Title

Associations between early-life and in utero infections and cytomegalovirus-positive acute lymphoblastic leukemia in children

Journal

International Journal of Cancer, 152(5)

ISSN

0020-7136

Authors

Gallant, Rachel E

Arroyo, Katti

Metayer, Catherine

et al.

Publication Date

2023-03-01

DOI

10.1002/ijc.34292

Peer reviewed

Received: 3 June 2022 | Revised: 12 August 2022 | Accepted: 2 September 2022

DOI: 10.1002/ijc.34292

CANCER EPIDEMIOLOGY

IJC INTERNATIONAL JOURNAL of CANCER

Associations between early-life and in utero infections and cytomegalovirus-positive acute lymphoblastic leukemia in children

Rachel E. Gallant[1,2] | Katti Arroyo[2] | Catherine Metayer[3] | Alice Y. Kang[3] | Adam J. de Smith[2] | Joseph L. Wiemels[2]

[1]Children's Hospital Los Angeles and Keck School of Medicine, University of Southern California, Los Angeles, California, USA

[2]Center for Genetic Epidemiology, Department of Population and Public Health Sciences, University of Southern California Keck School of Medicine, Los Angeles, California, USA

[3]Division of Epidemiology and Biostatistics, School of Public Health, University of California Berkeley, Berkeley, California, USA

Correspondence
Joseph L. Wiemels, Center for Genetic Epidemiology, Department of Population and Public Health Sciences, University of Southern California Keck School of Medicine, 1450 Biggy Street, NRT-1506, Los Angeles, CA 90089, USA.
Email: wiemels@usc.edu

Funding information
National Institute of Environmental Health Sciences, Grant/Award Numbers: 2R01ES009137-11A1, P50ES018172/EPA RD83615901; National Institute of Health, Grant/Award Numbers: 5T32CA009659-25, R01CA268494

Abstract

Childhood infections and cytomegalovirus (CMV) are associated with pediatric acute lymphoblastic leukemia (ALL). CMV dysregulates the host immune system and alters the immune response to subsequent antigenic exposures. We suspect that this immune dysregulation contributes to increased numbers of symptomatic infections in childhood allowing for expansion of pre-leukemic clones. We explored the association between childhood infections, maternal infections during pregnancy and CMV-positive ALL. Using a droplet digital PCR assay, we screened diagnostic ALL bone marrow samples from the California Childhood Leukemia Study (1995-2015) for the presence of CMV DNA identifying CMV-positive and CMV-negative cases. We performed a case-only analysis (n = 524) comparing the number and types of childhood infections and maternal infections during pregnancy between CMV-positive and CMV-negative ALL cases using logistic regression. With increasing numbers of infections in the first 12 months of life, children were more likely to classify to the highest tertile of CMV DNA in the bone marrow at diagnosis (OR: 1.04, 95% CI: 1.01-1.08). Specifically, those reporting cough or flu in the first 12 months were more likely to be CMV-positive at ALL diagnosis (OR: 2.15, 95% CI: 1.06-4.37 and OR: 2.06, 95% CI: 1.17-3.63 respectively). Furthermore, those with a history of maternal infection during pregnancy were more likely to be CMV-positive (OR: 2.12, 95% CI: 1.24-3.62). We hypothesize that children with underlying immune dysregulation develop more symptomatic infections in childhood and ultimately CMV-positive ALL; this underlying immune dysregulation may be due to early immune system alterations via CMV exposure (in utero or early infancy) proposing a potential link between CMV and ALL etiology.

KEYWORDS
childhood infection, cytomegalovirus, leukemia etiology, maternal infection, pediatric acute lymphoblastic leukemia

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; B-ALL, B-cell acute lymphoblastic leukemia; bp, base pair; CCLS, California Childhood Leukemia Study; CI, confidence interval; CMV, cytomegalovirus; ddPCR, droplet digital polymerase chain reaction; OR, odds ratio; T-ALL, T-cell acute lymphoblastic leukemia.

What's new?

Cytomegalovirus (CMV) infection early in life likely contributes to the development of childhood acute lymphoblastic leukemia (ALL), owing to CMV-induced alterations in immune function. In the present study, the relationship between risk of childhood ALL and exposure to CMV in utero or early life was investigated among children with ALL who participated in the California Childhood Leukemia Study. Case-only analysis shows that children exposed to maternal infections during pregnancy are more likely to have CMV-positive ALL than CMV-negative ALL. Children with increased numbers of infections in early life are more likely to be CMV-positive at the time of ALL diagnosis.

1 | INTRODUCTION

Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy.[1] Due to advances in therapy, outcomes have improved dramatically for childhood ALL. Even so, there are subgroups considered to be high risk who experience poorer outcomes, and children with relapsed disease have much lower survival rates.[1] Furthermore, those who do achieve a long-term remission are still subject to the acute and chronic toxicities associated with chemotherapy.[1] A better understanding of the risk factors associated with ALL etiology (and possibly progression), could allow for the development of interventions for prevention of ALL or even targeted therapies thereby minimizing toxicity of treatment.

Antigenic exposures and childhood infections have been shown to play a role in ALL development such that children with more frequent infections have a higher incidence of childhood ALL.[2-5] Somewhat counterintuitively, daycare attendance (a surrogate for exposure to childhood infections) is associated with *lower* risk of ALL development.[6,7] One hypothesis to explain these seemingly contradictory associations is that children in daycare are exposed to antigens early and often in childhood and thereby develop a more competent immune response; whereas those who do not attend daycare subsequently have delayed infection exposures and in the absence of a well-primed immune system experience a dysregulated immune response thereby increasing their risk of ALL.[2] It is theorized that a dysregulated immune response to childhood infections can provide the opportunity for pre-leukemic clones present at birth to expand, acquire new mutations and proliferate into overt leukemia.[8-14]

Cytomegalovirus (CMV) infection early in life is associated with increased risk of ALL development in childhood.[4,5] CMV is a herpesvirus common in the general population that after acute infection survives in a latent state in a host. To achieve this, CMV interacts with the host immune system thereby altering its response to infection,[14-16] and individuals with latent CMV have also been shown to have aberrant immune responses to other antigens into adulthood.[17-21] In prior studies, we found that approximately half of ALL cases have CMV DNA present in the bone marrow at diagnosis further suggesting that CMV could contribute to ALL development.[4,22] Congenital CMV infection occurs in 0.5% to 1.3% of children in the United States.[23] Severe sequelae such as neurologic deficits, hearing loss and liver dysfunction are features of a congenital CMV infection, but <10% of children with congenital CMV will exhibit these signs and symptoms, and the majority of children are asymptomatic at birth.[23] Although CMV status at birth is not assessed in the present study, we hypothesize that children exposed to CMV in utero or early in life may not have symptoms of CMV in the immediate neonatal period, but go on to have an exaggerated immune response to subsequent infections, leading to more immune dysregulation and thereby higher risk of childhood ALL. In our study, we performed a case-only analysis to explore the association between childhood infection history, maternal infections during pregnancy and the presence of CMV DNA in the bone marrow at the time of ALL diagnosis.

2 | MATERIALS AND METHODS

2.1 | Study design and population

A case-only analysis was performed using data collected from the California Childhood Leukemia Study (CCLS), a case-control study conducted in 35 counties in California between the years 1995 and 2015.[24,25] Children were eligible to participate in the CCLS if they were less than 15 years of age with no prior cancer diagnosis, living in the study region and had an English or Spanish-speaking parent.[24,25] Diagnostic leukemia bone marrow samples were obtained from participating treating hospitals as well as demographic and clinical data.[26] Although the CCLS included both acute myeloid leukemia cases and ALL cases, the present study was limited to ALL given the previously identified associations between childhood infections and ALL.[2,3,8] Of 744 diagnostic bone marrow samples screened for CMV DNA in our prior work,[22] 524 were included in this analysis based on the availability of childhood and maternal infection data (Figure 1). Notably, these samples, as well as the entire study set, are independent of those reported in Francis et al.[4]

2.2 | Cytomegalovirus DNA screening

Diagnostic leukemia bone marrow samples were screened for the presence of CMV DNA using a custom droplet digital PCR (ddPCR) assay as previously described.[23] The QIAamp DNA Blood Mini Kit (Qiagen) was used to isolate DNA from 50 μL bone marrow aspirate

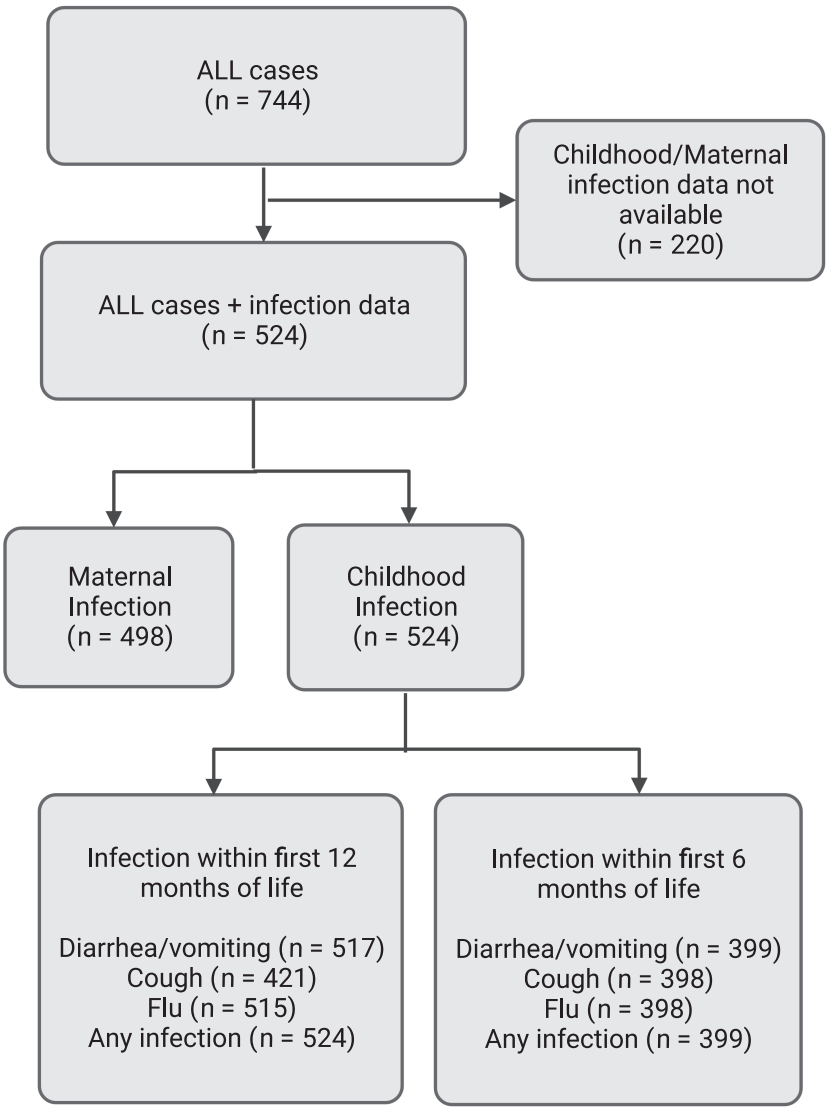

FIGURE 1 Flow chart of selection of cases for inclusion. ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CMV, cytomegalovirus. Created with BioRender.com

samples acquired at diagnosis from 931 childhood leukemia patients including 744 ALL cases. A subset (n = 33) of these cases previously had undergone whole genome sequencing (average ×50), and aligned to the Merlin CMV sequence (Ref Seq ID: NC_006273.2). In two of these samples, a ~700 base pair (bp) segment of CMV was recurrently identified (range: 174 980-175 691), and was chosen as a target for further study. PCR primers and probe were designed to target a segment (range: 175 040-175 279) of this 700 bp CMV sequence (forward primer: agacctcccatagtaaacgc, reverse primer: gctattaccatggtgatgcg, probe: /5HEX/tccccgtgagt/ZEN/caaaccgct/3IABkFQ/). This droplet digital PCR assay was optimized using quantified CMV sequence (Strain: AD-169) (catalog NATCMV-0003, ZeptoMetrix Corp) spiked into normal CMV-negative DNA, to ensure a reproduceable detection level of single copy CMV target in a background of non-CMV DNA. Each patient sample (~50 ng) was screened with this assay. To normalize the CMV positive droplet count based on the amount of DNA in the reaction, a second ddPCR reaction (C-LESS) was run on all samples using primers and probe to detect single copy human DNA target.[27]

Cases were considered CMV-positive if any CMV DNA was detected in the bone marrow samples and CMV-negative if no CMV DNA was detected. A ratio of CMV positive to human haploid target positive droplets (CMV-ratio) was calculated to determine the level of CMV-positivity and normalize for the concentration of DNA in the samples. CMV ratio was then categorized into tertiles based on the CMV-ratio distribution in the overall cohort with a CMV-ratio of 0 defined as negative and the referent for analyses of associations with CMV viral DNA load.

2.3 | Covariates

Demographic and clinical features of the cases were collected as part of the CCLS including age at diagnosis, sex, leukemia subtype, self-reported ethnicity (Latino vs Non-Latino), childhood infections and maternal infections during pregnancy (Table 1). Infection history was reported by the biological mother (or father in the event the mother

TABLE 1 Demographic and clinical features of cytomegalovirus (CMV) positive and negative acute lymphoblastic leukemia (ALL) patients, California Childhood Leukemia Study (CCLS) conducted from 1995 to 2015

	CMV-negative, n = 293 (56%)	CMV-positive, n = 231 (44%)	OR (95% CI)[a]	P[b]
Median age at diagnosis (IQR)	4.46 (2.87, 7.40)	4.84 (3.01, 7.88)		.19
Sex (n = 524)				
Male	168 (57%)	128 (55%)	Ref	
Female	125 (43%)	103 (45%)	1.08 (0.75, 1.55)	.65
Ethnicity (n = 522)				
Non-Latino	151 (52%)	107 (47%)	Ref	
Latino	142 (48%)	122 (53%)	1.21 (0.84, 1.74)	.28
ALL cell type (n = 524)				
T-cell	25 (8%)	16 (7%)	Ref	
B-cell	193 (66%)	167 (72%)	1.35 (0.67, 2.81)	.37
Not specified	75 (26%)	48 (21%)	1.00 (0.46, 2.23)	1.00
B-ALL subtype (n = 360)				
ETV6-RUNX1	67 (35%)	44 (26%)	Ref	
High hyperdiploid	52 (27%)	60 (36%)	1.76 (1.03, 2.98)	.04
Other	74 (38%)	63 (38%)	1.30 (0.78, 2.15)	.31

Note: High hyperdiploid: chromosomal gains resulting in 51 to 67 chromosomes.
Abbreviations: ALL, acute lymphoblastic leukemia; CMV, cytomegalovirus.
[a]OR (95% CI): odds ratio and 95% confidence interval for unadjusted logistic regression analysis.
[b]Wilcoxon *P* for continuous variables, Chi-square *P* for categorical variables and Wald Chi-square *P* when OR (95% CI) are presented.

was unavailable) during a standardized interview. Childhood infections were reported as either categorical (yes/no) or continuous numeric (number of episodes) variables. The type of infection including diarrhea/vomiting, cough and flu was also collected. The CCLS was conducted in phases as the number of participating sites expanded. Importantly, interview questions regarding maternal infections were asked differently in the first three phases (1995-2009) compared to the last phase of the study (2011-2015). In the first three phases of the study, maternal infections were reported as both categorical (yes/no) and continuous numeric (number of episodes) variables while in the last phase only categorical (yes/no) data were collected. For the purposes of our study, the maternal infections during pregnancy data were treated as a binary categorical variable. ALL subtype was categorized as T-cell ALL (T-ALL) or B-cell ALL (B-ALL) and B-ALL was further categorized into high hyperdiploidy (n = 112), *ETV6-RUNX1* (n = 111) and "other" subtypes (n = 137) as previously described.[28]

2.4 | Statistical Methods

Odds ratio (OR) and 95% confidence interval (CI) were estimated for associations between CMV-positivity at the time of ALL diagnosis and early childhood infections and maternal infections during pregnancy using logistic regression. Potential confounders included sex, age at diagnosis, ethnicity, birth order and ALL subtype. Variables were kept in the model as confounders if they caused ≥15% change in the parameter estimate for childhood or maternal infection. Effect modification of ethnicity and ALL subtype were examined for all models using the likelihood ratio test. A *P*-value of <.05 for the likelihood ratio test was considered effect modification. If effect modification was present, models were stratified by the levels of the effect modifier. Statistical analyses were performed using RStudio version 1.2.1335 (*RStudio Team [2020]. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, URL* http://www/rstudio/com).

3 | RESULTS

A total of 524 cases were included in the analysis based on availability of bone marrow DNA available for CMV screening and childhood and/or maternal infection data (Figure 1). Of these, 44% were CMV-positive (Table 1). The CMV-positive and CMV-negative groups did not differ in age at diagnosis, sex or ethnicity (Table 1). As previously described, there was a higher proportion of high hyperdiploid ALL cases in the CMV-positive group (OR: 1.76, 95% CI: 1.03-2.98, *P* = .037, Table 1).[23]

3.1 | Childhood infections and CMV status

With increasing number of infections in the first year of life, children were more likely to be in the highest CMV-tertile (3rd tertile) than to be CMV negative (OR: 1.04, 95% CI: 1.01-1.08, *P* = .022, Table 2, Figure 2). Likewise, those experiencing higher numbers of infections during the first 6 months of life, were more likely to be in the highest CMV-tertile at ALL diagnosis (OR: 1.08, 95% CI: 1.002-1.16, *P* = .044, Table 2). Childhood infections were further categorized into diarrhea/vomiting, cough and flu. We found that children who were reported to have had at least one episode of either cough or flu in the first year of life were greater

TABLE 2 Association of childhood infection history and CMV-status at ALL diagnosis, California Childhood Leukemia Study (CCLS) conducted from 1995 to 2015

Infection type	History of infection (n)	No infection (n)	CMV-positive, OR (95% CI), P[b]	CMV-ratio[a]		
				First tertile, OR (95% CI), P[c]	Second tertile, OR (95% CI), P[c]	Third tertile, OR (95% CI), P[c]
Ever infection (categorical)						
First 12 months of life						
Any type of infection	206	318	1.16 (0.82-1.16) $P = .404$	1.01 (0.60-1.69) $P = .985$	1.06 (0.63-1.78) $P = .819$	1.46 (0.88-2.67) $P = .142$
Diarrhea	97	421	0.93 (0.59-1.47) $P = .770$	0.69 (0.34-1.40) $P = .303$	1.01 (0.53-1.90) $P = .987$	0.92 (0.48-1.77) $P = .809$
Cough	58	363	1.63 (0.94-2.85) $P = .085$	1.42 (0.61-3.33) $P = .417$	1.28 (0.55-2.97) $P = .570$	**2.15 (1.06-4.37)** **$P = .035$**
Flu	111	405	1.43 (0.94-2.19) $P = .095$	1.43 (0.79-2.60) $P = .238$	0.93 (0.48-1.82) $P = .835$	**2.06 (1.17-3.63)** **$P = .013$**
First 6 months of life						
Any type of infection	81	318	1.25 (0.76-2.04) $P = .372$	1.23 (0.60-2.53) $P = .570$	1.00 (0.47-2.15) $P = .995$	1.55 (0.78-3.09) $P = .216$
Diarrhea	46	353	0.86 (0.45-1.61) $P = .644$	0.74 (0.27-2.01) $P = .553$	0.75 (0.28-2.05) $P = .580$	1.10 (0.45-2.67) $P = .830$
Cough	18	378	1.88 (0.76-4.76) $P = .172$	2.04 (0.60-6.87) $P = .252$	2.08 (0.61-7.02) $P = .239$	1.53 (0.40-5.84) $P = .537$
Flu	30	368	1.56 (0.73-3.30) $P = .246$	1.52 (0.53-4.39) $P = .435$	0.91 (0.25-3.28) $P = .891$	2.27 (0.88-5.88) $P = .091$
Number of infections (numeric)						
First 12 months of life						
Any type of infection	181	318	1.03 (1.00-1.07) $P = .060$	1.04 (0.995-1.08) $P = .084$	1.02 (0.96-1.06) $P = .545$	**1.04 (1.01-1.08)** **$P = .022$**
Diarrhea	78	394	1.01 (0.96-1.06) $P = .752$	0.96 (0.86-1.08) $P = .503$	1.02 (0.95-1.09) $P = .576$	1.02 (0.96-1.09) $P = 0.504$
Cough	48	426	1.07 (0.96-1.27) $P = .295$	1.06 (0.89-1.28) $P = 0.550$	0.97 (0.72-1.30) $P = .833$	1.12 (0.96-1.31) $P = .147$
Flu	73	398	1.08 (0.93-1.29) $P = .349$	1.14 (0.94-1.37) $P = .180$	0.78 (0.48-1.26) $P = .309$	1.13 (0.93-1.37) $P = .235$
First 6 months of life						
Any type of infection	73	322	1.05 (0.99-1.12) $P = 0.109$	1.03 (0.94-1.13) $P = .479$	1.03 (0.95-1.13) $P = .458$	**1.08 (1.002-1.16)** **$P = .044$**
Diarrhea	42	353	1.01 (0.97-1.10) $P = 0.345$	1.00 (0.89-1.12) $P = 0.951$	1.04 (0.95-1.13) $P = 0.413$	1.05 (0.97-1.14) $P = .242$
Cough	18	378	1.13 (0.94-1.61) $P = .323$	1.13 (0.82-1.55) $P = .460$	0.97 (0.54-1.74) $P = 0.925$	1.20 (0.90-1.60) $P = .214$
Flu	28	368	1.29 (0.96-1.95) $P = .144$	1.44 (0.96-2.17) $P = .076$	0.79 (0.31-2.01) $P = .624$	1.40 (0.92-2.12) $P = .114$

Note: Odds ratios for categorical variables including Any infection, Diarrhea, Cough and Flu are compared to those without history of infection as the referent group. Number of episodes of infection is treated as continuous variables; odds ratios represent odds of being CMV-positive with increasing number of infections. Associations with $P < .05$ in bold. Not all 524 cases are included in each model due to missing data.

Abbreviations: ALL, acute lymphoblastic leukemia; CMV, cytomegalovirus.

[a]CMV-ratio: ratio of number of CMV-positive droplets to human haploid positive droplets by droplet digital PCR, indicating amount of CMV DNA normalized for total DNA concentration of the sample.

[b]OR (95% CI): odds ratio and 95% confidence interval for unadjusted binomial logistic regression analysis.

[c]OR (95% CI): odds ratio and 95% confidence interval for unadjusted multinomial logistic regression analysis.

than two times more likely to have the highest levels of CMV DNA (highest tertile) at the time of ALL diagnosis (OR: 2.15, 95% CI: 1.06-4.37, $P = .035$ and 2.06, 95% CI: 1.17-3.63, $P = .013$, respectively, Table 2). Age, sex, birth order and ethnicity were not found to be effect modifiers or confounders and therefore are not included in these models. However, ALL subtype was identified as a possible effect modifier, and

therefore we stratified the analysis by subtype including high hyperdiploidy, *ETV6-RUNX1*, T-ALL and B-ALL without specified subtype (Table S1). Although the sample sizes of the strata are small, we did note that among the high hyperdiploid subtype increasing number of childhood infections tends to result in lower odds of being in the highest CMV-tertile at diagnosis (Table S1). Whereas, among *ETV6-RUNX1* cases, unspecified B-ALL subtypes and T-ALL the trend is in the opposite direction with increasing number of infections having higher odds of being strongly CMV-positive (highest CMV-tertile) (Table S1).

3.2 | Maternal infections and CMV status

Data on maternal infection during pregnancy was available for 498 cases (Figure 1), and was found to be associated with CMV-status at the time of ALL diagnosis. Children whose mother's experienced at least one infection during pregnancy were over two times more likely to be in the highest CMV-tertile compared to those with no history of maternal infection during pregnancy (OR: 2.12, 95% CI: 1.24-3.62, $P = .003$, Table 3, Figure 2). Questions regarding maternal infection during pregnancy in the last phase of the study differed from phases 1 to 3. Therefore, stratified analysis by phase of the CCLS was performed, and the results were similar to the overall analysis (Table 3). We further explored the effect of the combination of maternal infections during pregnancy and childhood infections on CMV-status at ALL diagnosis. We found that those with a history of self-reported maternal infection during pregnancy and at least one childhood infection in the first year of life were more likely to be CMV-positive at the time of ALL diagnosis (OR: 1.75, 95% CI: 1.04-2.97, $P = .037$, Table 4). Furthermore, this group was also 2.5 times more likely to be

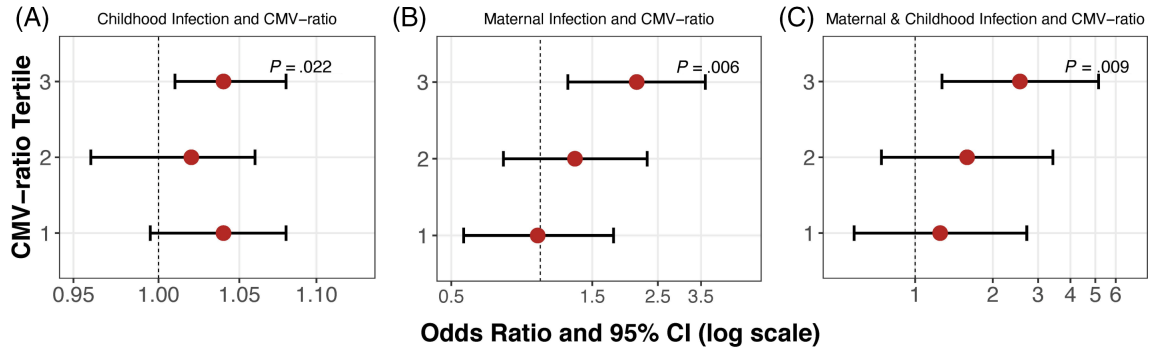

FIGURE 2 Association between childhood infections in the first year of life, maternal infections during pregnancy and CMV-ratio, California Childhood Leukemia Study (1995-2015). CMV, cytomegalovirus. Red dot represents the estimated odds ratio (OR) from a multinomial logistic regression with 95% confidence interval (95% CI) for each OR estimate. Vertical dashed line represents 1 (null estimate) and CIs that include 1 are not significant. Wald Chi-square *P*-values shown for statistically significant OR estimates. (A) Relationship between the number of infections in the first year of life and CMV-ratio tertile among acute lymphoblastic leukemia (ALL) cases (n = 499; cases with childhood infection = 180, cases without childhood infection = 319). (B) Relationship between maternal infection during pregnancy and CMV-ratio tertile among ALL cases (n = 498; cases with maternal infection = 147, cases without maternal infection = 351). (C) Relationship between the combination of maternal and childhood infection and CMV-ratio tertile among ALL cases (n = 497; cases with maternal infection and childhood infection = 75, cases with either maternal infection or childhood infection = 195, cases without either maternal or childhood infection = 227)

TABLE 3 Association of maternal infection during pregnancy and CMV-status at ALL diagnosis, California Childhood Leukemia Study (CCLS) conducted from 1995 to 2015

Maternal Infection during pregnancy	History of infection (n)	No infection (n)	CMV-positive, OR (95% CI), P^b	CMV-ratio[a]		
				First tertile, OR (95% CI), P^c	Second tertile, OR (95% CI), P^c	Third tertile, OR (95% CI), P^c
Overall	147	351	1.42 (0.96-2.09) $P = 0.077$	0.98 (0.55-1.77) $P = .958$	1.31 (0.75-2.30) $P = .342$	**2.12 (1.24-3.62) $P = 0.006$**
Phases 1-3	114	295	1.43 (0.92-2.21) $P = .109$	1.01 (0.51-1.97) $P = .987$	1.21 (0.63-2.31) $P = 0.571$	**2.26 (1.23-4.14) $P = .008$**
Phase 5	33	56	1.18 (0.50-2.83) $P = .714$	0.77 (0.23-2.64) $P = .683$	0.84 (0.24-2.92) $P = .776$	2.39 (0.73-7.79) $P = .149$

Note: Associations with $P < .05$ in bold.

Abbreviations: ALL, acute lymphoblastic leukemia; CMV, cytomegalovirus.

[a]CMV-ratio: ratio of number of CMV-positive droplets to human haploid positive droplets by droplet digital PCR, indicating amount of CMV DNA normalized for total DNA concentration of the sample.

[b]OR (95% CI): odds ratio and 95% confidence interval for unadjusted binomial logistic regression analysis.

[c]OR (95% CI): odds ratio and 95% confidence interval for unadjusted multinomial logistic regression.

TABLE 4 Association of maternal infection during pregnancy, childhood infection and CMV-status at ALL diagnosis, California Childhood Leukemia Study (CCLS) conducted from 1995 to 2015

		CMV-ratio[a]		
Maternal infection during pregnancy and/or childhood infection	**CMV-positive, OR (95% CI), *P*[b]**	**First tertile, OR (95% CI), *P*[c]**	**Second tertile, OR (95% CI), *P*[c]**	**Third tertile, OR (95% CI), *P*[c]**
No maternal or childhood infection (n = 227)	Ref	Ref	Ref	**Ref**
Maternal OR childhood infection (n = 195)	1.06 (0.72-1.57) *P* = .759	1.17 (0.49-1.52) *P* = .607	1.17 (0.66-2.06) *P* = .598	1.21 (0.67-2.15) *P* = .528
Maternal AND childhood infection (n = 75)	**1.75 (1.04-2.97)** ***P* = .037**	1.25 (0.58-2.71) *P* = .575	1.59 (0.74-3.42) *P* = .238	**2.55 (1.27-5.15)** ***P* = .009**

Note: Associations with $P < .05$ in bold.
Abbreviations: ALL, acute lymphoblastic leukemia; CMV, cytomegalovirus.
[a]CMV-ratio: ratio of number of CMV-positive droplets to human haploid positive droplets by droplet digital PCR, indicating amount of CMV DNA normalized for total DNA concentration of the sample.
[b]OR (95% CI): odds ratio and 95% confidence interval for unadjusted binomial logistic regression analysis.
[c]OR (95% CI): odds ratio and 95% confidence interval for unadjusted multinomial logistic regression.

in the highest CMV-tertile (95% CI: 1.27-5.15, $P = .009$, Table 4, Figure 2). The effect of age at diagnosis, sex, birth order and ethnicity were evaluated in all models and not found to be confounders or effect modifiers and therefore are not included in the models. ALL subtype was found to be a possible effect modifier in the associations between maternal infection during pregnancy and CMV-status at ALL diagnosis, therefore the results were stratified by subtype (Tables S2 and S3).

4 | DISCUSSION

We found that children with higher levels of CMV DNA present in bone marrow at the time of ALL diagnosis had a history of more frequent infections during infancy than children without CMV at diagnosis. The types of infections captured by a questionnaire in the CCLS include diarrhea, cough and flu and may not necessarily represent symptomatic CMV infections. However, as CMV can alter the host immune response to subsequent antigenic exposures, we hypothesize that the children with prior CMV infections may experience more symptomatic infections throughout childhood due to a dysregulated immune response. We also suspect that it is by way of this CMV-induced dysregulated immune response that early CMV infection contributes to ALL development.

In the analysis stratified by subtype, high hyperdiploid ALL cases were less likely to be in the highest CMV tertile with increasing number of childhood infections, whereas in the other subtypes the trend was in the opposite direction. This raises the possibility of a heterogenous effect of CMV on ALL development based on subtype. Though the frequency of T-ALL, high hyperdiploid and *ETV6-RUNX1* in our sample was similar to that of the general pediatric ALL population suggesting that this is a representative sample,[29] these results must be interpreted with caution as the sample sizes of the strata are small. Nonetheless, it is intriguing to see that the relationship between childhood infections and CMV status in the high hyperdipoid cases seems to differ from the overall group and the other individual subtypes. In our prior study we found that high hyperdiploid cases were the most likely subtype to have the highest amounts of CMV DNA present in the bone marrow at ALL diagnosis.[23] Though the timing of CMV infection in these cases is not known, it is possible that higher amounts of CMV DNA at ALL diagnosis represent an active infection which raises the possibility that an active CMV infection occurring close to the time of ALL diagnosis could be driving lymphoblast proliferation in the high hyperdiploid subtype; alternatively, high CMV content within high hyperdiploid leukemias may represent a dependency of this subtype on CMV-associated gene expression regardless of the timing of infection. Our current analysis was not able to investigate neonatal CMV infection status. CMV infection later in childhood may not affect responses to common infections prior to diagnosis in the high hyperdiploid subtype (as observed here) but result in leukemic proliferation and development of ALL, by a similar mechanism proposed in Greaves's hypothesis.[8] On the other hand, in the case of T-ALL, ETV6-RUNX1 and uncategorized B-ALL subtypes, CMV may be an early event that leads to dysregulated inflammatory reactions to subsequent antigenic exposures thereby contributing to the development of ALL. The subtype strata in this analysis are likely not appropriately powered to definitively identify associations between childhood infections and CMV-status at diagnosis within individual subtypes and ultimately a larger sample is necessary to appropriately evaluate this association within each subtype. Even so, the results exhibit a possible manifestation of the heterogeneity of effect of CMV. Our analysis also leaves open the question on the etiologic pathways to ALL without CMV involvement, which may be half of our case sample set.

We also found that children whose mothers had at least one infection during pregnancy were more likely to be in the highest CMV tertile at ALL diagnosis. For subjects in phases 1 to 3 of the CCLS, maternal infections during pregnancy were characterized as either pneumonia, flu or cold, and in the last phase of the study the type of infection was not obtained. Given the nonspecific symptomology of an acute CMV infection, high prevalence in the general population,[22] and that most children with congenital CMV are asymptomatic,[24] we

cannot rule out that the maternal infections reported in our study could represent new CMV infections or CMV superinfection.[30] Furthermore, Francis et al detected CMV in the neonatal blood spots of 3% of control subjects which is higher than the reported incidence suggesting that rates of congenital CMV exposure may be higher than what is clinically reported.[4,24] This is particularly interesting considering the previous finding associating maternal CMV infection during pregnancy with childhood ALL.[5] We also found that children with a history of maternal infection during pregnancy and at least one childhood infection were even more likely to be CMV-positive at diagnosis. This could represent a population of children who were exposed to CMV in utero (maternal infection during pregnancy) who then have aberrant immune responses to childhood infections ultimately increasing their risk of developing ALL associated with CMV.

Our study is limited by the fact that the timing of the CMV infection is unknown as we have only detected CMV at the time of diagnosis nor do we have information regarding clinical signs and symptoms of congenital CMV infection, though we suspect that higher levels of CMV DNA could represent an active infection whereas lower levels of CMV DNA could represent a latent infection. Though we have identified an association between increased frequency of childhood infections and CMV-positivity at ALL diagnosis, we acknowledge that early CMV infection may not be the only explanation for this relationship. Those with an increased number of childhood infections could represent a population of children with an underlying immunodeficiency or other immune dysregulation disorder that are more susceptible to childhood infections and potentially hematologic malignancies. Likewise, those with a history of maternal infection during pregnancy and childhood infections could indicate a familial immunodeficiency disorder contributing to the phenotype. We did note, however, that there is not a strong correlation between maternal and childhood infections in this cohort of cases (Table S4), suggesting that it is less likely that CMV-positivity at ALL diagnosis is related to a familial immunodeficiency disorder. Furthermore, ALL subtype is not complete for all of the subjects in the study and the ALL subtype "B-ALL unspecified subtype" is likely a heterogenous group. Given that many B-ALL subtypes have distinct intracellular pathway alterations driving leukemic proliferation, it is possible that CMV may impact these subtypes differently as well. As the collective knowledge of ALL biology is rapidly changing and new subtypes have been identified since the time of the CCLS classification, the limited subtype data in our study impairs our ability to draw conclusions about the effect of CMV on individual ALL subtypes. Future studies using contemporary patient samples may allow for the identification of more complete subtype-specific associations.

Despite its limitations, our study is the first to explore the impact of childhood infections on CMV-positive ALL. Prior studies have shown an association between pediatric ALL and CMV or childhood infections in general.[3-5] Our results generate a hypothesis for a possible mechanism relating these two risk factors; that is the dysregulation of the immune response caused by early CMV exposure leads to an increased number of symptomatic childhood infections and ultimately proliferation of leukemic lymphoblasts resulting in CMV-positive ALL. A better understanding of the effect of CMV on ALL development could lead to opportunities to alter therapy and develop strategies for the prevention of CMV-associated ALL, and the findings from our study certainly inspire further investigation into the mechanism by which CMV increases the risk of childhood ALL.

AUTHOR CONTRIBUTIONS

Rachel E. Gallant: Designed the study, performed laboratory experiments, analyzed data, created figures and wrote the primary and subsequent versions of the article. **Katti Arroyo:** Performed experiments and helped prepare the article. **Catherine Metayer:** Conducted the CCLS study and managed the collection of data and samples; helped refine the study and assisted in article preparation. **Alice Y. Kang:** Assisted with data extraction and article preparation. **Adam J. de Smith:** Helped refine the study and assisted in article preparation. **Joseph L. Wiemels:** Designed the research and assisted in writing the article. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The authors thank the families that participated in the California Childhood Leukemia Study (CCLS). Without their time and effort, none of our studies would have been possible. For recruitment of subjects enrolled in the CCLS, the authors gratefully acknowledge the clinical investigators at the following collaborating hospitals: University of California Davis Medical Center (Dr. Jonathan Ducore), University of California San Francisco (Drs. Mignon Loh and Katherine Matthay), Children's Hospital of Central California (Dr. Vonda Crouse), Lucile Packard Children's Hospital (Dr. Gary Dahl), Children's Hospital Oakland (Drs. James Feusner and Carla Golden), Kaiser Permanente Roseville (formerly Sacramento) (Drs. Kent Jolly and Vincent Kiley), Kaiser Permanente Santa Clara (Drs. Carolyn Russo, Alan Wong and Denah Taggart), Kaiser Permanente San Francisco (Dr. Kenneth Leung), Kaiser Permanente Oakland (Drs. Daniel Kronish and Stacy Month), California Pacific Medical Center (Dr. Louise Lo), Cedars-Sinai Medical Center (Dr. Fataneh Majlessipour), Children's Hospital Los Angeles (Dr. Cecilia Fu), Children's Hospital Orange County (Dr. Leonard Sender), Kaiser Permanente Los Angeles (Dr. Robert Cooper), Miller Children's Hospital Long Beach (Dr. Amanda Termuhlen), University of California, San Diego Rady Children's Hospital (Dr. William Roberts) and University of California, Los Angeles Mattel Children's Hospital (Dr. Theodore Moore).

FUNDING INFORMATION

This work was supported by the National Institute of Health [5T32CA009659-25 to R.G., and R01CA268494 to JW], the U.S. National Institute of Environmental Health Sciences of the National Institute of Health [2R01ES009137-11A1, P50ES018172/EPA RD83615901 to C.M.] and the Impact Grant via University of California San Francisco Cancer Center (J.L.W.). We acknowledge additional funding from University of Southern California (USC) Center for Genetic Epidemiology, USC Cancer Center CCSG Development

Funds, USC Cancer Center Norris Chair Funds and the Shirley McKernan Courage Foundation (Children's Hospital Los Angeles).

CONFLICT OF INTEREST

The authors have no possible conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of our study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The study was approved by the California Department of Health Services and the Institutional Review Boards of the participating hospitals, the University of California Berkeley and San Francisco and the University of Southern California and was conducted in accordance with the principles of the Declaration of Helsinki.

ORCID

Rachel E. Gallant https://orcid.org/0000-0003-0860-7184
Adam J. de Smith https://orcid.org/0000-0003-4880-7543
Joseph L. Wiemels https://orcid.org/0000-0003-4838-9951

REFERENCES

1. Orkin S, Fisher D, Ginsburg D, Look A, Lux S, Nathan D. *Nathan and Oski's Hematology and Oncology of Infancy and Childhood*. Vol 2. 8th ed. Philadelphia, PA: Elsevier; 2015.
2. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006;6(3):193-203.
3. Chang JS, Tsai CR, Tsai YW, Wiemels JL. Medically diagnosed infections and risk of childhood leukaemia: a population-based case-control study. *Int J Epidemiol*. 2012;41(4):1050-1059.
4. Francis SS, Wallace AD, Wendt GA, et al. In utero cytomegalovirus infection and development of childhood acute lymphoblastic leukemia. *Blood*. 2017;129(12):1680-1684.
5. Wiemels JL, Talbäck M, Francis S, Feychting M. Early infection with cytomegalovirus and risk of childhood hematologic malignancies. *Cancer Epidemiol Prev Biomark*. 2019;28(6):1024-1027.
6. Ma X, Buffler PA, Selvin S, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. *Br J Cancer*. 2002;86(9):1419-1424.
7. Gilham C, Peto J, Simpson J, et al. Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from UK case-control study. *BMJ*. 2005;330(7503):1294.
8. Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat Rev Cancer*. 2018;18(8):471-484.
9. Wiemels JL, Cazzaniga G, Daniotti M, et al. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet*. 1999;354(9189):1499-1503.
10. Nielsen AB, Zhou M, de Smith AJ, et al. Increased neonatal level of arginase 2 in cases of childhood acute lymphoblastic leukemia implicates immunosuppression in the etiology. *Haematologica*. 2019;104(11):e514-e516.
11. Chang JS, Zhou M, Buffler PA, Chokkalingam AP, Metayer C, Wiemels JL. Profound deficit of IL10 at birth in children who develop childhood acute lymphoblastic leukemia. *Cancer Epidemiol Prev Biomark*. 2011;20(8):1736-1740.
12. Hauer J, Martín-Lorenzo A, Sánchez-García I. Infection causes childhood leukemia. *Aging*. 2015;7(9):607-608.
13. Rudant J, Lightfoot T, Urayama KY, et al. Childhood acute lymphoblastic leukemia and indicators of early immune stimulation: a childhood leukemia international consortium study. *Am J Epidemiol*. 2015;181(8):549-562.
14. Schraw JM, Bailey HD, Bonaventure A, et al. Infant feeding practices and childhood acute leukemia: findings from the Childhood Cancer & Leukemia International Consortium. *Int J Cancer*. 2022;151(7):1013-1023.
15. Picarda G, Benedict CA. Cytomegalovirus: shape-shifting the immune system. *J Immunol*. 2018;200(12):3881-3889.
16. La Rosa C, Diamond DJ. The immune response to human CMV. *Future Virol*. 2012;7(3):279-293.
17. Miller-Kittrell M, Sparer TE. Feeling manipulated: cytomegalovirus immune manipulation. *Virol J*. 2009;6(1):4.
18. Wang C, Liu Y, Xu LT, et al. Effects of aging, CMV infection, and EBV infection on human B cell repertoires. *J Immunol*. 2014;192(2):603-611.
19. Furman D, Jojic V, Sharma S, et al. Cytomegalovirus infection improves immune responses to influenza. *Sci Transl Med*. 2015;7:281-243.
20. Soderberg C, Sumitran-Karuppan S, Ljungman P, Moller E. CD13-specific autoimmunity in cytomegalovirus-infected immunocompromised patients. *Transplantation*. 1996;61(4):594-600.
21. Frasca D, Diaz A, Romero M, Landin AM, Blomberg BB. Cytomegalovirus (CMV) seropositivity decreases B cell responses to the influenza vaccine. *Vaccine*. 2015;33(12):1433-1439.
22. 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(9):1143-1151.
23. Gallant RE, Arroyo K, Bracci PM, et al. Clinical characteristics of cytomegalovirus-positive pediatric acute lymphoblastic leukemia at diagnosis. *Am J Hematol*. 2022;97:E198-E201.
24. Jenks CM, Hoff SR, Mithal LB. Congenital cytomegalovirus infection: epidemiology, timely diagnosis, and management. *NeoReviews*. 2021;22(9):e606-e613.
25. Metayer C, Zhang L, Wiemels JL, et al. Tobacco smoke exposure and the risk of childhood acute lymphoblastic and myeloid leukemias by cytogenetic subtype. *Cancer Epidemiol Biomark Prev*. 2013;22(9):1600-1611.
26. de Smith AJ, Kaur M, Gonseth S, et al. Correlates of prenatal and early-life tobacco smoke exposure and frequency of common gene deletions in childhood acute lymphoblastic leukemia. *Cancer Res*. 2017;77(7):1674-1683.
27. Weisenberger DJ, Trinh BN, Campan M, et al. DNA methylation analysis by digital bisulfite genomic sequencing and digital MethyLight. *Nucleic Acids Res*. 2008;36(14):4689-4698.
28. Aldrich MC, Zhang L, Wiemels JL, et al. Cytogenetics of Hispanic and White children with acute lymphoblastic leukemia in California. *Cancer Epidemiol Prev Biomark*. 2006;15(3):578-581.
29. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet*. 2019;51(2):296-307.
30. Boppana SB, Rivera LB, Fowler KB, Mach M, Britt WJ. Intrauterine transmission of cytomegalovirus to infants of women with preconceptional immunity. *N Engl J Med*. 2001;344(18):1366-1371.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gallant RE, Arroyo K, Metayer C, Kang AY, de Smith AJ, Wiemels JL. Associations between early-life and in utero infections and cytomegalovirus-positive acute lymphoblastic leukemia in children. *Int J Cancer*. 2023;152(5):845-853. doi:10.1002/ijc.34292