Ethnic Difference in Daycare Attendance, Early Infections, and Risk of Childhood Acute Lymphoblastic Leukemia

Xiaomei Ma,[1] Patricia A. Buffler,[2] Joseph L. Wiemels,[3] Steve Selvin,[2] Catherine Metayer,[2] Mignon Loh,[3] Monique B. Does,[2] and John K. Wiencke[3]

[1]Yale University School of Medicine, New Haven, Connecticut; [2]School of Public Health, University of California, Berkeley, California; and [3]School of Medicine, University of California, San Francisco, California

Abstract

A role for infectious agents has been proposed in the etiology of childhood acute lymphoblastic leukemia (ALL), particularly for common ALL (c-ALL; ALL diagnosed in children ages 2-5 years and expressing CD10 and CD19 surface antigens). We evaluated the possible etiologic role of daycare attendance (a proxy measure for exposure to infectious agents) and infections during infancy in the Northern California Childhood Leukemia Study. A total of 294 incident ALL cases (ages 1-14 years) and 376 individually matched controls were included in this analysis. In non-Hispanic White children, daycare attendance measured by child-hours was associated with a significantly reduced risk of ALL. Compared with children who did not attend any daycare, the odds ratio (OR) for those who had >5,000 child-hours during infancy was 0.42 [95% confidence interval (95% CI), 0.18-0.99] for ALL and 0.33 (95% CI, 0.11-1.01) for c-ALL. Test for trend is also significant, which supports a dose-response relationship. The magnitude of effect associated with the same number of child-hours was stronger for daycare attendance during infancy than for daycare attendance before diagnosis. In addition, self-reported ear infection during infancy was associated with a significantly reduced risk of c-ALL (OR, 0.32; 95% CI, 0.14-0.74) in non-Hispanic White children. In Hispanic children, no association was observed among daycare attendance, early infections, and risk of childhood ALL or c-ALL. These results offer indirect yet strong support for the infectious disease hypothesis in the etiology of ALL in non-Hispanic White children and highlight an important ethnic difference. (Cancer Epidemiol Biomarkers Prev 2005;14(8):1928–34)

Introduction

Infectious agents have long been suspected of playing a role in the etiology of childhood leukemia, especially acute lymphoblastic leukemia (ALL). Greaves hypothesized that delayed exposure to common infections leads to an inadequately stimulated immune system and a subsequently increased risk of childhood leukemia, particularly common ALL (c-ALL; ALL diagnosed in children ages 2-5 years and expressing both CD10 and CD19 surface antigens), which has a maximum incidence in 2- to 5-year-olds (1, 2). Alternatively, Kinlen suggested that childhood leukemia might result from a rare response to a common infection(s), which has not been identified to date (3). Increased risks would occur when populations from different geographic areas were mixed resulting in an increased level of contact between infected and susceptible individuals (3). Testing these hypotheses has been challenging, mainly because it is difficult to directly measure exposure or delayed exposure to infectious agents.

In developed countries, most exposure to common childhood infections results from contact with other children, and daycare attendance has been suggested as an indirect measure of early exposure to infectious agents (4). It has been well documented that daycare attendance increases the risk of infections and the risk of infections increases with the number of children in a group (5-7). In the Northern California Childhood Leukemia Study (NCCLS), a comprehensive questionnaire was designed to collect relevant data, including daycare and preschool (''daycare'') attendance and the occurrence of several types of infections during infancy (i.e., 0-11 months) to assess the infectious disease hypothesis. An earlier analysis of the NCCLS data (140 ALL cases and 140 controls) has indicated that extensive contact with other children in a daycare setting is associated with a reduced risk of ALL (8). This current report expands the previous research by including a larger number of subjects (294 ALL cases and 376 controls), further characterizing ALL using flow cytometry characteristics, studying the timing of daycare attendance, and adding early infections to the analysis. In addition, a possible difference in risk between Hispanic and non-Hispanic White children is evaluated.

Materials and Methods

Study Design. The NCCLS commenced in 1995 and is currently ongoing. The study area includes 17 counties in the Greater San Francisco Bay Area (1995 to present) and in 1999 was expanded to include 18 additional counties in northern and central California. In the NCCLS, incident cases of newly diagnosed childhood leukemia (ages 0-14 years) are rapidly ascertained from major pediatric clinical centers, usually within 72 hours after diagnosis. Although case ascertainment is hospital-based, a comparison with all population-based cases ascertained by the statewide California Cancer Registry (2000) confirms that the NCCLS protocol successfully identified 95% of all age-eligible newly diagnosed childhood leukemia cases among residents of the five-county San Francisco metropolitan area and 76% of such cases in the other 30 counties. Controls are randomly selected from the

Received 2/16/05; revised 5/6/05; accepted 5/23/05.

Grant support: National Institute of Environmental Health Sciences grants R01 ES09137 and PS42 ES04705 and Hope Street Kids (X. Ma). J.L. Wiemels is a scholar of the Leukemia and Lymphoma Society.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Xiaomei Ma, Department of Epidemiology and Public Health, Yale University School of Medicine, 60 College Street, New Haven, CT 06520-8034. Phone: 203-785-3880; Fax: 203-785-6980. E-mail: xiaomei.ma@yale.edu

doi:10.1158/1055-9965.EPI-05-0115

statewide birth certificate files maintained by the Center for Health Statistics in the California Department of Health Services and individually matched to cases on date of birth, sex, mother's race (White, African American, or other), and Hispanic status (a child is considered Hispanic if either parent is Hispanic). For cases diagnosed before December 1, 1999, the case-control ratio was 1:1, and mother's county of residence at the time of child's birth was used as an additional matching criterion. For cases diagnosed on or after December 1, 1999, the case-control ratio is 1:2. A detailed protocol for control selection has been reported elsewhere (9).

To be eligible, each case or control had to (*a*) reside in the study area; (*b*) be <15 years at the reference date (date of diagnosis for cases and the corresponding date for matched controls); (*c*) have at least one parent or guardian who speaks English or Spanish; and (*d*) have no previous history of malignancy. Approximately 86% of the eligible cases consented to participate. The consent rate did not differ by language spoken (English or Spanish). To enroll the 376 controls included in the study, a total of 806 potential controls were searched (yield = 376 / 806 = 46.6%), among which 139 (17.2%) could not be located and 126 (14.6%) refused without providing eligibility information. Of the 541 potential controls who were successfully located and whose eligibility was assessed, 448 (82.8%) were eligible. If the same percentage of eligible subjects among those whose eligibility was assessed (n = 541) was assumed in those whose eligibility could not be assessed (n = 139 + 126 = 265), a total of 667 potential controls were presumed eligible [n = 448 + (265 × 82.8%) = 667]. Therefore, the overall control participation rate in the study was ~56% (376 / 667 = 56.4%).

To enroll 172 non-Hispanic White controls, a total of 346 potential controls were searched; to enroll 153 Hispanic controls, a total of 349 potential controls were searched. The yield in non-Hispanic White controls (172 / 346 = 49.7%) was somewhat higher than the yield in Hispanic controls (153 / 349 = 43.8%), but the two proportions are not statistically significant (P = 0.12).

The study protocol was approved by the institutional review boards of the University of California-Berkeley and all collaborating institutions, and a written informed consent was obtained for all participating subjects.

Study Population and Data Collection. A personal interview with the primary caretaker of each case or control subject, usually the biological mother, was scheduled as soon as consent was obtained. By the end of 2002, a total of 382 cases and 482 matched controls had been interviewed in the NCCLS. Among the 864 subjects, 25 (12 cases and 13 controls) were excluded because the respondents were not biological mothers of the index children (n = 17) or the biological mothers could not recall the details of their children's daycare attendance (n = 8). Among the remaining 839 subjects, there were 270 matched pairs (1 case and 1 control) and 92 matched triplets (1 case and 2 controls). After excluding non-ALL cases (n = 55) as well as ALL cases diagnosed during infancy (which is believed to have a distinct etiology; n = 13), a total of 294 ALL cases who were diagnosed at age ≥12 months and 376 matched controls (212 pairs and 82 triplets) were included in the present analysis.

Daycare attendance data were censored on the reference date (date of diagnosis for cases and corresponding date for matched controls), or at age 6 years, whichever occurred first. For this analysis, the following variables were included or constructed: (*a*) history of attending any daycare facilities (yes/no); (*b*) age when first started daycare (in months); (*c*) duration of stay at all daycare facilities (in months) regardless of hours attended; (*d*) mean hours per week attending daycare facilities; (*e*) mean number of other children in attendance at each daycare facility; and (*f*) total child-hours, which is an overall summary of daycare attendance. Child-hours at each daycare is calculated as follows: number of months attending a daycare × mean hours per week at this daycare × number of other children at this daycare × 4.35 (i.e. number of weeks per month). The child-hours at each daycare were added to obtain the total child-hours for each child. For children who never attended daycare, an age of 72 months was assigned as the age when first started daycare, and a value of 0 was assigned as duration of stay, mean hours per week, mean number of children, total number of children, and total child-hours.

To eliminate the possibility that leukemia-related symptoms may have affected daycare attendance just before diagnosis, daycare attendance data were also censored 1 year before the reference date. Furthermore, for the purpose of assessing the role of early exposure to infectious agents, we also censored daycare attendance on age 1 year and calculated the total child-hours for daycare attendance during infancy.

Data on the number of other children in household before the index child went to first grade were also collected. This variable is a potential indicator of exposure to infectious agents by contact with other children in the same household. In addition, the respondents were asked about child's infection history during infancy, including severe diarrhea/vomiting, ear infection, persistent cough, mouth infection, eye infection, influenza, and unspecified ''other infections.''

Case Characterization. In the NCCLS, clinical records of all cases were abstracted, usually within several months after diagnosis. Medical data abstracted included findings from pretreatment bone marrow aspirate, bone marrow biopsy, flow cytometry, and cytogenetics. Abstracted data were reviewed by a clinical hematologist/oncologist to ensure accuracy and completeness. These data were used to classify cases. In the present study, we not only analyzed ALL as a whole but also evaluated its major subtypes c-ALL and non-c-ALL. c-ALL is defined as ALL diagnosed in children ages 2 to 5 years and expressing both CD10 and CD19 surface antigens (10), and all other ALL is classified as non-c-ALL.

Statistical Analysis. Pearson's χ^2 test was used to compare demographic factors, socioeconomic characteristics, and daycare attendance between cases and controls as well as daycare attendance (ever versus never) between Hispanic and non-Hispanic White children. Nonparametric Wilcoxon rank-sum test was used to compare characteristics of daycare attendance (which were measured as continuous variables and did not seem to be normally distributed) between cases and controls of the same ethnicity as well as between Hispanic and non-Hispanic White children of the same case-control status. A quantile-quantile plot, which is an important graphic technique for comparing shapes of distributions (11), was used to illustrate the difference in total child-hours between cases and controls. Data were plotted separately for Hispanic and non-Hispanic White children. Conditional logistic regression models were used to estimate odds ratios (OR) and construct 95% confidence intervals (95% CI) as approximations of relative risks, adjusting for maternal education and annual household income. To avoid the loss of subjects with unknown household income data in analyses, the reported median annual household income of the study population (US $30,000-44,900) was assigned to 19 subjects (8 cases and 11 controls).

Results

Demographic, socioeconomic, and selected birth characteristics of cases and controls are presented in Table 1. Control families had a significantly higher household income than cases (P = 0.001), and mothers of controls had higher education levels than mothers of cases (P = 0.03). Maternal age at the time of the index child's birth appeared older in cases than in controls, but

the difference was not statistically significant. Cases and controls had similar birth weight and were of similar birth order. The difference in duration of breast-feeding between cases and controls was not statistically significant.

A slightly higher percentage of controls (67.6%) than cases (62.2%) attended daycare before the reference date. Similar difference was observed in both Hispanic and non-Hispanic White children. None of the differences reached statistical difference. However, in both cases and controls, significantly less Hispanic children attended daycare than their non-Hispanic White counterparts (Table 2). Compared with the cases, controls started daycare at a younger age, stayed at daycare for longer duration (more months of stay), spent more hours at daycare facilities per week, were exposed to more children at each daycare facility, and had more child-hours. Although not all the individual tests were statistically significant, a consistent pattern was observed (Table 3).

Table 1. Characteristics of 294 noninfant ALL cases and 376 matched controls, the NCCLS, 1995 to 2002

Characteristics	Cases *n* (%)	Controls *n* (%)	*P* (χ^2 test)
Age*			
12-23.9 mo	26 (8.8)	34 (9.0)	—
2-5 y	179 (60.9)	232 (61.7)	
6-10 y	60 (20.4)	72 (19.2)	
11-14 y	29 (9.9)	38 (10.1)	
Mean (SE), y	5.5 (0.2)	5.4 (0.2)	
Gender*			
Male	152 (51.7)	192 (51.1)	—
Female	142 (48.3)	184 (48.9)	
Race/ethnicity†			
Hispanic	120 (40.8)	153 (40.7)	—
Non-Hispanic White	136 (46.3)	172 (45.7)	
Non-Hispanic Black	6 (2.0)	9 (2.4)	
Other	32 (10.9)	42 (11.2)	
Household income ($/y)			
<15,000	40 (13.6)	35 (9.3)	0.001
15,000-29,900	59 (20.1)	54 (14.4)	
30,000-44,900	56 (19.0)	55 (14.6)	
45,000-59,900	44 (15.0)	48 (12.8)	
60,000-74,900	31 (10.5)	49 (13.0)	
≥75,000	64 (21.8)	135 (35.9)	
Maternal education			
High school or less	130 (44.2)	130 (34.6)	0.03
High school or higher but less than bachelor's	82 (27.9)	135 (35.9)	
Bachelor's or higher	82 (27.9)	111 (29.5)	
Maternal age (y)			
<25	102 (34.7)	103 (27.4)	0.12
25-34	153 (52.0)	220 (58.5)	
≥35	39 (13.3)	53 (14.1)	
Mean (SE)	28.0 (0.3)	28.8 (0.3)	
Birth weight (g)			
<2,500	18 (6.1)	20 (5.3)	0.90
2,500-3,999	223 (75.8)	288 (76.6)	
≥4,000	53 (18.0)	68 (18.1)	
Mean (SE)	3436 (35)	3475 (30)	
Birth order			
First	123 (41.8)	146 (38.8)	0.80
Third	52 (17.7)	65 (17.3)	
Fourth or higher	23 (7.8)	35 (9.3)	
Mean (SE)	1.99 (0.07)	2.04 (0.06)	
Duration of breast-feeding (mo)			
0 (never breast-fed)	56 (19.0)	59 (15.7)	0.45
>0 and <3	68 (23.1)	74 (19.7)	
3-5	48 (16.3)	75 (20.0)	
6-11	61 (20.8)	87 (23.1)	
≥12	61 (20.8)	81 (21.5)	
Mean (SE)	6.25 (0.47)	6.77 (0.40)	

*These are matching variables. Age is age at the reference date (date of diagnosis for cases and the corresponding date for matched controls).

†Controls were individually matched to cases on maternal race (White, African American, or other) and parental Hispanic status as recorded on birth certificates. The race/ethnicity information presented was reported by the biological mothers at interview.

When the analyses were stratified by ethnicity, it was clear that the relationship between daycare attendance and risk of childhood ALL was not the same in Hispanic and non-Hispanic White children. Furthermore, in both cases and controls, Hispanic and non-Hispanic White children were significantly different with respect to age first started daycare, months of stay, mean number of other children at each facility, total child-hours, and number of other children in household before the index child went to first grade (Table 3).

The quantile-quantile plot (cases versus controls) clearly showed that non-Hispanic White cases had fewer total child-hours than matched controls (Fig. 1). If there were no difference in the distribution of total child-hours between non-Hispanic White cases and controls, the points would have randomly deviated from the dotted reference line as opposed to clustering in one side of the reference line. The points on the quantile-quantile plot for Hispanic cases and controls were on both sides of the reference line (Fig. 1).

Due to the striking differences observed between Hispanic and non-Hispanic White subjects (Tables 2 and 3; Fig. 1), the infectious disease hypothesis is tested separately in the two ethnic groups. In our previously published analysis of daycare attendance and risk of childhood ALL (8), child-hours was the dominant predictor of case-control status, and other variables measuring daycare attendance add little additional independent information to the evaluation of association between daycare attendance and the risk of ALL. We confirmed this in the present analysis (data not shown). After the confirmation, child-hours was used as the primary measure for estimating ORs and constructing 95% CIs (Table 4).

In non-Hispanic White children, there was a significant dose-response relationship between child-hours censored on the reference date and a reduced risk of childhood ALL (*P* for trend = 0.03) or c-ALL (*P* for trend = 0.01; Table 4). Comparing with children who did not attend any daycare facilities, the OR for children who had ≥15,000 child-hours was 0.28 (95% CI, 0.08-0.95) for c-ALL. A child would have 15,000 child-hours if he or she attended a daycare facility with 10 other children for 20 hours per week for ~17.3 months. The association between daycare attendance and the risk of c-ALL persisted when daycare attendance was censored 1 year before the reference date (Table 4). An association in the same direction was also observed for child-hours censored at age 1 year. In addition, the magnitude of effect associated with the same number of child-hours was stronger for daycare attendance during infancy than for daycare attendance before the reference date. The OR associated with each thousand child-hours was 0.880 (95% CI, 0.786-0.985) and 0.978 (95% CI, 0.961-0.996) for attendance during infancy and attendance before the reference date, respectively. Five thousand child-hours during infancy correspond to an OR of $(0.880)^5 = 0.53$ (95% CI, 0.30-0.93), whereas the same level of exposure before the reference date corresponds to an OR of $(0.978)^5 = 0.89$ (95% CI, 0.82-0.98).

Daycare attendance did not seem to be associated with the risk of non-c-ALL in non-Hispanic White children or the risk of any type of ALL (total, c-ALL, or non-c-ALL) in Hispanic children (Table 4).

In non-Hispanic White children, self-reported ear infection during infancy was associated with a significantly reduced risk of c-ALL (OR, 0.32; 95% CI, 0.14-0.74). The ORs associated with having 1 to 4 and ≥5 infections during infancy (seven different types combined) were <1 for c-ALL among non-Hispanic White children, but both 95% CIs indicated that the observed results may be due to random variation (Table 5). None of the other kinds of infections during infancy (severe diarrhea/vomiting, persistent cough, mouth infection, eye infection, influenza, or unspecified "other infections") seemed to be associated with the risk of ALL in non-Hispanic White children (Table 5; some data not shown). There was also no

Table 2. Daycare attendance censored at different time points by case-control status and ethnicity, the NCCLS, 1995 to 2002

Total no. cases/controls	Attended daycare		P^* (cases vs controls)
	Cases (%)	Controls (%)	
Censored on the reference date†			
Total (294/376)	183 (62.2)	254 (67.6)	0.15
Hispanic (120/153)	63 (52.5)	86 (56.2)	0.54
Non-Hispanic White (136/172)	99 (72.8)	137 (79.6)	0.16
P^* (Hispanic vs White)	0.001	<0.001	
Censored 1 y before the reference date†			
Total (294/376)	152 (51.7)	211 (56.1)	0.25
Hispanic (120/153)	54 (45.0)	66 (43.1)	0.76
Non-Hispanic White (136/172)	82 (60.3)	116 (67.4)	0.19
P (Hispanic vs White)	0.01	<0.001	
Censored at age 1 y†			
Total (294/376)	62 (21.1)	94 (25.0)	0.23
Hispanic (120/153)	22 (18.3)	21 (13.7)	0.30
Non-Hispanic White (136/172)	34 (25.0)	56 (32.6)	0.15
P (Hispanic vs White)	0.20	<0.001	

*Ps are from χ^2 tests.
† "Censored" at a specific point in time indicates that exposure occurred after that point is not considered. The "reference date" is the date of diagnosis for cases and the corresponding date for matched controls.

indication of any association between infections and the risk of ALL in Hispanic children.

Adjusting for history of breast-feeding or the number of other children living in household before the index child went to first grade had no significant impact on the results of the analyses described above.

Discussion

In the present study, daycare attendance and ear infection during infancy were associated with a significantly reduced risk of ALL, especially c-ALL, in non-Hispanic White children. This observation offers support to the infectious disease hypothesis regarding the etiology of childhood ALL. In Hispanic children, however, no evidence supporting this hypothesis was observed. It is not clear whether the observed ethnic difference could be due to cultural and environmental differences or biological characteristics.

Four previous studies have reported a statistically significant protective effect of daycare attendance (12-15). In a Greek study, childhood leukemia cases in metropolitan Athens were less likely than controls to have attended daycare, especially for

Table 3. Characteristics of daycare attendance (censored on the reference date) and other children living in the household by case-control status and ethnicity, the NCCLS, 1995 to 2002

Variables	Cases, mean (median)*	Controls, mean (median)*	P † (cases vs controls)
Age first started daycare (mo)			
Total	43.1 (42.0)	39.5 (37.5)	0.08
Hispanic	48.5 (54.0)	48.9 (49.0)	0.96
Non-Hispanic White	37.0 (30.0)	31.6 (30.0)	0.06
P† (Hispanic vs White)	<0.001	<0.001	
Months of stay (regardless of hours)			
Total	15.8 (8.5)	16.6 (11.0)	0.25
Hispanic	11.4 (3.0)	9.4 (3.0)	0.77
Non-Hispanic White	20.7 (12.5)	22.4 (18.8)	0.32
P† (Hispanic vs White)	<0.001	<0.001	
Mean hours per week			
Total	12.6 (7.5)	16.1 (10.0)	0.01
Hispanic	12.5 (6.8)	13.8 (8.0)	0.43
Non-Hispanic White	12.8 (8.5)	18.1 (13.7)	0.01
P† (Hispanic vs White)	0.21	<0.01	
Mean no. other children at each daycare			
Total	8.2 (7.0)	9.4 (9.0)	0.10
Hispanic	6.7 (3.8)	8.3 (5.0)	0.25
Non-Hispanic White	9.6 (9.4)	10.6 (10.1)	0.26
P† (Hispanic vs White)	<0.01	<0.01	
Total thousand child-hours			
Total	19.3 (2.9)	25.4 (7.1)	0.07
Hispanic	15.2 (1.9)	16.6 (2.1)	0.90
Non-Hispanic White	22.9 (8.0)	32.0 (12.3)	0.06
P† (Hispanic vs White)	0.01	<0.001	
No. other children in household before the index child went to first grade			
Total	2.24 (2.00)	1.98 (2.00)	0.74
Hispanic	2.90 (2.00)	2.48 (2.00)	0.42
Non-Hispanic White	1.74 (1.00)	1.63 (1.00)	0.72
P† (Hispanic vs White)	<0.001	<0.001	

*None of the variables seemed to be normally distributed, hence the decision to present both means and medians.
† Ps are from nonparametric Wilcoxon rank-sum tests.

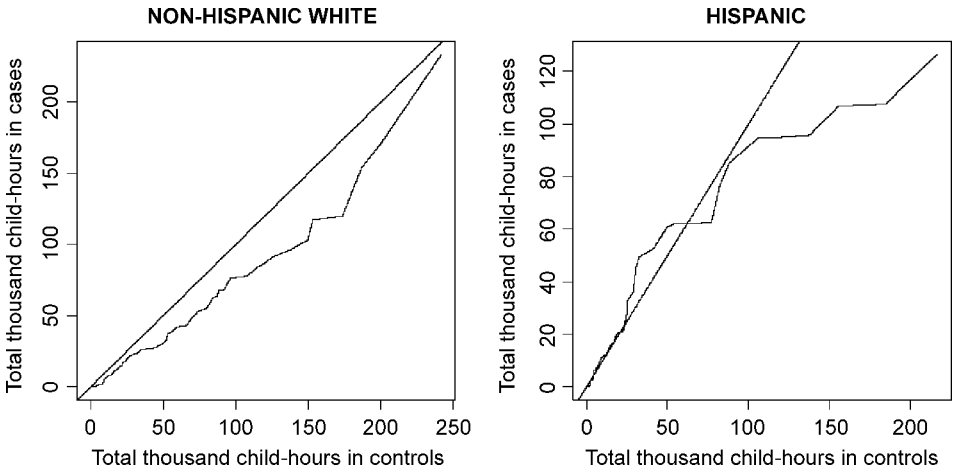

Figure 1. Quantile-quantile plot of child-hours by case-control status and ethnicity.

attendance of at least 3 months before age 2 years (12). A significantly protective effect of daycare attendance was observed in a population-based Canadian study with ALL cases ages 0 to 9 years. Compared with children who never attended daycare, the ORs for children who started daycare at ≤2 and >2 years were 0.49 (95% CI, 0.31-0.77) and 0.67 (95% CI, 0.45-1.01), respectively (13). In a French study, a significantly lower percentage of childhood acute leukemia cases ages >2 years than controls attended any daycare, and an early age at start of daycare was associated with a reduced risk of childhood acute leukemia (14). In another French study, the OR associated with daycare attendance was 0.7 (95% CI, 0.6-1.0) for ALL and 0.8 (95% CI, 0.6-1.0) for c-ALL, respectively, and an effect of early age at start of daycare (<3 months) was observed (15).

Several other studies did not find a statistically significant association between daycare attendance and childhood leukemia. The OR for history of daycare attendance was 0.83 (95% CI, 0.51-1.37) for childhood leukemia in a Greek study (16) and 0.67 (95% CI, 0.40-1.12) for childhood ALL in a New Zealand study (17). In a U.S. study with a large sample size and detailed disease characterization, neither attendance at daycare, age first started daycare, nor time enrolled at daycare was associated with risk of childhood ALL or c-ALL (10). In another U.S. study, longer duration of daycare attendance

Table 4. Daycare attendance measured by child-hours and the risk of childhood ALL, the NCCLS, 1995 to 2002

Child-hours in thousand	Total ALL		c-ALL		Non-c-ALL	
	Case/control	OR (95% CI)*	Case/control	OR (95% CI)*	Case/control	OR (95% CI)*
Censored on the reference date†						
Hispanic						
0	57/67	1.00	39/37	1.00	18/30	1.00
>0 and <15	33/50	0.98 (0.53-1.84)	19/30	0.85 (0.35-2.08)	14/20	1.25 (0.50-3.10)
≥15	30/36	1.27 (0.62-2.56)	13/19	1.02 (0.37-2.85)	17/17	2.05 (0.68-6.12)
P for trend‡		0.73		0.38		0.54
Non-Hispanic White						
0	37/35	1.00	19/22	1.00	18/13	1.00
>0 and <15	50/56	0.93 (0.44-1.95)	33/35	0.97 (0.33-2.85)	17/21	0.68 (0.22-2.13)
≥15	49/81	0.60 (0.28-1.27)	22/40	0.28 (0.08-0.95)	27/41	0.80 (0.28-2.26)
P for trend‡		0.03		0.01		0.41
Censored 1 y before the reference date†						
Hispanic						
0	66/87	1.00	46/53	1.00	20/34	1.00
>0 and <15	28/41	1.22 (0.64-2.32)	14/23	1.14 (0.45-2.87)	14/18	1.37 (0.54-3.48)
≥15	26/25	1.81 (0.78-4.18)	11/10	1.87 (0.54-6.39)	15/15	2.27 (0.66-7.74)
P for trend‡		0.68		0.37		0.56
Non-Hispanic White						
0	54/56	1.00	34/41	1.00	20/15	1.00
>0 and <15	36/50	0.93 (0.47-1.86)	21/29	1.08 (0.44-2.66)	15/21	0.69 (0.21-2.27)
≥15	46/66	0.75 (0.36-1.55)	19/27	0.34 (0.10-1.16)	27/39	0.88 (0.30-2.58)
P for trend‡		0.08		0.01		0.46
Censored at age 1 y†						
Hispanic						
0	98/132	1.00	58/74	1.00	40/58	1.00
>0 and <5	12/14	1.78 (0.66-4.83)	6/8	1.90 (0.43-8.45)	6/6	2.22 (0.54-9.24)
≥5	10/7	2.10 (0.70-6.34)	7/4	2.53 (0.60-10.7)	3/3	1.46 (0.26-8.03)
P for trend‡		0.23		0.21		0.81
Non-Hispanic White						
0	102/116	1.00	55/65	1.00	47/51	1.00
>0 and <5	20/26	1.27 (0.60-2.69)	9/11	0.95 (0.27-3.39)	11/15	1.52 (0.58-4.01)
≥5	14/30	0.42 (0.18-0.99)	10/21	0.33 (0.11-1.01)	4/9	0.50 (0.09-2.65)
P for trend‡		0.04		0.03		0.91

*ORs and 95% CIs are derived from conditional logistic regression models, adjusting for annual household income and maternal education.
† "Censored" at a specific point in time indicates that exposure occurred after that point is not considered. The "reference date" is the date of diagnosis for cases and the corresponding date for matched controls.
‡ *P* for trend is obtained by putting child-hours in the conditional logistic regression model as a continuous variable.

Table 5. Infections during infancy and the risk of childhood ALL by ethnicity, the NCCLS, 1995 to 2002

	Hispanic			Non-Hispanic White		
	Cases *n* (%)	Controls *n* (%)	OR (95% CI)*	Cases *n* (%)	Controls *n* (%)	OR (95% CI)*
Severe diarrhea/vomiting						
ALL						
No	78 (65.6)	112 (73.7)	1.00	106 (80.3)	141 (81.5)	1.00
Yes	41 (34.4)	40 (26.3)	1.42 (0.73-2.77)	26 (19.7)	32 (18.5)	1.31 (0.69-2.48)
c-ALL						
No	46 (64.8)	67 (77.9)	1.00	59 (79.7)	79 (82.3)	1.00
Yes	25 (35.2)	19 (22.1)	1.92 (0.72-5.08)	15 (20.3)	17 (17.7)	1.21 (0.50-2.91)
Ear infection						
ALL						
No	59 (49.2)	81 (53.3)	1.00	64 (47.8)	80 (46.0)	1.00
Yes	61 (50.8)	71 (46.7)	1.01 (0.61-1.69)	70 (52.2)	94 (54.0)	0.65 (0.39-1.10)
c-ALL						
No	36 (50.7)	42 (48.8)	1.00	37 (50.0)	41 (42.3)	1.00
Yes	35 (49.3)	45 (51.2)	0.74 (0.37-1.50)	37 (50.0)	56 (57.7)	0.32 (0.14-0.74)
Persistent cough						
ALL						
No	107 (88.4)	135 (87.7)	1.00	124 (92.5)	167 (96.0)	1.00
Yes	14 (11.6)	19 (12.3)	0.71 (0.31-1.60)	10 (7.5)	7 (4.0)	1.80 (0.50-6.55)
c-ALL						
No	63 (88.7)	75 (87.2)	1.00	67 (91.8)	93 (95.9)	1.00
Yes	8 (11.3)	15 (12.8)	0.60 (0.19-1.93)	6 (8.2)	4 (4.1)	2.24 (0.50-10.0)
Total no. infections during infancy†						
ALL						
0	23 (19.0)	39 (25.3)	1.00	40 (29.0)	51 (29.3)	1.00
1-4	48 (39.7)	73 (47.4)	1.00 (0.50-2.01)	61 (44.2)	76 (43.7)	0.92 (0.52-1.63)
≥5	50 (41.3)	42 (27.3)	1.74 (0.80-3.76)	37 (26.8)	47 (27.0)	0.79 (0.40-1.57)
c-ALL						
0	13 (18.3)	22 (25.6)	1.00	20 (27.0)	23 (23.7)	1.00
1-4	27 (38.0)	38 (44.2)	0.75 (0.27-2.08)	34 (46.0)	45 (46.4)	0.55 (0.23-1.31)
≥5	31 (43.7)	26 (30.2)	1.40 (0.48-4.06)	20 (27.0)	29 (29.9)	0.41 (0.15-1.14)

*ORs and 95% CIs are derived from conditional logistic regression models, adjusting for annual household income and maternal education.
† Those included severe diarrhea/vomiting, ear infection, persistent cough, mouth infection, eye infection, influenza, and unspecified ''other infections.'' Each episode of infection is counted once. For example, if a child had ear infections thrice, mouth infection twice, and no other infections, then the total number of infections would be 5.

seemed to confer a reduced risk of childhood ALL, especially B-lineage ALL, but the association was not consistently statistically significant (18). More recently, daycare attendance was not associated with the risk of childhood acute leukemia in a small case-control study conducted in Hong Kong, China (19).

The potential etiologic role of infections has been investigated in several studies. In a study of ALL in the Netherlands, children who had infections during infancy that required hospitalization had a OR of 0.6 (95% CI, 0.4-1.0; ref. 20). A German study evaluated 18 different types of infectious diseases before diagnosis. The OR for children ages ≥2 years who had no infectious disease during the first year of life but at least one infectious disease during the year preceding diagnosis, compared with those who had at least one infectious illness during the first year of life or no infectious illness during the year before diagnosis or both, was 1.2 (95% CI, 1.0-1.5) for c-ALL (21). It was observed in a Scottish study that the presence of a neonatal infection significantly reduced the risk of ALL, particularly in children diagnosed before age 4 years (22). Ear infection during infancy was less common among ALL cases than among controls in a U.S. study, with ORs of 0.86, 0.83, 0.71, and 0.69 for 1, 2 to 4, and 5+ episodes and continuous infections, respectively (P for trend = 0.026). In addition, the significantly inverse association with ear infection during infancy was more evident for c-ALL (10). As observed in a study conducted in Hong Kong, China, roseola and/or fever and rash during infancy were associated with a reduced risk of childhood acute leukemia (19). In a French study, having more than three infections per year before age 2 years was associated with a significantly reduced risk of childhood acute leukemia, as was surgical operation for ear-nose-throat infections before age 2 years (14). In another French study, having more than three infections during infancy was linked to a reduced risk of ALL (15). More recently, any infection in the first 2 years of life was linked to a reduced risk of ALL in children with Down syndrome (23).

Petridou et al. conducted a study with 94 ALL cases and 94 controls from Greece (24). In this study, past exposure to common infections was assessed using 10 serologic markers. There was little evidence for an association of ALL with the serology of any of the studied infectious agents among the very young children. In contrast, among children ages ≥5 years, leukemia was inversely associated with seropositivity to EBV, human herpes virus-6, *Mycoplasma pneumoniae*, and parvovirus B19. The findings indicate that among children ages ≥5 years the risk of ALL may be higher when the low herd immunity for several agents is challenged by late infection from an agent that, as a rule, would attack children at a younger age (24).

To our knowledge, the present study is the first to report an ethnic difference in the relationship among daycare attendance, infections, and risk of childhood ALL. Compared with non-Hispanic White children included in the study, Hispanic children in the study had significantly more other children living in the same household before they went to first grade. In addition, few Hispanic children started attending daycare before age 1 year. Thus, it is possible that contact with other children in a daycare setting was not the primary source of exposure to infectious agents for the Hispanic children included in this study. If the ethnic difference observed in the NCCLS is confirmed by other studies, it would be interesting to further explore underlying mechanisms. In our previously published analysis of daycare attendance (8), no ethnic difference was observed, which was probably due to the small number of Hispanic subjects included (30 cases and 30 controls).

A major strength of the present study is the detailed and timely assessment of daycare exposures. Comprehensive data on the time a subject started and stopped attending each

daycare facility, number of hours on average spent at the facility, and number of other children at the facility were obtained through personal interviews within several months after the reference date. In contrast to many previous studies that focused primarily on history of daycare attendance (ever versus never) and age beginning daycare attendance, we analyzed several different variables measuring daycare attendance and constructed child-hours, a variable that effectively summarized different aspects of daycare attendance. As for infections, respondents were asked specifically whether the index child had had any severe diarrhea/vomiting, ear infection, persistent cough, mouth infection, eye infection, and influenza during infancy, and if yes, when the infection occurred, how long it lasted, whether it required contacting a doctor and obtaining a prescription. These types of close-ended questions for specific infections generate higher-quality data than open-ended nonspecific questions (25).

Data included in the present analysis were restricted to information provided by biological mothers who were the primary caretakers of the index children. The availability of data on the timing of daycare attendance enabled censoring on the reference date, which ensures that cases and controls had a comparable time window for exposure. Censoring daycare attendance 1 year before the reference date eliminated the possibility that leukemia-related symptoms may have affected daycare attendance right before diagnosis. Furthermore, censoring daycare attendance at age 1 year offered a more detailed classification of exposure and a unique opportunity to study the timing of exposure.

Improved disease characterization is also strength of the present study. Based on biological plausibility, c-ALL is likely more appropriate than ALL for testing the infectious disease hypothesis (1, 2). For variables that seemed to be associated with the risk of ALL, the magnitude of effect with c-ALL was consistently stronger.

In the NCCLS, the matched population-based controls were selected from statewide birth records. An earlier evaluation indicated that controls enrolled in the NCCLS were comparable to ideal controls who would have been enrolled under optimal circumstances with respect to parental ages, maternal history of fetal loss, birth weight, birth order, and time since last live birth (9). However, cases and controls included in the present analysis were significantly different in annual household income and maternal education, two variables that reflect socioeconomic status and may also affect daycare attendance and early infections. Although these two variables were adjusted for in the analyses, possible residual confounding remained a concern. In a subgroup analysis of matched cases and controls who had the same annual household income, which were about a third of the overall study population, the pattern of association with daycare attendance and infections persisted. This result provided some assurance that the findings were not due to socioeconomic differences between cases and controls (i.e., due to selection bias). In addition, the striking differences between Hispanic and non-Hispanic White children and the differences observed between c-ALL and non-c-ALL provided evidence against selection bias as well as recall bias in the reporting of exposures by case mothers and control mothers.

A limitation of the study is the moderate number of cases and controls included, especially after stratifying by ethnicity and ALL subtype. Stratifying observations inevitably reduces sample size (statistical power). Nevertheless, more homogeneous subgroups reduce potential misclassification of the disease and therefore likely improve the quality of the study. The use of a matched design and a matched analysis also serves to increase the precision of the risk estimates. Another limitation of the study is that information on daycare attendance and infections during infancy was all self-reported.

To summarize, findings from the NCCLS data indicate a strong ethnic difference in the relationship among daycare attendance, early infections, and risk of childhood ALL. Daycare attendance and ear infection during infancy were associated with a significantly decreased risk of ALL among non-Hispanic White children in northern and central California but not among Hispanic children in the same region. In the future, studies with larger sample size, diverse ethnic groups, detailed exposure assessment, and more direct measures of infections and immune response are needed to elucidate the immunologic mechanisms in the etiology of childhood ALL.

Acknowledgments

We thank the clinical collaborators at the participating hospitals (Children's Hospital of Central California, Children's Hospital Oakland, Stanford Packard Children's Hospital, University of California-San Francisco Children's Hospital, University of California-Davis Children's Hospital, and Kaiser Permanente Hospitals in Oakland, San Francisco, Sacramento, and Santa Clara) for help with recruiting patients.

References

1. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. Leukemia 1988;2:120–5.
2. Greaves M. Aetiology of acute leukaemia. Lancet 1997;349:344–9.
3. Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain [see comments]. Lancet 1988;2:1323–7.
4. Alexander FE, McKinney PM, Cartwright RA. Migration patterns of children with leukaemia and non-Hodgkin's lymphoma in three areas of northern England. J Public Health Med 1993;15:9–15.
5. Istre GR, Conner JS, Broome CV, et al. Risk factors for primary invasive *Haemophilus influenzae* disease: increased risk from day care attendance and school-aged household members. J Pediatr 1985;106:190–5.
6. Ferson MJ. Infections in day care. Curr Opin Pediatr 1993;5:35–40.
7. Nystad W, Skrondal A, Magnus P. Day care attendance, recurrent respiratory tract infections and asthma. Int J Epidemiol 1999;28:882–7.
8. Ma X, Buffler PA, Selvin S, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. Br J Cancer 2002;86:1419–24.
9. Ma X, Buffler PA, Layefsky M, et al. Control selection strategies in case-control studies of childhood diseases. Am J Epidemiol 2004;159:915–21.
10. Neglia JP, Linet MS, Shu XO, et al. Patterns of infection and day care utilization and risk of childhood acute lymphoblastic leukaemia. Br J Cancer 2000;82:234–40.
11. Hoaglin DC, Mosteller F, Tukey JW, editors. Exploring data, table, trends and shapes. New York: John Wiley and Sons; 1985.
12. Petridou E, Kassimos D, Kalmanti M, et al. Age of exposure to infections and risk of childhood leukaemia. BMJ 1993;307:774.
13. Infante-Rivard C, Fortier I, Olson E. Markers of infection, breast-feeding and childhood acute lymphoblastic leukaemia. Br J Cancer 2000;83:1559–64.
14. Perrillat F, Clavel J, Auclerc MF, et al. Day-care, early common infections and childhood acute leukaemia: a multicentre French case-control study. Br J Cancer 2002;86:1064–9.
15. Jourdan-Da Silva N, Perel Y, Mechinaud F, et al. Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. Br J Cancer 2004;90:139–45.
16. Petridou E, Trichopoulos D, Kalapothaki V, et al. The risk profile of childhood leukaemia in Greece: a nationwide case-control study. Br J Cancer 1997;76:1241–7.
17. Dockerty JD, Skegg DC, Elwood JM, et al. Infections, vaccinations, and the risk of childhood leukaemia. Br J Cancer 1999;80:1483–9.
18. Rosenbaum PF, Buck GM, Brecher ML. Early child-care and preschool experiences and the risk of childhood acute lymphoblastic leukemia. Am J Epidemiol 2000;152:1136–44.
19. Chan LC, Lam TH, Li CK, et al. Is the timing of exposure to infection a major determinant of acute lymphoblastic leukaemia in Hong Kong? Paediatr Perinat Epidemiol 2002;16:154–65.
20. van Steensel-Moll HA, Valkenburg HA, van Zanen GE. Childhood leukemia and infectious diseases in the first year of life: a register-based case-control study. Am J Epidemiol 1986;124:590–4.
21. Schüz J, Kaletsch U, Meinert R, et al. Association of childhood leukaemia with factors related to the immune system. Br J Cancer 1999;80:585–90.
22. McKinney P, Juszczak E, Findlay E, et al. Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study. Br J Cancer 1999;80:1844–51.
23. Canfield KN, Spector LG, Robison LL, et al. Childhood and maternal infections and risk of acute leukaemia in children with Down syndrome: a report from the Children's Oncology Group. Br J Cancer 2004;91:1866–72.
24. Petridou E, Dalamaga M, Mentis A, et al. Evidence on the infectious etiology of childhood leukemia: the role of low herd immunity (Greece). Cancer Causes Control 2001;12:645–52.
25. Menegaux F, Olshan AF, Neglia JP, et al. Day care, childhood infections, and risk of neuroblastoma. Am J Epidemiol 2004;159:843–51.

Ethnic Difference in Daycare Attendance, Early Infections, and Risk of Childhood Acute Lymphoblastic Leukemia

Xiaomei Ma, Patricia A. Buffler, Joseph L. Wiemels, et al.

Cancer Epidemiol Biomarkers Prev 2005;14:1928-1934.