

105 Choi KY, Schimmenti LA, Jurek AM, Sharon B, Daly K, Khan C, *et al.* Detection of cytomegalovirus DNA in dried blood spots of minnesota infants who do not pass newborn hearing screening. *Pediatr Infect Dis J* 2009;**28**:1095–8. https://doi.org/10.1097/INF.0b013e3181af6230.

106 Dollard SC, Dreon M, Hernandez-Alvarado N, Amin MM, Wong P, Lanzieri TM, *et al.* Sensitivity of Dried Blood Spot Testing for Detection of Congenital Cytomegalovirus Infection. *JAMA Pediatr* 2021;**175**:. https://doi.org/10.1001/JAMAPEDIATRICS.2020.5441.

107 Cheeran MCJ, Lokensgard JR, Schleiss MR. *Neuropathogenesis of congenital cytomegalovirus infection: Disease mechanisms and prospects for intervention*. vol. 22. 2009.

108 Xiaofei E, Kowalik TF. The DNA damage response induced by infection with human cytomegalovirus and other viruses. *Viruses* 2014;**6**:2155–85. https://doi.org/10.3390/V6052155.

109 Hahn G, Jores R, Mocarski ES. Cytomegalovirus remains latent in a common precursor of dendritic and myeloid cells. *Proc Natl Acad Sci U S A* 1998;**95**:3937–42. https://doi.org/10.1073/pnas.95.7.3937.

110 Crawford LB, Caposio P, Kreklywich C, Pham AH, Hancock MH, Jones TA, *et al.* Human Cytomegalovirus US28 ligand binding activity is required for latency in CD34+ hematopoietic progenitor cells and humanized NSG mice. *MBio* 2019;**10**:. https://doi.org/10.1128/mBio.01889-19.

111 Albano MS, Ciubotariu R, Dobrila L, Tarnawski M, DeLeon M, Watanabe C, *et al.* Cytomegalovirus viral load in cord blood and impact of congenital infection on markers of hematopoietic progenitor cell potency. *Transfusion* 2017;**57**:2768–74. https://doi.org/10.1111/trf.14257.

112 Hayden RT, Gu Z, Ingersoll J, Abdul-Ali D, Shi L, Pounds S, *et al.* Comparison of droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. *J Clin Microbiol* 2013;**51**:540–6.

https://doi.org/10.1128/JCM.02620-12.

113 Sedlak RH, Cook L, Cheng A, Magaret A, Jerome KR. Clinical Utility of Droplet Digital PCR for Human Cytomegalovirus. *J Clin Microbiol* 2014;**52**:2844. https://doi.org/10.1128/JCM.00803-14.

114 Hall EM, Flores SR, De Jesús VR. Influence of Hematocrit and Total-Spot Volume on Performance Characteristics of Dried Blood Spots for Newborn Screening. *Int J Neonatal Screen* 2015;**1**:69. https://doi.org/10.3390/IJNS1020069.

115 Kenneson A, Cannon MJ. Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev Med Virol* 2007;**17**:253–76. https://doi.org/10.1002/RMV.535.

116 Ssentongo P, Hehnly C, Birungi P, Roach MA, Spady J, Fronterre C, *et al.* Congenital Cytomegalovirus Infection Burden and Epidemiologic Risk Factors in Countries With Universal Screening: A Systematic Review and Meta-analysis. *JAMA Netw Open* 2021;**4**:e2120736–e2120736. https://doi.org/10.1001/JAMANETWORKOPEN.2021.20736.

117 Boppana SB, Ross SA, Fowler KB. Congenital Cytomegalovirus Infection: Clinical Outcome. *Clin Infect Dis An Off Publ Infect Dis Soc Am* 2013;**57**:S178. https://doi.org/10.1093/CID/CIT629.

118 Kruc R, Osterholm EA, Hernandez-Alvarado N, Rosendahl S, McCann M, Sidebottom A, *et al.* #40: Does Cytomegalovirus (CMV) Viral Load Correlate with Disease Severity in the Setting of Congenital CMV (cCMV) Infection? Results from a Universal cCMV Screening Study. *J Pediatric Infect Dis Soc* 2021;**10**:S13–4. https://doi.org/10.1093/JPIDS/PIAB031.029.

119 Walter S, Atkinson C, Sharland M, Rice P, Raglan E, Emery VC, *et al.* Congenital cytomegalovirus: association between dried blood spot viral load and hearing loss. *Arch Dis Child - Fetal Neonatal Ed* 2008;**93**:F280–5. https://doi.org/10.1136/ADC.2007.119230.

120 Boppana SB, Ross SA, Novak Z, Shimamura M, Tolan RW, Palmer AL, *et al.* Dried Blood Spot Real-Time Polymerase Chain Reaction Assays to Screen Newborns for Congenital Cyotemegalovirus Infection. *JAMA* 2010;**303**:1375. https://doi.org/10.1001/JAMA.2010.423.

121 *Congenital cytomegalovirus approved for addition to newborn screening panel*. Minnesota Department of Health. 2022. URL: https://www.health.state.mn.us/news/pressrel/2022/newborn020222.html (Accessed 11 March 2022).

122 *Update on congenital cytomegalovirus infection: Prenatal prevention, newborn diagnosis, and management*. Canadian Paediatric Society. n.d. URL: https://cps.ca/documents/position/update-on-congenital-cytomegalovirus-infection-prenatal-prevention-newborn-diagnosis-and-management (Accessed 11 March 2022).

123 Zhang LJ, Hanff P, Rutherford C, Churchill WH, Crumpacker CS. Detection of Human Cytomegalovirus DNA, RNA, and Antibody in Normal Donor Blood. *J Infect Dis* 1995;**171**:1002–6. https://doi.org/10.1093/infdis/171.4.1002.

124 Herman D, Han H. Cytomegalovirus in liver transplant recipients. *Curr Opin Organ Transplant* 2017:345–50. https://doi.org/10.1097/MOT.0000000000000433.

125 Engels EA, Pfeiffer RM, Fraumeni JF, Kasiske BL, Israni AK, Snyder JJ, *et al.* Spectrum of cancer risk among US solid organ transplant recipients. *JAMA - J Am Med Assoc* 2011;**306**:1891–901. https://doi.org/10.1001/jama.2011.1592.

126 Jorgenson MR, Parajuli S, Marka N, Leverson GE, Smith JA, Mandelbrot DA, *et al.* Geographic Distribution of Cytomegalovirus Serology in Kidney and Pancreas Transplant Recipients in the United States. *Transplant Direct* 2021;**7**:. https://doi.org/10.1097/TXD.0000000000001147.

127 Swerdlow S.H., Campo E., Harris N.L., Jaffe E.S., Pileri S.A., Stein H., *et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. IARC: Lyon 2008*. 2008.

128 *Site Recode ICD-O-3/WHO 2008 - SEER Data Reporting Tools*. n.d. URL: https://seer.cancer.gov/siterecode/icdo3_dwhoheme/index.html (Accessed 25 August 2021).

129 Turner J, Morton L, Linet M, Clarke C, Kadin M, Vajdic C, *et al.* InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. *Blood*

2010;**116**:. https://doi.org/10.1182/BLOOD-2010-06-289561.

130 Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control* 2001;**12**:703–11. https://doi.org/10.1023/A:1011240019516.

131 Vanichanan J, Udomkarnjananun S, Avihingsanon Y, Jutivorakool K. Common viral infections in kidney transplant recipients. *Kidney Res Clin Pract* 2018:323–37. https://doi.org/10.23876/j.krcp.18.0063.

132 Vajdic CM, McDonald SP, McCredie MRE, Van Leeuwen MT, Stewart JH, Law M, *et al.* Cancer incidence before and after kidney transplantation. *J Am Med Assoc* 2006;**296**:2823–31. https://doi.org/10.1001/jama.296.23.2823.

133 Litjens NHR, van der Wagen L, Kuball J, Kwekkeboom J. Potential beneficial effects of cytomegalovirus infection after transplantation. *Front Immunol* 2018;**9**:389.

134 Kaminski H, Garrigue I, Couzi L, Taton B, Bachelet T, Moreau JF, *et al.* Surveillance of gd T cells predicts cytomegalovirus infection resolution in kidney transplants. *J Am Soc Nephrol* 2016;**27**:637–45. https://doi.org/10.1681/ASN.2014100985/-/DCSUPPLEMENTAL.

135 D'Offizi G, Gioia C, Martini F, Volpi I, Solmone M, Poccia F, *et al.* γδ T cells and resolution of cytomegalovirus infection in an HIV/HCV coinfected patient after liver transplantation [2]. *Transplantation* 2005;**80**:1523–4. https://doi.org/10.1097/01.TP.0000180524.28964.E9.

136 Khairallah C, Déchanet-Merville J, Capone M. γδ T Cell-Mediated Immunity to Cytomegalovirus Infection. *Front Immunol* 2017;**8**:105. https://doi.org/10.3389/fimmu.2017.00105.

137 Khairallah C, Netzer S, Villacreces A, Juzan M, Rousseau B, Dulanto S, *et al.* γδ T Cells Confer Protection against Murine Cytomegalovirus (MCMV). *PLOS Pathog* 2015;**11**:e1004702. https://doi.org/10.1371/JOURNAL.PPAT.1004702.

138 Gaballa A, Alagrafi F, Uhlin M, Stikvoort A. Revisiting the role of γδ t cells in anti-cmv immune response after transplantation. *Viruses* 2021;**13**:1–13. https://doi.org/10.3390/v13061031.

139 Scheper W, Van Dorp S, Kersting S, Pietersma F, Lindemans C, Hol S, *et al.* γδT

cells elicited by CMV reactivation after allo-SCT cross-recognize CMV and leukemia. *Leukemia* 2013;**27**:1328–38. https://doi.org/10.1038/LEU.2012.374.

140 Knight A, Madrigal AJ, Grace S, Sivakumaran J, Kottaridis P, Mackinnon S, *et al.* The role of Vδ2-negative γδ T cells during cytomegalovirus reactivation in recipients of allogeneic stem cell transplantation. *Blood* 2010;**116**:2164–72. https://doi.org/10.1182/BLOOD-2010-01-255166.

141 Löunnqvist B, Ringdegn O, Ljungman P, Wahren B, Gahrton G. Reduced risk of recurrent leukaemia in bone marrow transplant recipients after cytomegalovirus infection. *Br J Haematol* 1986;**63**:671–9. https://doi.org/10.1111/J.1365-2141.1986.TB07551.X/FORMAT/PDF.

142 Godder KT, Henslee-Downey PJ, Mehta J, Park BS, Chiang KY, Abhyankar S, *et al.* Long term disease-free survival in acute leukemia patients recovering with increased gammadelta T cells after partially mismatched related donor bone marrow transplantation. *Bone Marrow Transplant* 2007;**39**:751–7. https://doi.org/10.1038/SJ.BMT.1705650.

143 Elmaagacli AH, Steckel NK, Koldehoff M, Hegerfeldt Y, Trenschel R, Ditschkowski M, *et al.* Early human cytomegalovirus replication after transplantation is associated with a decreased relapse risk: evidence for a putative virus-versus-leukemia effect in acute myeloid leukemia patients. *Blood* 2011;**118**:1402–12. https://doi.org/10.1182/BLOOD-2010-08-304121.

144 Green ML, Leisenring WM, Xie H, Walter RB, Mielcarek M, Sandmaier BM, *et al.* CMV reactivation after allogeneic HCT and relapse risk: evidence for early protection in acute myeloid leukemia. *Blood* 2013;**122**:1316–24. https://doi.org/10.1182/BLOOD-2013-02-487074.

145 Takenaka K, Nishida T, Asano-Mori Y, Oshima K, Ohashi K, Mori T, *et al.* Cytomegalovirus Reactivation after Allogeneic Hematopoietic Stem Cell Transplantation is Associated with a Reduced Risk of Relapse in Patients with Acute Myeloid Leukemia Who Survived to Day 100 after Transplantation: The Japan Society for Hematopoietic Cell Transplantation Transplantation-related Complication Working Group. *Biol Blood Marrow Transplant* 2015;**21**:2008–16. https://doi.org/10.1016/J.BBMT.2015.07.019.

146 Ito S, Pophali P, Co W, Koklanaris EK, Superata J, Fahle GA, *et al.* CMV reactivation is associated with a lower incidence of relapse after allo-SCT for CML. *Bone Marrow Transplant* 2013;**48**:1313–6. https://doi.org/10.1038/BMT.2013.49.

147 Bigley AB, Baker FL, Simpson RJ. Cytomegalovirus: an unlikely ally in the fight against blood cancers? *Clin Exp Immunol* 2018;**193**:265–74. https://doi.org/10.1111/CEI.13152.

148 Cichocki F, Cooley S, Davis Z, DeFor TE, Schlums H, Zhang B, *et al.* CD56dimCD57+NKG2C+ NK cell expansion is associated with reduced leukemia relapse after reduced intensity HCT. *Leukemia* 2016;**30**:456–63. https://doi.org/10.1038/LEU.2015.260.

149 Monsiváis-Urenda A, Noyola-Cherpitel D, Hernández-Salinas A, García-Sepúlveda C, Romo N, Baranda L, *et al.* Influence of human cytomegalovirus infection on the NK cell receptor repertoire in children. *Eur J Immunol* 2010;**40**:1418–27. https://doi.org/10.1002/EJI.200939898.

150 Bigley AB, Rezvani K, Shah N, Sekine T, Balneger N, Pistillo M, *et al.* Latent cytomegalovirus infection enhances anti-tumour cytotoxicity through accumulation of NKG2C+ NK cells in healthy humans. *Clin Exp Immunol* 2016;**185**:239–51. https://doi.org/10.1111/CEI.12785.

151 Sun JC, Beilke JN, Lanier LL. Adaptive immune features of natural killer cells. *Nature* 2009;**457**:557–61. https://doi.org/10.1038/NATURE07665.

152 Gumá M, Angulo A, Vilches C, Gómez-Lozano N, Malats N, López-Botet M. Imprint of human cytomegalovirus infection on the NK cell receptor repertoire. *Blood* 2004;**104**:3664–71. https://doi.org/10.1182/BLOOD-2004-05-2058.

153 Tomasec P, Braud VM, Rickards C, Powell MB, McSharry BP, Gadola S, *et al.* Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. *Science* 2000;**287**:1031–3. https://doi.org/10.1126/SCIENCE.287.5455.1031.

154 Couzi L, Levaillant Y, Jamai A, Pitard V, Lassalle R, Martin K, *et al.* Cytomegalovirus-Induced γδ T Cells Associate with Reduced Cancer Risk after Kidney Transplantation. *J Am Soc Nephrol* 2010;**21**:181–8.

https://doi.org/10.1681/ASN.2008101072.

155 Browne BJ, Young JA, Dunn TB, Matas AJ. The impact of cytomegalovirus infection ≥1 year after primary renal transplantation. *Clin Transplant* 2010;**24**:572. https://doi.org/10.1111/J.1399-0012.2010.01208.X.

156 Lizaola-Mayo BC, Rodriguez EA. Cytomegalovirus infection after liver transplantation. *World J Transplant* 2020;**10**:183. https://doi.org/10.5500/WJT.V10.I7.183.

157 Hughes D, Hafferty J, Fulton L, Friend P, Devaney A, Loke J, *et al.* Donor and recipient CMV serostatus and antigenemia after renal transplantation: An analysis of 486 patients. *J Clin Virol* 2008;**41**:92. https://doi.org/10.1016/J.JCV.2007.10.006.

158 Brayman K, Dafoe D, Smythe W, Barker C, Perloff L, Naji A, *et al.* Prophylaxis of serious cytomegalovirus infection in renal transplant candidates using live human cytomegalovirus vaccine. Interim results of a randomized controlled trial. *Arch Surg* 1988;**123**:1502–8. https://doi.org/10.1001/ARCHSURG.1988.01400360072012.

159 Hirata M, Teraski I, Cho Y. Cytomegalovirus antibody status and renal transplantation: 1987-1994. *Transplantation* 1996;**62**:34–7. https://doi.org/10.1097/00007890-199607150-00007.

160 Warrell M, Chinn I, Morris P, Tobin J. The Effects of Viral Infections on Renal Transplants and Their Recipients. *QJM An Int J Med* 1980;**49**:219–31. https://doi.org/10.1093/OXFORDJOURNALS.QJMED.A067618.

161 Kasiske BL, Snyder JJ, Gilbertson DT, Wang C. Cancer after Kidney Transplantation in the United States. *Am J Transplant* 2004;**4**:905–13. https://doi.org/10.1111/j.1600-6143.2004.00450.x.

162 Taher C, de Boniface J, Mohammad A, Religa P, Hartman J, Yaiw K, *et al.* High prevalence of human cytomegalovirus proteins and nucleic acids in primary breast cancer and metastatic sentinel lymph nodes. *PLoS One* 2013;**8**:. https://doi.org/10.1371/JOURNAL.PONE.0056795.

163 Cobbs CS, Harkins L, Samanta M, Gillespie GY, Bharara S, King PH, *et al.* Human cytomegalovirus infection and expression in human malignant glioma.

Cancer Res 2002;**62**:3347–50.

164 Söderberg-Nauclér C. Does cytomegalovirus play a causative role in the development of various inflammatory diseases and cancer? *J Intern Med* 2006;**259**:219–46. https://doi.org/10.1111/J.1365-2796.2006.01618.X.

165 Humar A, Snydman D. Cytomegalovirus in solid organ transplant recipients. *Am J Transplant* 2009;**9 Suppl 4**: https://doi.org/10.1111/J.1600-6143.2009.02897.X.

166 Courivaud C, Bamoulid J, Gaugler B, Roubiou C, Arregui C, Chalopin JM, *et al.* Cytomegalovirus exposure, immune exhaustion and cancer occurrence in renal transplant recipients. *Transpl Int* 2012;**25**:948–55. https://doi.org/10.1111/J.1432-2277.2012.01521.X.

167 Desai R, Collett D, Watson CJE, Johnson PJ, Moss P, Neuberger J. Impact of Cytomegalovirus on Long-term Mortality and Cancer Risk after Organ Transplantation. *Transplantation* 2015;**99**:1989–94. https://doi.org/10.1097/TP.0000000000000641.

168 Ruiz JC, Lücke L, Minarowski Ł, Kowalczuk O. Human cytomegalovirus infection effects on lung cancer prognosis after surgical resection. *Eur Respir J* 2019;**54**:PA3053. https://doi.org/10.1183/13993003.CONGRESS-2019.PA3053.

169 Alavanja M, Baron JA, Brownson RC, Buffler PA, DeMarini DM, Djordjevic M V., *et al.* Tobacco Smoke and Involuntary Smoking. *IARC Monogr Eval Carcinog Risks to Humans* 2004;**83**:1–1413.

170 Bai B, Wang X, Chen E, Zhu H. Human cytomegalovirus infection and colorectal cancer risk: a meta-analysis. *Oncotarget* 2016;**7**:76735. https://doi.org/10.18632/ONCOTARGET.12523.

171 Lv YL, Han FF, An ZL, Jia Y, Xuan LL, Gong LL, *et al.* Cytomegalovirus Infection Is a Risk Factor in Gastrointestinal Cancer: A Cross-Sectional and Meta-Analysis Study. *Intervirology* 2020;**63**:10–6. https://doi.org/10.1159/000506683.

172 Zhang L, Guo G, Xu J, Sun X, Chen W, Jin J, *et al.* Human cytomegalovirus detection in gastric cancer and its possible association with lymphatic metastasis. *Diagn Microbiol Infect Dis* 2017;**88**:62–8. https://doi.org/10.1016/J.DIAGMICROBIO.2017.02.001.

173 *Surveillance, Epidemiology, and End Results (SEER) Program Populations (1969-*

2019). National Cancer Institute, DCCPS, Surveillance Research Program. 2021. URL: www.seer.cancer.gov/popdata.

174 Ferrarini M, Ferrero E, Dagna L, Poggi A, Zocchi MR. Human gammadelta T cells: a nonredundant system in the immune-surveillance against cancer. *Trends Immunol* 2002;**23**:14–8. https://doi.org/10.1016/S1471-4906(01)02110-X.

Appendix 1: Strain variability as a predictor of acute lymphoblastic leukemia

1.1 Introduction

Variation of viral genes has been studied to predict cCMV disease severity and may provide insight into the pathogenesis of cancer. Among the genes meriting investigation is *UL144*, a truncated tumor necrosis factor-∝-like receptor gene. *UL144* polymorphisms have been studied among cCMV infected children and certain genotypes have been associated with symptomatic disease.[58] Other genes of interest include *UL146*, *UL55* (gB), and *UL9*. Like *UL144*, these candidate genes have, to varying degrees, been associated with disease severity and have been studied in congenitally infected children.[49,58] However, little is known how specific genotypes interact and nothing is known about whether one or a combination of these genes may be associated with ALL.

As an addendum to "Chapter 2: *Limited evidence for an association between congenital cytomegalovirus infection and pediatric acute lymphoblastic leukemia in a large, population-based study*, here we attempted to evaluate whether specific genotypes of CMV genes *UL144*, *UL146*, *UL55* (gB), and *UL9* were associated with ALL.

1.2 Methods

1.2.1 Sample selection

Dried blood spots (DBS) from cases and controls identified as cCMV-positive in the primary analysis were evaluated by qualitative PCR and subsequent nested PCR for amplification of *UL144*, *UL146*, *UL55* (gB), and *UL9* genes.

1.2.2 DNA extraction

DNA was extracted from one 6mm punch from the DBS cards using the GenTegra GenSolve DNA Complete Kit (GenTegra LLC, California, USA) following the manufacturer's protocol. Briefly, 609 μL lysis solution (GenTegra) and 11μL of Proteinase K was added to each DBS sample tube and incubated at 56°C for 1.5 hours with agitation at 1400 rpm. Samples were then transferred to a spin basket and centrifuged, with a Recovery Solution (GenTegra) added to the product. The product then

underwent simple DNA purification (GenTegra). Samples were eluted in 50 μL of elution buffer, and stored as necessary at -20°C.

1.2.3 PCR and nested PCR assay

Primers for PCR amplification were taken from Murthy *et al*. 2011 (**Table A- 1**). The sizes of the amplified loci were 400 base pair (bp) encompassing the proteolytic cleavage site of gB, and 531 bp, 480 bp, and 687 bp representing the full length *UL144, UL146,* and *UL9* protein coding regions.

Towne and Toldeo laboratory strains ($1x10^{-4}$ copies/mL stock) were used as positive controls to assess the validity of the assay. The reaction mixture consisted of 1μL template, 1μL each primer (forward and reverse), 45 μL Supermix (Invitrogen, Waltham, MA) for a total volume of 50 μL per sample. A master mix solution was made for each gene. The conditions for amplification with all primer sets were 94°C for 5 min, followed by 34 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min. This was followed by a single extension cycle of 72°C for 7 min. PCR products (2uL dye and 10uL PCR template) with 100 bp ladder were separated on a 1.5% agarose gel with 5uL ethidium bromide for 75 minutes and imaged under UV light. A nested PCR reaction was also performed using the PCR product from the first reaction. The reaction mixture consisted of 1μL PCR template, 1μL each primer (forward and reverse), 45 μL Supermix (Invitrogen, Waltham, MA) for a total volume of 50 μL per sample. The conditions for amplification were the same as previously used and the PCR products were separated on a 1.5% agarose gel as described previously.

When amplification failed, conditions were altered. The volume of template DNA of the study samples was increased to 2 μL and the Supermix volume was reduced to 46 μL to keep the reaction volume at 50 μL. The conditions for amplification were the same as described above. We then proceeded directly to a nested PCR, using 2 μL of PCR template and the same amplification conditions.

When those conditions did not yield a product, a new reaction mixture was used. The new mixture consisted of SuperFi Taq Polymerase (2U/μL) with 10mM dNTP mix, and 5X SuperFi Buffer (Invitrogen). The template volume was increased to 5 μL and the

supermix was reduced to 43 μL. The conditions for amplification were adjusted based on the calculated T_m for the gB primers: 94°C for 5 min, followed by 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min, followed by a single extension cycle of 72°C for 7 min. For the nested reaction, 10 μL of PCR template was used and the same PCR amplification conditions were applied.

1.2.4 Sequencing

PCR products with visible bands were purified using the QiaQuick PCR purification kit (Qiagen, Carlsbad, CA) following the manufacturer's instructions. The product was then sent to University of Minnesota Genomics Center (UMGC) for Sanger Sequencing, according to the guidelines of UMGC on an Applied Biosystems 3730xl DNA analyzer using ABI BigDye Terminator version 3.1 chemistry (Applied Biosystems, Massachusets, USA).

To assess genotype, sequences were aligned to reference files (UL55- GU365817-GU365825, UL144- GU365826-GU365834, UL146- GU365835-GU365846, and UL09- HM542481) using CodonCode Aligner Software (Version 10.0.1).

1.2.4 gB genotyping

Genotyping of gB using multiplex real-time PCR, previously described by de Vries et al., was conducted on all samples to confirm gB genotype or when nested PCR failed. Briefly, amplification of CMV gB genotypes 1, 2, 3, and 4 was performed in a reaction containing 10 μL of DNA template, 12.5 μL Roche 480 Taqman Probe Master (Roche Diagnostics, Rotkreuz, Switzerland), 0.4 μL uracil-DNA glycosylase (UNG) (Thermo Scientific, Waltham, Massachusetts, USA), 0.5 μL of each gB probe set (gB 1 and 3, gB 2 and 4), and 0.35 μL of water.

1.3 Results and Discussion

Both the Toledo and Towne positive controls showed visible bands and were purified and sent to UMGC. Sequences were a perfect match with the reference strains, indicating PCR amplification was successful.

Table A- 2 shows the results of amplification and genotyping. Among the study samples, no product was amplified for *UL146*. For UL9, only one sample had amplification and the remaining samples couldn't amplify after multiple attempts. For *UL144* and gB, 4/27 and 9/27 samples successfully amplified for *UL144* and gB, respectively. However, 2 of the 4 *UL144* samples failed Sanger sequencing and the other 2 could not align to the reference.

Using data from both the Sanger sequencing and multiplex PCR, we compared the distribution of gB genotypes among ALL cases and controls (**Figure A- 1**). For both cases and controls, the most prevalent genotype was gB1 (2/3 cases; 5/10 controls). When comparing the odds of being gB1 compared to gB 3 (since data was unavailable for other comparisons), Compared to controls, ALL cases had 0.40 times odds of being gB1 than gB3 but this difference was not statistically significant (95% CI: 0.02 – 10.02).

Overall, there was not sufficient data to be able to make comparisons between CMV genotype and case or control status. There are several potential explanations for why this assay failed to produce results. First, the PCR conditions were not ideal for amplification. The reaction volumes and PCR method were taken from the protocol used by Murthy et al., however, differences in the type of starting material used and viral load may not have been satisfactory. Murthy *et al*. utilized tissue culture isolates of CMV from urine, saliva, and vaginal fluid which typically have a higher viral load compared to blood. Second, the quality of the starting material may have been degraded. The samples were DBS that were stored at ambient temperature for anywhere between 10-33 years. However, DBS stored at room temperature in other studies has not been associated with diminished quality, and may only be due to the age of the specimen. Lastly, the size of the target amplicon may have been too large to successfully amplify, particularly for *UL144*, *UL146*, and *UL9*. While the sizes, 531 bp, 480 bp, and 687 bp, respectively, may not be considered large by other applications, it is possible in combination with the other potential limitations is what hindered its success.

Future directions could include the use of TOPO cloning for a more precise isolation of amplified gene products. The advantage to this technique is the reduction of potential sequence errors that may be introduced by repeated PCR amplification.

Table A- 1. PCR primers used for CMV genotyping

Primer	**Sequence**
UL144 forward	*5'-CGTATTACAAACCGCGGAGAGGAT-3'*
UL144 reverse	*5'-CTCAGACACGGTTCCGTAAAGTC-3'*
UL144 nested forward	*5'-CTTCCGGTAGGAGGCATGAAG-3'*
UL144 nested reverse	*5'-GACTTCATCGTACCGTGATC-3'*
UL146-147 forward	*5'- GTCATGGACGCAGTTTTG-3'*
UL146-147 reverse	*5'- GAACGATCTCGTCCGGTTC-3'*
UL146 hemi-nested reverse 1	*5'- CTAAAASATGGACGGCTAGG-3'*
UL146 hemi-nested reverse 2	*5'- GTCGTAATCTTCCARTTC-3'*
UL55 forward	*5'-TCCGAAGCCGAAGACTCGTA-3'*
UL55 reverse	*5'-GATGTAACCGCGCAACGTGT-3'*
UL55 nested forward	*5'-CATAGGTGAACTGCAGCTG-3'*
UL55 nested reverse	*5'-AGCATGGTGAAAAGAAGACG-3'*
UL09 forward	*5'-CATCTGTCTRCGAGCACCTC -3'*
UL09 reverse	*5'-GACCATCGGAAAAGATCATGG -3'*
UL09 nested forward	*5'-CAGTACGGACAAGTGTTYATG -3'*
UL09 nested reverse	*5'-GGTTCACGATATGGTTAATCAG -3'*

Source: Murthy et al, 2011. doi: 10.1371/journal.pone.0015949

Table A- 2. Genotyping results of ALL cases and controls

Sample	Viral load (copies/mL)	gB amplification PCR yes/no	gB genotype	gB real-time Multiplex Typing Yes/No	gB Multiplex genotype	UL144 amplification Yes/No	UL144 genotype	UL9 amplification Yes/No	UL9 genotype	UL146 amplification Yes/No	UL146 genotype
0079	7380	N		Y	No signal	N		N		N	
0836	2610	N		N		N		N		N	
0985	3800	N		N		N		N		N	
1863	45100	N		N		N		N		N	
2010	414000	Y	5	Y	No signal	Y	Could not be sequenced	Y	Could not align to ref	N	
2886	1440	Y	1	Y	No signal	N		N		N	
2951	29500	N		Y	1	N		N		N	
3018	14700	N		N		N		N		N	
3213	6360	Y	1	Y	3	Y	Could not be sequenced	N		N	
3774	13800	Y	1	Y	No signal	N		N		N	
3800	58800	Y	1	Y	No signal	N		N		N	
3904	24400	N		N		N		N		N	
3921	16000	Y	1A	Y	No signal	N		N		N	
3936	45600	N		N		N		N		N	
4002	2800	Y	4	Y	4	N		N		N	
4106	6100	N		N		N		N		N	

4586	4820	Y	2	Y	No signal	Y	Could not be sequenced	N		N	
4686	12200	Y	1	Y	No signal	N		N		N	
4702	1030	N		Y	No signal	N		N		N	
4828	121000	N		Y	1	N		N		N	
4895	29700	N		Y	3	N		N		N	
4933	9790	N		Y	No signal	N		N		N	
4962	1500	N		Y	No signal	N		N		N	
4973	10300	N		Y	No signal	Y	Could not be sequenced	N		N	
4995	2400	N		Y	No signal	N		N		N	
5028	18700	N		N		N		N		N	
5803	33700	N		Y	3	N		N		N	

Figure A- 1. Distribution of gB genotypes among ALL cases and controls.