A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia

Kevin Y Urayama,[1]* Patricia A Buffler,[1] Emily R Gallagher,[2] Julie M Ayoob[3] and Xiaomei Ma[4]

[1]School of Public Health, University of California, Berkeley, CA, USA, [2]Seattle Children's Hospital, University of Washington, Seattle, WA, USA, [3]School of Nursing, University of California, San Francisco, CA, USA and [4]Yale University School of Medicine, New Haven, CT, USA

*Corresponding author: School of Public Health, University of California, Berkeley, 50 University Hall, MC-7360, Berkeley, CA 94720, USA. E-mail: urayamak@fellows.iarc.fr

Accepted 25 November 2009

Background Childhood acute lymphoblastic leukaemia (ALL) may be the result of a rare response to common infection(s) acquired by personal contact with infected individuals. A meta-analysis was conducted to examine the relationship between day-care attendance and risk of childhood ALL, specifically to address whether early-life exposure to infection is protective against ALL.

Methods Searches of the PubMed database and bibliographies of publications on childhood leukaemia and infections were conducted. Observational studies of any size or location and published in English resulted in the inclusion of 14 case–control studies.

Results The combined odds ratio (OR) based on the random effects model indicated that day-care attendance is associated with a reduced risk of ALL [OR = 0.76, 95% confidence interval (CI): 0.67, 0.87]. In subgroup analyses evaluating the influence of timing of exposure, a similarly reduced effect was observed for both day-care attendance occurring early in life ($\leqslant$2 years of age) (OR = 0.79, 95% CI: 0.65, 0.95) and day-care attendance with unspecified timing (anytime prior to diagnosis) (OR = 0.81, 95% CI: 0.70, 0.94). Similar findings were observed with seven studies in which common ALL were analysed separately. The reduced risk estimates persisted in sensitivity analyses that examined the sources of study heterogeneity.

Conclusions This analysis provides strong support for an association between exposure to common infections in early childhood and a reduced risk of ALL. Implications of a 'hygiene'-related aetiology suggest that some form of prophylactic intervention in infancy may be possible.

Keywords Childhood, leukaemia, day care, epidemiology, infection, meta-analysis, case–control studies

Introduction

Evidence is growing in support of a role for infections in the aetiology of childhood leukaemia, particularly for the most common subtype, acute lymphoblastic leukaemia (ALL).[1–3] Two infection-related hypotheses have gained popularity and are currently supported by substantial, yet inconsistent, epidemiologic findings. Kinlen first proposed the 'population mixing' hypothesis in response to the observed childhood leukaemia clusters occurring in the early 1980s in Seascale and Thurso, two remote and isolated communities in the UK that experienced a rapid influx of professional workers.[4] He proposed that childhood leukaemia may result from an abnormal immune response to specific, although unidentified, infections commonly seen with the influx of infected persons into an area previously populated with non-immune and susceptible individuals. This hypothesis suggests a mechanism that involves a direct pathological role of specific infectious agents, presumably viruses, in the development of childhood leukaemia and that an immunizing effect may be acquired through previous exposure. Supportive data include several subsequent studies conducted by Kinlen and others examining similar examples of population mixing including rural new towns, situations of wartime population change and other circumstances contributing to unusual patterns of personal contact.[4–11] Currently, there is no molecular evidence implicating cell transformation by a specific virus.[12]

The 'delayed infection' hypothesis proposed by Greaves emphasizes the critical nature of the timing of exposure and is intended to apply mostly to common B-cell precursor ALL (c-ALL), which largely accounts for the observed peak incidence of ALL between 2 and 5 years of age in developed countries.[13,14] He described a role for infections in the context of a 'two-hit' model of the natural history of c-ALL,[15] where the first 'hit' or initiating genetic event occurs *in utero* during fetal haematopoiesis producing a clinically covert pre-leukemic clone. The transition to overt disease occurs, in a small fraction (~1%) of pre-leukaemia carriers, after a sufficient postnatal secondary genetic event, which may be caused by a proliferative stress-induced effect of common infections on the developing immune system of the child.[1,13] This adverse immune response to infections is thought to be the result of insufficient priming of the immune system usually influenced by a delay in exposure to common infectious agents during early childhood. With the assumption that improved socio-economic conditions may lead to delay in exposure to infections, the Greaves hypothesis provides one plausible explanation for the notably higher incidence rates of ALL with its characteristic peak age between 2 and 5 years observed only in more socio-economically developed countries.[16,17] Although different in hypothesized mechanism, both the 'population mixing' and 'delayed infection' hypotheses propose childhood leukaemia to be caused by an abnormal immune response to infection(s) acquired by personal contacts, and are compatible with available evidence. In some populations, it is possible that both mechanisms may be operating.

Several previous epidemiological studies have used day-care attendance as an indicator of the increased likelihood of early exposure to infections,[18] since it is well documented that in developed countries exposures to common infections, particularly those affecting the respiratory and gastrointestinal tracts, occur more frequently in this type of setting.[19] The immaturity of children's immune systems in combination with the lack of appropriate hygienic behaviour is believed to promote the transmission of infectious agents in this social setting.[19–21] In the current analysis, we took a meta-analytic approach to summarize the findings to date on the relationship between day-care attendance and risk of childhood ALL.

Methods

Identification of studies

Literature searches were conducted in PubMed to identify original research and review articles related to childhood leukaemia and day-care attendance and/or social contacts published between January 1966 and October 2008. The searches were conducted using the term 'childhood leukaemia' in combination with other terms including 'infection', 'child care', 'day care' and 'social contact'. In addition, the bibliographies of epidemiology publications on childhood leukaemia and infections were searched to identify studies that may not have been captured through the initial database search. This included the review published in 2004 by McNally and Eden on the infectious aetiology of childhood acute leukaemia (AL).[2]

Inclusion/exclusion criteria and definitions

Among the studies identified, inclusion in the meta-analysis was limited to observational studies of case–control or cohort design of any size, geographic location and race/ethnicity of study participants. When more than one publication from an individual study was available, either the most recent publication or the publication that performed the analysis most applicable to evaluating the 'delayed infection' hypothesis was selected. Studies needed to have reported a relative risk (RR) or odds ratio (OR) and confidence intervals (CIs), or original data by disease status from which a measure of effect could be calculated. The outcome of interest was defined as clinically diagnosed leukaemia in children between the ages of 0 and 19 years. In the very few studies that did not distinguish between specific leukaemia subtypes,[22–26] it was assumed that ALL was the primary subtype since it accounts for the majority (~80%) of leukaemia diagnoses in children.[27]

The exposure of interest generally referred to as 'day-care attendance', which, in addition to formal day care, may have included preschool, nursery school, play groups, mother–toddler groups and other early social contacts. A strict criterion for the meaning of 'regular attendance' was not defined *a priori* since it was assumed that this would vary between studies. Of the primary studies identified, four were excluded for various reasons, including study emphasis on evaluating leukaemia prognosis and outcome,[28] an earlier analysis of data from a study for which a more complete and recent publication is available,[29] and not reporting a risk estimate for day-care attendance.[30,31] After the exclusions, a total of 14 studies, all case–control in design, were retained for the meta-analysis.[22–24,32–42]

Data extraction and statistical approach

For most studies,[23,24,34–40] the ORs and 95% CIs for leukaemia, AL, ALL or c-ALL among those who attended day care compared with those who did not attend day care were extracted. Among the few studies that did not provide this estimate, the OR for a similar measure was extracted, including those for no deficit in social contacts,[42] regular contact outside the home,[22] >36 months duration of day-care attendance,[41] increasing index and family day-care measure,[32] and social activity.[33] In two instances, the reported OR was recalculated to reflect the risk associated with the highest level of day-care attendance and/or social activity measure compared with the lowest.[41,42] Furthermore, several studies reported risk estimates for stratified analyses by specific subtype of leukaemia,[22,32,33,35–38,40–42] age at diagnosis,[34,35,38] specific age of day-care attendance,[22,33,32–38,40] or race/ethnicity;[37] multiple estimates were extracted from these studies for the purposes of subgroup and sensitivity evaluations in the meta-analysis, including specific leukaemia subtypes, particularly ALL and c-ALL and timing of day-care attendance. In general, studies referred to the common precursor B-cell ALL subtype (CD10 and CD19 positive ALL) as c-ALL. Four studies defined c-ALL with an added criterion that specified an age range between 2 and 5 years.[33,37,38,43] Risk estimates by specific diagnosis age groups were not extracted since there were only a few studies that provided this information and the age cut-points varied. For the one study that stratified by race/ethnicity,[37] two separate risk estimates were included in the meta-analysis since the reported estimates were based on independent populations.

The between-study heterogeneity was assessed using the Q statistic, which tests the null hypothesis that the estimated effect is homogenous across all studies.[44] Acknowledging that the eligible studies have been conducted independently and may represent only a random sample of the distribution of all possible effect sizes for this association, the random effects model was utilized, which incorporates an estimate of both between-study and within-study variation into the calculation of the summary effect measure.[45] Compared with the fixed effects model,[46] this method is more conservative and generally results in a wider CI. Finally, publication bias was evaluated visually using the funnel graph method that displays the distribution of all included studies by their point estimates and standard errors.[47] In addition, the Begg and Mazumdar adjusted rank correlation test was used to test for correlation between the effect estimates and their variances which, if present, provides an indication of publication bias.[48]

The association with c-ALL was evaluated with a meta-analysis of 7 of the 14 studies.[32,33,35–38,42] If a study reported multiple ORs and 95% CIs by timing of day-care attendance, the risk estimate associated with the earliest timing (e.g. age $\leqslant$ 2 years) was used to be consistent with the 'delayed infection' hypothesis.[23,32,34,37,38] The effect of the timing of exposure was evaluated in subgroup meta-analyses of studies reporting risk estimates for early day-care attendance (age $\leqslant$ 2 years)[22,23,32–34,36–38,42] and studies reporting risk estimates for day-care attendance anytime before diagnosis.[23,24,35,37–39,41] Finally, a series of sensitivity analyses were conducted to evaluate the sources of study heterogeneity, namely, the influences of potential selection bias, and heterogeneity in disease classification and exposure definition. The analyses were conducted using the statistical software, STATA Version 9.[49]

Results

Table 1 presents selected characteristics of the 14 studies included in this meta-analysis. The studies, all case–control in design, were published between 1993 and 2008 and were conducted in many different geographic areas. Most studies achieved a population-based ascertainment of cases utilizing a national registry or a regional network of all major paediatric oncology centres. A population-based control selection strategy was most common with the exception of three studies that selected hospital-based controls.[23,24,39] Only 1 of the 14 studies utilized a records-based day-care assessment protocol,[36] whereas the remaining studies relied on standardized questionnaires administered either in person, by telephone or by mail. All studies have accounted for major confounding factors such as age, sex, race and socio-economic status through a matched study design and/or statistical adjustment in the analysis. Of the 14 studies identified, 11 studies have reported either a statistically significant reduced risk associated with day-care attendance and/or social contact measures[23,33–37,39] or provide some evidence of a reduced risk.[22,24,40,41]

As shown in Table 2, the 14 studies included a total of 6108 cases and generated a combined OR estimate indicating that day-care attendance is associated with

Table 1 Select characteristics of studies included in the meta-analyses of day-care attendance and childhood leukemia

Author, year (location)	Case ascertainment	Control selection	Data collection	Select results	Confounding addressed
Petridou et al., 1993[23] (Greece, Attica and Crete)	• 125 leukaemia (Attica) 11 leukaemia (Crete) • Age 0–14 years • Children's hospitals of the University of Athens (1987–91) and the University of Crete (1990–92)	187 frequency matched children who attended the outpatient clinic of the hospitals where the children with leukaemia were treated	Telephone interviews with parents of children	Attendance at crèche (yes/no) • At any time Leukaemia: 0.67 (0.41, 1.11) • In infancy (for ⩾3 months in the first 2 years of life) Leukaemia: 0.28 (0.09, 0.88)	• Frequency match: sex, age, hospital • Adjustment: place of residence, social class
Roman et al., 1994[40] (UK)	• 38 ALL • Age 0–4 years • Diagnosed 1972–89 • Born in west Berkshire or north Hampshire, and were residents there when diagnosed	112 individually matched children selected from hospital delivery registers	Parents were interviewed	Preschool playgroup (yes/no) • In the year before diagnosis (for ⩾3 months) ALL: 0.6 (0.2, 1.8)	• Individual match: sex, date of birth, mother's age, area of residence at birth, time of diagnosis • Adjustment: not specified
Petridou et al., 1997[24] (Greece)	• 153 leukaemia • Age 0–14 years • Diagnosed 1993–94 • Nation wide network of paediatric haematologists/oncologists	300 individually matched children hospitalized at the same time as the corresponding case for acute conditions	Guardians of all subjects completed an interviewer-administered questionnaire	Day-care attendance (ever/never) Leukaemia: 0.83 (0.51, 1.37)	• Individual match: sex, age, geographic region • Adjustment: maternal age at birth, maternal education, sibship size, birth order, persons per room
Schuz et al., 1999[42] (Germany)	• 1010 AL (686 c-ALL) • Diagnosed 1980–94 • Age 0–14 years • Nation wide German Children's Cancer Registry at the University of Mainz	2588 matched children randomly selected from complete files of local offices of registration of residents	• Structured questionnaire based on the US Children's Cancer Group • Day-care attendance not directly assessed	Deficit in social contacts (yes/no, age ⩽18 months excluded) AL: 1.1 (0.9, 1.3) c-ALL: 1.0 (0.8, 1.2)	Analysis of AL • Individual match: date of birth, sex, district • Adjustment: SES Analysis of c-ALL • Adjustment: sex, age, year of birth, study setting • Adjustment: SES, urbanization

Table 1 Continued

Author, year (location)	Case ascertainment	Control selection	Data collection	Select results	Confounding addressed
Dockerty et al., 1999[22] (New Zealand)	• 97 ALL • Diagnosed 1990–93 • Age 0–14 years • New Zealand Cancer registry, public hospital admission/discharge computer system, and the Children's Cancer Registry. Nationwide	• 97 individually matched children randomly selected from the New Zealand born and resident childhood population using national birth records • 209 solid cancer cases	Mothers interviewed in their homes using a questionnaire adapted from Patricia McKinney and Eve Roman in the UK	Regular contact with other children from outside home at <12 months (yes/no, age <15 months excluded) ALL: 0.65 (0.36, 1.17)	• Individual match: age and sex • Adjustment: sex, age, several others including social class
Infante-Rivard et al., 2000[34] (Canada)	• 491 ALL • Diagnosed 1980–93 • Age 0–9 years • Tertiary care centre similar to population-based ascertainment	491 individually matched children chosen from the most complete census of children for the study years	Structured questionnaire administered to mothers by phone	Day-care attendance by age at entry • Entry ≤2 years old vs no ALL: 0.49 (0.31, 0.77) • Entry at >2 years old vs no ALL: 0.67 (0.45, 1.01)	• Individual match: age, sex, region of residence at diagnosis • Adjustment: maternal age, maternal education
Neglia et al., 2000[38] (USA)	• 1744 ALL (633 c-ALL; excludes cases <1 year) • Diagnosed 1989–93 • Age 0–14 years • Children's Cancer Group member institutions throughout the USA	1879 individually matched children randomly selected using the random digit dialing (RDD) methodology	Structured interview	Day-care attendance (age <1 year excluded) • Yes vs no ALL: 0.96 (0.82, 1.12) c-ALL: 0.96 (0.75, 1.24) • Day care before age 2 vs no ALL: 0.99 (0.84, 1.17) c-ALL: 1.05 (0.80, 1.37)	• Individual match: age, race, telephone area code, exchange, sex (T-cell leukaemia only) • Adjustment: maternal race, education, family income
Rosenbaum et al., 2000[41] (USA)	• 255 ALL • Diagnosed 1980–91 • Age 0–14 years • Four clinical centers in a 31-county study region. Institutional tumour registries and department of paediatric haematology-oncology records	760 frequency matched children randomly selected through the Live Birth Certificate Registry maintained by the New York State Department of Health	Self-administered questionnaire mailed to the parents of subjects	Total duration of out-of-home care (duration vs >36 months) ALL: Stay home: 1.32 (0.70, 2.52) 1–18 months: 1.74 (0.89, 3.42) 19–36 months: 1.32 (0.70, 2.52)	• Frequency match: sex, age, race, birth year • Adjustment: maternal age, maternal education, birth year, maternal employment, breastfeeding, birth order

Table 1 Continued

Author, year (location)	Case ascertainment	Control selection	Data collection	Select results	Confounding addressed
Chan et al., 2002[32] (Hong Kong)	• 98 AL (66 c-ALL) • Diagnosed 1994–97 • Age 2–14 years • Hong Kong Pediatric Hematology and Oncology Study Group	228 children selected using RDD methodology	In-person interview using a structured questionnaire adapted from UKCCS and translated into Chinese	Index and family day-care attendance (3-category variable) • First year of life AL: 0.96 (0.70, 1.32) Child peak: 0.63 (0.38, 1.07) c-ALL: 0.93 (0.63, 1.36)	• Matching: None • Adjustment: age, number of children in household at reference date
Perrillat et al., 2002[39] (France)	• 280 AL (240 ALL) • Diagnosed 1995–99 • Age 0–15 years • Hospitals of Lille, Lyon, Nancy and Paris. Cases need to have resided in the hospital catchment area	288 frequency matched children hospitalized in the same hospital as the cases, and residing in the catch-ment area of the hospital	In-person standardized questionnaire	Day-care attendance (age <2 years excluded) • Ever vs never AL: 0.6 (0.4, 1.0) • Age started vs no day care >12 months: 0.5 (0.3, 1.0) 7–12 months: 0.6 (0.2, 1.7) ≤6 months: 0.5 (0.3, 1.0)	• Frequency match: age, sex, hospital, hospital catchment area, ethnic origin • Adjustment: age, sex, hospital, ethnic origin, maternal education, parental professional category
Jourdan-Da Silva et al., 2004[35] (France)	• 473 AL (408 ALL, 304 c-ALL) • Diagnosed 1995–98 • Age 0–14 years • National Registry of Childhood Leukaemia and Lymphoma (NRCL)	567 frequency matched children randomly selected using age, sex, and region quotas from a sample of 30 000 phone numbers representa-tive of the French population on area of residence and munic-ipality size	Standardized self-administered questionnaire on mothers	Day-care attendance (age <1 year excluded) • Ever vs never ALL: 0.7 (0.6, 1.0) c-ALL: 0.8 (0.6, 1.0) • Started at age <3 months vs never ALL: 0.6 (0.4, 0.8) c-ALL: 0.6 (0.4, 0.9)	• Frequency match: age, sex, region • Adjustment: age, sex, region

(continued)

Table 1 Continued

Author, year (location)	Case ascertainment	Control selection	Data collection	Select results	Confounding addressed
Gilham et al., 2005[33] (UK)	• 1286 ALL (791 c-ALL; excludes cases <2 years) Diagnosed 1991–96 • Age 0–14 years • Nation-wide ascertainment through pediatric oncology units	6238 individually matched children randomly selected from primary care population registers	In-person interview with parents using a structured questionnaire	Social activity in the first year of life (age <2 years excluded) • Any vs no social activity ALL: 0.66 (0.56, 0.77) c-ALL: 0.67 (0.55, 0.82) • Age started vs no day care ALL- <3 months: 0.71 (0.60, 0.85) 3–5 months: 0.71 (0.56, 0.90) 6–11 months: 0.76 (0.63, 0.92)	• Individual match: sex, month and year of birth, region of residence at diagnosis • Adjustment: age at diagnosis, sex, region, maternal age, mother working at time of birth, deprivation
Ma et al., 2005[37] (USA)	• 294 ALL (145 c-ALL; excludes case <1 year) Diagnosed 1995– 2002 • Age 0–14 years • Population-based ascertainment from major paediatric clinical centres in Northern and Central California	376 individually matched children randomly selected from statewide birth certificate files maintained by the California Department of Health Services	Personal interview with primary caretaker	Child-hours of exposure at day care (age <1 year excluded) • ≥5000 child-hours (1st year) vs 0 ALL- Hispanic: 2.10 (0.70, 6.34) White: 0.42 (0.18, 0.99) c-ALL- Hispanic: 2.53 (0.60, 10.7) White: 0.33 (0.11, 1.01)	• Individual match: date of birth, sex, mother's race, Hispanic status • Adjustment: annual household income, maternal education
Kamper- Jorgensen et al., 2008[36] (Denmark)	• 559 ALL (199 c-ALL) Diagnosed 1989– 2004 • Age 0–15 years • All cases of childhood leukaemia identified in a cohort of all children in Denmark	5590 individually matched children selected from population registers	Records-based data from three population-based registries: The Nordic Society of Paediatric Haematology and Oncology, the Danish Civil Registration System, and the Childcare database	Child-care attendance in children during first 2 years of life (yes/no) ALL: 0.68 (0.48, 0.95) c-ALL: 0.58 (0.36, 0.93)	• Individual match: date of birth, sex, birth cohort • Adjustment: several demographic characteristics was considered but none were major confounders

SES, socioeconomic status; RDD, random digit dialing; UKCCS, United Kingdom Childhood Cancer Study.

Table 2 Meta-analysis of studies examining the association between day-care attendance and risk of childhood ALL

Study, year	Outcome, age in years	Day-care definition	Timing	Cases	OR	95% CI	Wi (%)[a]
Petridou et al., 1993[23]	Leukaemia, 0–14	Attendance at crèche: yes/no	Before age 2 years	136	0.28	0.09, 0.88	1.2
Roman et al., 1994[40]	ALL, 0–4	Pre school playgroup: yes/no	Year before dx	38	0.60	0.20, 1.80	1.3
Petridou et al., 1997[24]	Leukaemia, 0–14	Day care: ever/never	Birth to dx	153	0.83	0.51, 1.37	5.0
Schuz et al., 1999 [42,b]	AL, 1.5–14	Deficit in social contacts: yes/no	Before age 2 years	921	0.91	0.90, 1.30	12.7
Dockerty et al., 1999[22]	ALL, 1.25–14	Reg. contact outside home: yes/no	First year of life	90	0.65	0.36, 1.17	3.8
Infante-Rivard et al., 2000[34]	ALL, 0–9	Day care: entry at ≤2 years/never	At or before age 2 years	433	0.49	0.31, 0.77	5.6
Neglia et al., 2000[38]	ALL, 1–14	Day care before age 2 years: yes/no	Before age 2 years	1744	0.99	0.84, 1.17	13.3
Rosenbaum et al., 2000[41,b]	ALL, 0–14	Out-of-home care: >36 months/none	Birth to dx	158	0.76	0.70, 2.52	3.3
Chan et al., 2002[32]	AL, 2–14	Index and family day care: 3-level	First year of life	98	0.96	0.70, 1.32	8.5
Perrillat et al., 2002[39]	AL, 2–15	Day-care attendance: yes/no	Birth to dx	246	0.60	0.40, 1.00	5.5
Jourdan-Da Silva et al., 2004[35]	ALL, 1–14	Day-care attendance: yes/no	Birth to dx	387	0.70	0.60, 1.00	10.3
Gilham et al., 2005[33]	ALL, 2–14	Social activity: any/none	First year of life	1272	0.66	0.56, 0.77	13.6
Ma et al., 2005[37]—White	ALL, 1–14	Day care first year of life: yes/no	First year of life	136	0.77	0.43, 1.40	3.8
Ma et al., 2005[37]—Hispanic[c]	ALL, 1–14	Day-care attendance: yes/no	Birth to dx	120	1.09	0.62, 1.90	4.1
Kamper-Jorgensen et al., 2008[36]	ALL, 0–15	Attendance to child care: yes/no	Before age 2 years	176	0.68	0.48, 0.95	7.9
			Combined:	6108	0.76	0.67, 0.87	100.0
			P-value (heterogeneity) = 0.014				

[a]Percent weight assigned to each OR in the random effects model.
[b]Schuz et al.: changed reference to 'Yes-deficit in social contacts' by calculating the inverse of the OR provided for 'No-deficit in social contacts'; Rosenbaum et al.: estimated the OR for '>36 months' by calculating the inverse of the originally provided OR for 'stayed home'.
[c]Day-care attendance censored on reference date was used due to the low number of Hispanic subjects attending day care during the first year of life.
dx, diagnosis; Wi, weight.

a reduced risk of childhood ALL (OR = 0.76, 95% CI: 0.67, 0.87). Figure 1 provides a visual portrayal of the relationship between day-care attendance and the risk of childhood ALL. Three large studies conducted in Germany,[42] the USA[38] and the UK[33] appeared to carry a large proportion of the weight in the meta-analysis at ~13% each. The combined risk estimates excluding each of these studies individually remained similarly reduced indicating that no one large study was able to completely explain the protective effect observed (data not shown). No remarkable evidence of publication bias was apparent from the funnel plot since the data points for these 14 studies were, in general, randomly distributed around the combined OR estimate (plot not shown). This visual interpretation of the results was confirmed by the large P-value using the rank correlation method (P = 0.553).

We attempted to maintain a reasonable balance between maximizing the inclusion of studies and minimizing sources of heterogeneity, by relaxing the eligibility criteria to include estimates for broader leukaemia subtypes, other social contact measures and unspecified timing of exposure. The contribution of the influence of possible sources of heterogeneity on the combined risk estimate was evaluated. In subgroup meta-analyses presented in Table 3 examining the influence of the timing of exposure, the combined OR for seven studies reporting estimates for day-care attendance or social contacts before diagnosis showed a reduced risk of childhood ALL (OR = 0.81, 95% CI: 0.70, 0.94). When the meta-analysis was limited to the nine studies that specifically evaluated day-care attendance at or before age 1 or 2 years, a similarly reduced risk of ALL (OR = 0.79, 95% CI: 0.65, 0.95) was observed.

A series of sensitivity analyses were conducted on the meta-analysis of the 14 studies to examine the influence of individual study characteristics on the combined OR, namely, potential biases in the selection of controls, the categorization of leukaemia and the assessment of day-care attendance. Figure 2 presents a summary of these analyses showing that none of these factors was able to completely account for the reduced risk of ALL observed in the main analysis of the 14 studies. For example, in the evaluation of potential control selection bias, reduced risks were observed for the analyses excluding three studies that used hospital-based controls (OR = 0.78, 95% CI: 0.68, 0.90) and excluding two studies that used random digit dialing (RDD) to select controls (OR = 0.72, 95% CI: 0.63, 0.81). Similarly reduced combined ORs were observed when excluding studies that included infants (<1 year of age) in the study population (OR = 0.81, 95% CI: 0.70, 0.94), studies not specifically examining ALL (OR = 0.74, 95% CI: 0.63, 0.87), and studies that did not define the exposure strictly as attendance at a day care or a similar type of setting (OR = 0.74, 95% CI: 0.61, 0.88).
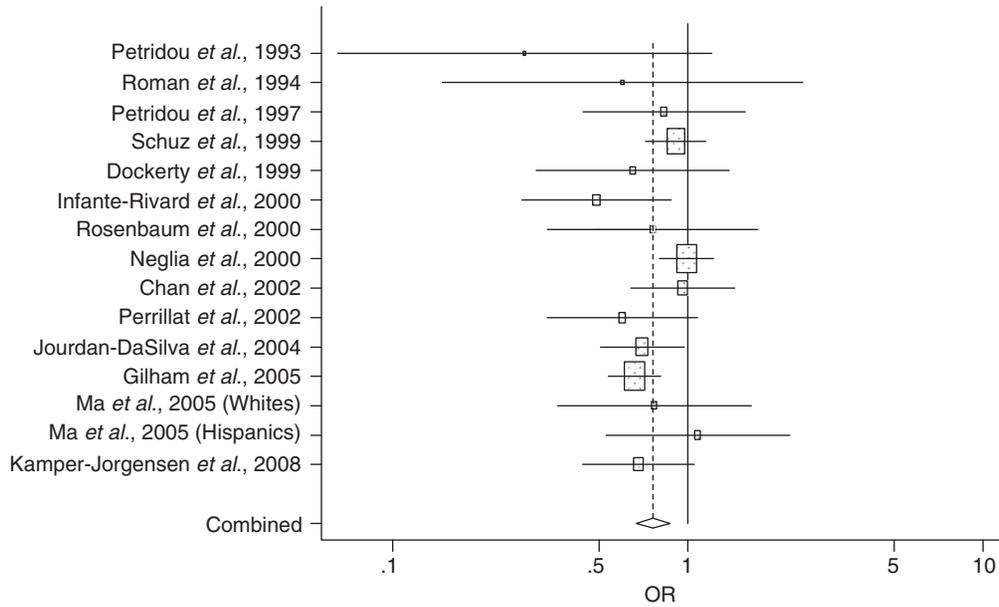
Figure 1 Forest plot displaying ORs and 95% CIs of studies examining the association between day-care attendance and risk of childhood ALL. The risk estimates are plotted with boxes and the area of each box is inversely proportional to the variance of the estimated effect. The horizontal lines represent the 95% CIs of the risk estimate for each study. The solid vertical line at 1.0 represents a risk estimate of no effect. The dashed vertical line represents the combined risk estimate (OR = 0.76), and the width of the diamond is the 95% CI for this risk estimate (0.67–0.87).

Table 3 Subgroup meta-analyses of day-care attendance and risk of childhood ALL evaluating the influence of timing of day-care attendance

Study, year	Day care any time				Day care at age ⩽ 2			
	Cases	OR	95% CI	Wi (%)[a]	Cases	OR	95% CI	Wi (%)[a]
Petridou et al., 1993[23]	136	0.67	0.41, 1.11	7.7	136	0.28	0.09, 0.88	2.3
Roman et al., 1994[40]								
Petridou et al., 1997[24]	153	0.83	0.51, 1.37	7.8				
Schuz et al., 1999[42,b]					921	0.91	0.90, 1.30	15.7
Dockerty et al., 1999[22]					90	0.65	0.36, 1.17	6.5
Infante-Rivard et al., 2000[34]					433	0.49	0.31, 0.77	8.8
Neglia et al., 2000[38]	1744	0.96	0.82, 1.12	38.0	1744	0.99	0.84, 1.17	16.1
Rosenbaum et al., 2000[41,b]	158	0.76	0.70, 2.52	4.9				
Chan et al., 2002[32]					98	0.96	0.70, 1.32	12.0
Perrillat et al., 2002[39]	246	0.60	0.40, 1.00	8.9				
Jourdan-Da Silva et al., 2004[35]	387	0.70	0.60, 1.00	22.1				
Gilham et al., 2005[33]					1272	0.66	0.56, 0.77	28.8
Ma et al., 2005—White[37]	136	0.75	0.38, 1.45	4.5	136	0.77	0.43, 1.40	2.1
—Hispanic[37]	120	1.09	0.62, 1.90	6.2	120	1.92	0.89, 4.13	1.2
Kamper-Jorgensen et al., 2008[36]					176	0.68	0.48, 0.95	6.3
Combined:	3080	0.81	0.70, 0.94	100.0	5126	0.79	0.65, 0.95	100.0
P-value (heterogeneity):	0.277				0.001			

[a]Percent weight assigned to each OR in the random effects model. Wi, weight.
[b]Schuz et al.: changed reference to 'Yes-deficit in social contacts' by calculating the inverse of the OR provided for 'No-deficit in social contacts'; Rosenbaum et al.: estimated the OR for '>36 months' by calculating the inverse of the originally provided OR for 'stayed home'.

Table 4 presents the results of the meta-analyses evaluating the association between childhood c-ALL and day-care attendance. The analysis of c-ALL contained fewer numbers of studies compared with the analysis of ALL. Similar to the result from the meta-analysis of ALL, the combined OR for the seven studies of c-ALL was also <1, although the CI was slightly wider (OR = 0.83, 95% CI: 0.70, 0.98). The subgroup analyses among studies of day-care attendance before age 1 or 2 years and c-ALL
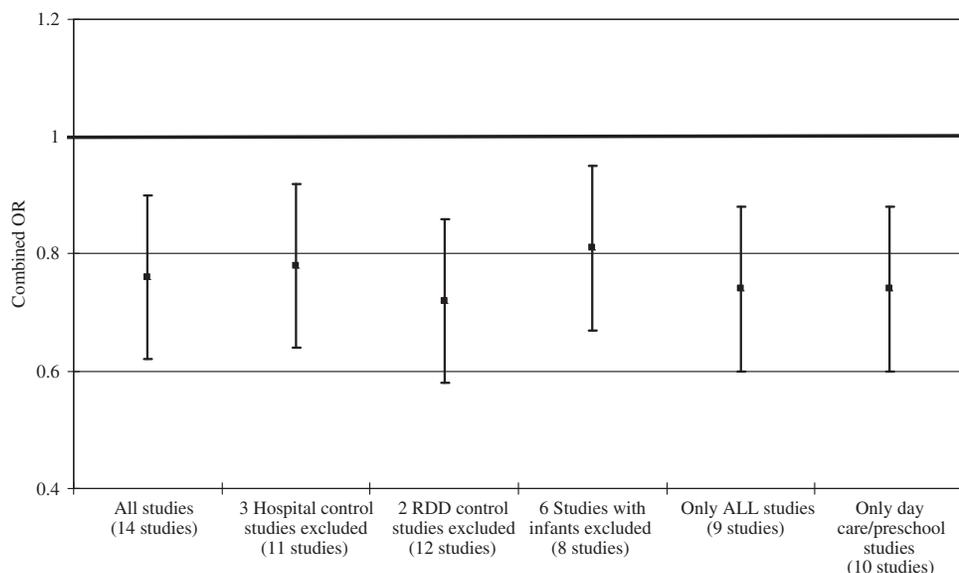
Figure 2 Plot showing results of sensitivity meta-analyses evaluating the influence of potential biases within individual studies on combined risk estimates. RDD, random digit dialing.

generated results similar to those for ALL (data not shown). No evidence of publication bias was observed for these analyses.

Discussion

The evidence from a large and growing body of literature related to the exposure to infectious agents, as measured by day-care attendance, and the risk of childhood leukaemia was systematically evaluated using a meta-analytic approach. Heterogeneity between epidemiologic studies and their results is common and constitutes one of the major challenges in such a synthesis. Although the random effects model was used in this analysis to account for some of the between-study variation, we acknowledge the importance of interpreting results together with a thorough consideration of the potential sources of heterogeneity.

All the studies included in this analysis were conducted with the *a priori* objective of testing the biologically plausible, 'delayed infection' hypothesis, which specifies a predicted direction of risk, timing of the exposure and the most applicable subtype of leukaemia. Overall, the studies show consistency in support of a reduced risk associated with day-care attendance or social contacts during early childhood, with the vast majority of studies either reporting an effect in the hypothesized direction or no association. A quantitative assessment using meta-analysis indicates that day-care attendance is associated with a reduced risk of childhood ALL, as well as c-ALL. The reduction in risk persisted despite a thorough consideration of potential sources of study heterogeneity. We did not conduct a meta-analysis specifically in non-c-ALL or acute myeloid leukaemia due to the limited number of studies reporting results for these associations. Of the four studies that present data for non-c-ALL,[35–38] three studies showed reduced ORs,[35–37] but lacked precision. Based on currently available data, it is difficult to determine whether the association applies to a specific subtype of ALL only or ALL in general.

The subgroup meta-analysis by timing of day-care attendance did not suggest a stronger reduction in risk for day care specifically at or before age 1 or 2 years as might have been expected based on the hypothesis. However, a few individual studies have shown that the strongest reduction in risk occurs when day-care attendance is started <6 months of age.[33,35,39] Although not formally evaluated in this meta-analysis, several individual studies that used detailed exposure assessment protocols demonstrated evidence of dose–response effects. Strong trends were observed for increasing levels of child-hours of day-care attendance,[37] levels of social activity[33] and age at start of day care.[35]

We were not able to conduct a comparable meta-analysis of studies pertaining to the related mechanism of rural 'population mixing' and the risk of childhood leukaemia. Although it was not possible to analyse the role of 'population mixing' in the same manner as was done for the 'delayed infection' hypothesis, it is recognized that these two processes may be interrelated or occur simultaneously and that both mechanisms may be operating in a given population. Thus, the results observed for the analyses of studies providing data relevant to the timing of infection in early life cannot be interpreted as 'ruling out' the possible role of 'population mixing', but rather lend further support to the role of immune related processes in the aetiology of ALL.

Table 4 Meta-analysis of studies examining the association between day-care attendance and risk of c-ALL

Study, year	Age	Day-care definition	Timing	Cases	OR	95% CI	Wi (%)[a]
Schuz et al., 1999[42]	1.5–14	Deficit in social contacts: Yes/no	Before age 2 years	658	1.00	0.80, 1.20	19.9
Neglia et al., 2000[38]	2–5	Day care before age 2 years: Yes/no	Before age 2 years	633	1.05	0.80, 1.37	16.3
Chan et al., 2002[32]	2–14	Index and family day care: 3-level	First year of life	66	0.93	0.63, 1.36	11.4
Jourdan-Da Silva, 2004[35]	1–14	Day-care attendance: Yes/no	Birth to dx	304	0.80	0.60, 1.00	17.0
Gilham et al., 2005[33]	2–5	Social activity: Any/none	First year of life	791	0.67	0.55, 0.82	20.1
Ma et al., 2005—White[37]	2–5	Day care first year of life: Yes/no	First year of life	74	0.49	0.19, 1.26	2.8
—Hispanic[37,b]	2–5	Day-care attendance: Yes/no	Birth to dx	71	0.91	0.41, 2.05	3.8
Kamper-Jorgensen et al., 2008[36]	0–15	Attendance to child care: Yes/no	Before age 2 years	101	0.58	0.36, 0.93	8.7
			Combined:	2698	0.83	0.70, 0.98	100.0
			P-value (heterogeneity) = 0.044				

aPercent weight assigned to each OR in the random effects model.
bDay-care attendance censored on reference date was used due to the low number of Hispanic subjects attending day-care during the first year of life.
dx, diagnosis; Wi, weight.

One major consideration in the evaluation of study validity is the possibility of selection bias, a type of systematic error that occurs when there is differential selection of either the cases or controls on the basis of characteristics which may affect exposure status. One way this may arise is if cases and controls do not originate from the same source population. A population-based ascertainment of cases is considered favourable since a defined source population, from which controls may be selected, is easily identifiable. Other strategies of case ascertainment may be appropriate as well, as long as the source population can be clearly defined. As implemented in three of the included studies, selection of controls among the inpatient cohort of the same hospital as the case diagnosis can fulfill this requirement, but can introduce bias if the illnesses/conditions of the control group are related to the exposure under study. Also, it has been suggested that the use of RDD, a population-based method of control recruitment, may result in a control group biased with respect to certain population characteristics that may be associated with exposures of interest.[50] Analysis excluding the three studies that selected hospital-based controls[23,24,39] or the two studies that used RDD to recruit population-based controls[32,38] produced similar results to those for the full set of studies.

Similar types of systematic biases resulting in socio-economic differences between cases and controls have been implicated in other studies as well, including the large United Kingdom Children Childhood Cancer Study (UKCCS)[33] and the Northern California Childhood Leukemia Study (NCCLS).[37,51] Adjustments for these differences have been implemented in the analyses; however, the possibility of residual effects cannot be ruled out. To alleviate some of this concern, results of a subgroup analysis conducted in the NCCLS among matched cases and controls who had the same annual household income showed that the pattern of association with day-care attendance persisted.[37]

The potential for information bias in case–control studies is of particular importance due to the retrospective nature of data collection, and the recall of past exposures may be influenced by disease status. Most studies collected exposure data based on respondent recall using a standardized questionnaire administered either in person, by telephone or by mail. Recall bias in the evaluation of c-ALL is expected to be less likely, since diagnoses of c-ALL are usually made between ages 2 and 5 years, and recall of early exposure histories may be easier for the primary caregiver. Although the influence of recall bias could not be formally evaluated in these meta-analyses, one records-based day-care study conducted by Kamper-Jorgensen et al. in Denmark reported a reduced risk of childhood ALL associated with childcare attendance during the first 2 years of life.[36] Several subtype specific analyses performed in this

study showed the strongest association in B-cell precursor ALL and c-ALL.

In addition to potential biases associated with the ability of respondents to accurately recall past events, there was variation between studies in the extent of exposure assessment and categorization of individual exposures to infectious agents. For example, Schuz et al. reported results from a matched case–control study conducted in Germany that used a 'deficit in social contacts' variable based on the assumption that children were likely to have attended day care if during the first 2 years of their life both parents were in full-time work.[42] The assumption made in the formulation of this social contact variable most likely contributed some non-differential misclassification, which tends to bias findings towards one of no effect. Their analysis did not indicate an association between deficit in social contact and AL or c-ALL.

In contrast, in the UKCCS, Gilham et al. created a hierarchical variable that reflected a child's overall social activity based on interview data incorporating information on frequency of regular activity with children outside the home, frequency of attendance at a day nursery or nursery school, and number of other children in attendance.[33] These analyses indicated that social activity/day-care attendance is associated with a reduced risk of childhood ALL. Ma et al., in the first publication on day-care attendance from the NCCLS, constructed a 'child-hours of exposure' variable incorporating information on the number of months attending a day care, mean hours per week at this day care and the number of children exposed to at this day care. They reported that children who had more total child-hours of exposure had a reduced risk of ALL.[29] These results were later confirmed in a follow-up analysis using a larger study population.[37] In non-Hispanic White children, children in the highest category of child-hours during infancy had a reduced risk of ALL and c-ALL compared with children who did not attend day care with strong evidence of a dose–response effect. This association was not observed in Hispanic children, which, as noted by the authors, had different socio-economic and demographic characteristics, including larger family size and different day-care utilization patterns. Although these types of refined exposure assessment strategies that account for duration, frequency and size of the day-care facility serve as examples for future studies, results from these analyses may have contributed to study heterogeneity. In a meta-analysis of 10 studies that strictly defined the exposure as attendance at a day care or other similar types of settings,[23,24,34–41] a reduced risk estimate was observed.

Current evidence suggests that different subtypes of leukaemia, defined by both immunophenotypic and molecular characteristics, may be associated with distinct aetiological mechanisms.[52,53] To minimize the bias associated with misclassification of the phenotype, most studies specifically evaluating the infectious hypothesis have reported results by subtype-specific leukaemia such as c-ALL, and have excluded infants since there is evidence suggesting these leukaemias may be associated with a causal mechanism involving transplacental chemical carcinogenesis.[54–56] This is not expected to be a major source of error, as observed in the sensitivity analysis, since infant leukaemias comprise only a very small proportion of all leukaemia diagnoses ($<5\%$).[57] It is believed that the hypothesis on infections, particularly the 'delayed infection' hypothesis, is most relevant to ALL and its most common subtype, c-ALL.[1] Limiting the meta-analysis to only those studies providing risk estimates for specific subtypes resulted in a reduced risk associated with both ALL[22,33–38,40,41] and c-ALL.[32,33,35–38,42]

The UKCCS recently published results from the first records-based study examining the relationship between clinically diagnosed infections in the first year of life and childhood ALL.[43,58] Contrary to what is expected based on the 'delayed infection' hypothesis and what was observed in this meta-analysis of day-care attendance, the results of this well-designed records-based study showed evidence of an increased risk of childhood ALL and c-ALL associated with clinically diagnosed infections in the first year of life. It is possible that these contrasting results reflect one of many mechanisms involved in the aetiology of childhood ALL. The authors explain that their findings may indicate that a dysregulated immune response to infections during the first few months of life leads to an increased risk of ALL.[43]

Alternatively, from a methodological perspective, it has been suggested that these contrasting results may be an indication that previous studies using self-reported data on infections and social contacts, many of which have found a reduced risk of ALL, may be biased due to differential recall/reporting between cases and controls.[58] Although more studies are needed to evaluate this apparent discrepancy, it is important to note at this juncture that infection based on clinical diagnosis may reflect a different infectious disease experience of the child compared with a self-reported infectious disease history, as mothers may not seek medical attention for all of the common infections experienced by the child.

Although still susceptible to recall bias, surrogate measures of exposure to infections such as day-care attendance and birth order, are recognized as strong alternative measures to testing the 'delayed infection' hypothesis, since they are highly associated with common childhood infectious diseases and have the added advantage of capturing a child's asymptomatic infections.[59] It is not known to what extent recall bias may have affected results of previous day-care studies, but there is evidence from a recent Denmark study also showing strong evidence of a reduced risk associated with a records-based assessment of day-care attendance.[36]

Overall, this meta-analysis of existing epidemiological data provides strong support for an association between exposure to common infections in early childhood and subsequent risk of ALL. As an indirect measure of exposure to infections, the ability of day-care attendance to serve as a surrogate measure may vary depending on characteristics of the facility attended and the child's pattern of attendance. Epidemiologic studies have shown that the transmission and development of infectious diseases are highly influenced by the age of the child, frequency and duration of attendance, structure and size of the facility.[19,21] Future epidemiologic studies of childhood leukaemia should attempt to obtain this type of detailed information on the facilities attended to refine the exposure classification.

Although inconsistent, there is evidence from studies of other surrogate measures of exposure to infections including birth order,[2] parental social contacts in the workplace,[60] and other immune-related factors (e.g. vaccination and breastfeeding history[61,62]), that support a role for infections and immune response in the aetiology of childhood leukaemia. The causal significance of the role of infections in childhood ALL would be strengthened by identification of a plausible biological mechanism for the conversion of pre-leukemic cells following infection[1] and by incorporation of genetic biomarkers of susceptibility and immune response into further epidemiological studies.[63,64] The protective effect of early infection on risk of subsequent childhood ALL parallels the similarly protective impact of parasitic infections on type I diabetes in both animal models and children.[65] An important implication of these 'hygiene'-related hypotheses and supportive data is that some form of prophylactic intervention in infancy may ultimately be possible.[1,65]

Funding

Acknowledgements

KEY MESSAGES

- There is growing evidence to suggest that childhood leukaemia, particularly ALL, may be the result of an abnormal immune response to common infections likely caused by a delay in exposure to infections early in life.

- Day-care attendance, one strong surrogate measure of exposure to infections early in life, is associated with a reduced risk of childhood ALL.

- Similarly reduced risk estimates were observed for c-ALL; however, further studies are needed to determine whether the association applies exclusively to this common subtype or to ALL or leukaemia in general.

References

1 Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer* 2006;**6**:193–203.

2 McNally RJ, Eden TO. An infectious aetiology for childhood acute leukaemia: a review of the evidence. *Br J Haematol* 2004;**127**:243–63.

3 O'Connor SM, Boneva RS. Infectious etiologies of childhood leukemia: plausibility and challenges to proof. *Environ Health Perspect* 2007;**115**:146–50.

4 Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. *Lancet* 1988;**2**:1323–27.

5 Alexander FE, Boyle P, Carli PM *et al*. Spatial clustering of childhood leukaemia: summary results from the EUROCLUS project. *Br J Cancer* 1998;**77**:818–24.

6 Kinlen L. Childhood leukaemia and ordnance factories in west Cumbria during the Second World War. *Br J Cancer* 2006;**95**:102–6.

7 Kinlen L, Doll R. Population mixing and childhood leukaemia: Fallon and other US clusters. *Br J Cancer* 2004;**91**:1–3.

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

9 Kinlen LJ. Infection and childhood leukemia. *Cancer Caus Cont* 1998;**9**:237–39.

10 Kinlen LJ, Balkwill A. Infective cause of childhood leukaemia and wartime population mixing in Orkney and Shetland, UK. *Lancet* 2001;**357**:858.

11 Kinlen LJ, Dickson M, Stiller CA. Childhood leukaemia and non-Hodgkin's lymphoma near large rural construction

sites, with a comparison with Sellafield nuclear site. *BMJ* 1995;**310:**763–68.

12 MacKenzie J, Greaves MF, Eden TO *et al*. The putative role of transforming viruses in childhood acute lymphoblastic leukemia. *Haematologica* 2006;**91:**240–43.

13 Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* 1988; **2:**120–25.

14 Greaves MF, Pegram SM, Chan LC. Collaborative group study of the epidemiology of acute lymphoblastic leukaemia subtypes: background and first report. *Leuk Res* 1985; **9:**715–33.

15 Greaves MF, Wiemels J. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer* 2003;**3:** 639–49.

16 Greaves MF, Alexander FE. An infectious etiology for common acute lymphoblastic leukemia in childhood? *Leukemia* 1993;**7:**349–60.

17 Hepworth SJ, Feltbower RG, McKinney PA. Childhood leukaemias and CNS tumours: correlation of international incidence rates. *Eur J Cancer* 2006;**42:**509–13.

18 Urayama K, Ma X, Buffler PA. Exposure to infections through day-care attendance and risk of childhood leukemia. *Rad Protec Dos* 2008;**132:**259–66.

19 Osterholm MT. Infectious disease in child day care: an overview. *Pediatrics* 1994;**94:**987–90.

20 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.

21 Holmes SJ, Morrow AL, Pickering LK. Child-care practices: effects of social change on the epidemiology of infectious diseases and antibiotic resistance. *Epidemiol Rev* 1996;**18:**10–28.

22 Dockerty JD, Skegg DC, Elwood JM, Herbison GP, Becroft DM, Lewis ME. Infections, vaccinations, and the risk of childhood leukaemia. *Br J Cancer* 1999;**80:** 1483–89.

23 Petridou E, Kassimos D, Kalmanti M *et al*. Age of exposure to infections and risk of childhood leukaemia. *BMJ* 1993;**307:**774.

24 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:**1241–47.

25 Schuz J, Kaatsch P, Kaletsch U, Meinert R, Michaelis J. Association of childhood cancer with factors related to pregnancy and birth. *Int J Epidemiol* 1999;**28:**631–39.

26 Shaw G, Lavey R, Jackson R, Austin D. Association of childhood leukemia with maternal age, birth order, and paternal occupation. A case–control study. *Am J Epidemiol* 1984;**119:**788–95.

27 Smith M, Ries L, Gurney J, Ross J. Leukemia. In: Ries L, Smith M, Gurney J *et al*. (eds). *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995*. Pub. No. 99-4649 ed. Bethesda, MD: National Cancer Institute, SEER Program, 1999, pp.17–34.

28 Charalampopoulou A, Petridou E, Spyridopoulos T *et al*. An integrated evaluation of socioeconomic and clinical factors in the survival from childhood acute lymphoblastic leukaemia: a study in Greece. *Eur J Cancer Prev* 2004; **13:**397–401.

29 Ma X, Buffler PA, Selvin S *et al*. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. *Br J Cancer* 2002;**86:**1419–24.

30 MacArthur AC, McBride ML, Spinelli JJ, Tamaro S, Gallagher RP, Theriault GP. Risk of childhood leukemia associated with vaccination, infection, and medication use in childhood: the Cross-Canada Childhood Leukemia Study. *Am J Epidemiol* 2008;**167:**598–606.

31 Rosenbaum PF, Buck GM, Brecher ML. Allergy and infectious disease histories and the risk of childhood acute lymphoblastic leukaemia. *Paediatr Perinat Epidemiol* 2005; **19:**152–64.

32 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.

33 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:**1294.

34 Infante-Rivard C, Fortier I, Olson E. Markers of infection, breast-feeding and childhood acute lymphoblastic leukaemia. *Br J Cancer* 2000;**83:**1559–64.

35 Jourdan-Da Silva N, Perel Y *et al*. Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. *Br J Cancer* 2004;**90:**139–45.

36 Kamper-Jorgensen M, Woodward A, Wohlfahrt J *et al*. Childcare in the first 2 years of life reduces the risk of childhood acute lymphoblastic leukemia. *Leukemia* 2008; **22:**189–93.

37 Ma X, Buffler PA, Wiemels JL *et al*. Ethnic difference in daycare attendance, early infections, and risk of childhood acute lymphoblastic leukemia. *Cancer Epidemiol Biomarkers Prev* 2005;**14:**1928–34.

38 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.

39 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:**1064–69.

40 Roman E, Watson A, Bull D, Baker K. Leukaemia risk and social contact in children aged 0–4 years in southern England. *J Epidemiol Community Health* 1994; **48:**601–602.

41 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.

42 Schuz J, Kaletsch U, Meinert R, Kaatsch P, Michaelis J. Association of childhood leukaemia with factors related to the immune system. *Br J Cancer* 1999;**80:**585–90.

43 Roman E, Simpson J, Ansell P *et al*. Childhood acute lymphoblastic leukaemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study. *Am J Epidemiol* 2007;**165:**496–504.

44 Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M Davey, Smith G, Altman DG (eds). *Systematic Rewiews in Health Care: Meta-Analysis in Context*. London: BMJ Publishing Group, 2001, pp. 285–312.

[45] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7:**177–88.

[46] Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987;**9:**1–30.

[47] Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001;**54:**1046–55.

[48] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;**50:**1088–101.

[49] Stata Statistical Software: Release 9 [program]. College Station, TX: StataCorp LP, 2005.

[50] Greenberg ER. Random digit dialing for control selection. A review and a caution on its use in studies of childhood cancer. *Am J Epidemiol* 1990;**131:**1–5.

[51] Ma X, Buffler PA, Layefsky M, Does MB, Reynolds P. Control selection strategies in case–control studies of childhood diseases. *Am J Epidemiol* 2004;**159:**915–21.

[52] Wiemels JL, Pagnamenta A, Taylor GM, Eden OB, Alexander FE, Greaves MF. A lack of a functional NAD(P)H:quinone oxidoreductase allele is selectively associated with pediatric leukemias that have MLL fusions. United Kingdom Childhood Cancer Study Investigators. *Cancer Res* 1999;**59:**4095–99.

[53] Wiemels JL, Smith RN, Taylor GM, Eden OB, Alexander FE, Greaves MF. Methylenetetrahydrofolate reductase (MTHFR) polymorphisms and risk of molecularly defined subtypes of childhood acute leukemia. *Proc Natl Acad Sci USA* 2001;**98:**4004–4009.

[54] Alexander FE, Patheal SL, Biondi A *et al*. Transplacental chemical exposure and risk of infant leukemia with MLL gene fusion. *Cancer Res* 2001;**61:**2542–46.

[55] Greaves MF. Aetiology of acute leukaemia. *Lancet* 1997;**349:**344–49.

[56] Ross JA, Davies SM, Potter JD, Robison LL. Epidemiology of childhood leukemia, with a focus on infants. *Epidemiol Rev* 1994;**16:**243–72.

[57] Ries LAG. Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975–1995. Bethesda, MD: National Cancer Institute, SEER Program, 1999.

[58] Simpson J, Smith A, Ansell P, Roman E. Childhood leukaemia and infectious exposure: A report from the United Kingdom Childhood Cancer Study (UKCCS). *Eur J Cancer* 2007;**43:**2396–403.

[59] Greaves M, Buffler PA. Infections in early life and risk of childhood ALL. *Br J Cancer* 2009;**100:**863.

[60] Chang JS, Metayer C, Fear NT *et al*. Parental social contact in the work place and the risk of childhood acute lymphoblastic leukaemia. *Br J Cancer* 2007;**97:**1315–21.

[61] Kwan ML, Buffler PA, Abrams B, Kiley VA. Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Rep* 2004;**119:**521–35.

[62] Ma X, Does MB, Metayer C, Russo C, Wong A, Buffler PA. Vaccination history and risk of childhood leukaemia. *Int J Epidemiol* 2005;**34:**1100–1109.

[63] Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. *Blood* 1999;**94:**694–700.

[64] Taylor GM, Dearden S, Ravetto P *et al*. Genetic susceptibility to childhood common acute lymphoblastic leukaemia is associated with polymorphic peptide-binding pocket profiles in HLA-DPB1*0201. *Hum Mol Genet* 2002;**11:**1585–97.

[65] Zaccone P, Burton OT, Cooke A. Interplay of parasite-driven immune responses and autoimmunity. *Trends Parasitol* 2008;**24:**35–42.