A systematic review and meta-analysis of the association between childhood infections and the risk of childhood acute lymphoblastic leukaemia

Jeremiah Hwee*,1, Christopher Tait1, Lillian Sung2,3,4, Jeffrey C Kwong1,5,6,7,8, Rinku Sutradhar1,5 and Jason D Pole1,9

1Dalla Lana School of Public Health, University of Toronto, Health Sciences Building, 155 College Street, Toronto, ON, Canada; 2Division of Haematology/Oncology, Department of Pediatrics, The Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada; 3Program in Child Health Evaluative Sciences, The Hospital for Sick Children, Peter Gilgan Centre for Research and Learning, 686 Bay Street, Toronto, ON, Canada; 4Institute of Health Policy, Management and Evaluation, University of Toronto, Health Sciences Building, 155 College Street, Toronto, ON, Canada; 5Institute for Clinical Evaluative Sciences, 2075 Bayview Avenue, Toronto, ON, Canada; 6Department of Family and Community Medicine, University of Toronto, Toronto, ON Canada; 7Public Health Ontario, Toronto, ON, Canada; 8Toronto Western Family Health Team, University Health Network, Toronto, ON, Canada and 9Pediatric Oncology Group of Ontario, 480 University Avenue, Suite 1014, Toronto, ON, Canada

**Background:** To determine whether childhood infections were associated with the development of childhood acute lymphoblastic leukaemia (ALL).

**Methods:** We included studies that assessed any infection in childhood prior to the diagnosis of ALL in children aged 0–19 years compared to children without cancer. The primary analysis synthesised any infection against the odds of ALL, and secondary analyses assessed the frequency, severity, timing of infections, and specific infectious agents against the odds of ALL. Subgroup analyses by data source were investigated.

**Results:** In our primary analysis of 12,496 children with ALL and 2,356,288 children without ALL from 38 studies, we found that any infection was not associated with ALL (odds ratio (OR) = 1.10, 95% CI: 0.95–1.28). Among studies with laboratory-confirmed infections, the presence of infections increased the odds of ALL by 2.4-fold (OR = 2.42, 95% CI: 1.54–3.82). Frequency, severity, and timing of infection were not associated with ALL.

**Conclusions:** The hypothesis put forward by Greaves and others about an infectious aetiology are neither confirmed nor refuted and the overall evidence remains inadequate for good judgement. The qualitative difference in the subgroup effects require further study, and future research will need to address the challenges in measuring infectious exposures.

The aetiology of childhood acute lymphoblastic leukaemia (ALL) is largely unknown, and likely arises from interactions between exogenous and/or endogenous exposures, genetic susceptibility, and chance. Genetic causes of ALL account for a small proportion of cases, and while the disease is usually initiated in utero, other promotional exposures are probably necessary for disease emergence (Greaves et al., 2003). There are two key hypotheses on infections and the development of ALL. Kinlen proposed the
‘population mixing’ hypothesis to describe the observed increased rates of childhood ALL following an influx of migrants into rural areas (Kinlen, 1988, 2012). Briefly, the mixing of rural, isolated individuals with the influx of mostly urban individuals into a rural area would create a localised epidemic of an underlying infection due to the increased level of contact between susceptible and infected individuals that may produce the rare response of ALL. Studies from Kinlen and others have found evidence to support the hypothesis (Kinlen, 1988, 2006, 2012; Alexander et al, 1998; Kinlen and Doll, 2004). The hypothesis suggests a direct pathological role of a specific infection, presumed to be viral, in the development of ALL and that a protective effect may be acquired from previous exposure. Currently, there is limited molecular evidence that implicates a specific infection (Martin-Lorenzo et al, 2015; da Conceicao Nunes et al, 2016). Greaves’ ‘delayed infection’ hypothesis for childhood ALL suggests a two-hit model that emphasises the timing of exposure and the child’s immune system (Greaves, 1997, 2006). The first hit occurs in utero through one’s genetic makeup that produces a pre-leukaemic clone. In a small number of pre-leukaemia carriers, it is the absence of exposure to infections in early life, and a postnatal secondary genetic event caused by a delayed, stress-induced infection (second hit) on the developing, ‘unprepared’ immune system that may increase the risk of childhood ALL. Although the mechanisms differ, both hypotheses suggest that ALL is a rare response to one or more common infections acquired through personal contact.

The difficulties in measuring exposure to infectious agents and subsequent responses make it challenging to directly test the hypotheses, especially since no specific leukaemogenic agent has been identified. Several previous epidemiological studies have used a history of infections as an indicator for early exposure to infections. Establishing the timing of the infections is critical to testing the hypotheses; however, birth cohort studies are not feasible given the rarity of childhood ALL. Thus, most studies used a case–control design and interviews to measure infections. Assessing a history of infections through interviews can be problematic due to the potential for recall bias and misclassification of children who had asymptomatic infections (Simpson et al, 2007). Other methods for measuring infections such as using administrative data overcome these limitations, but may lack information on important confounders. Other than narrative summaries (McNally and Eden, 2004; Buffler et al, 2005; Ma et al, 2009; Maia Rda and Wunsch Filho, 2013), no study has attempted to synthesise and quantitatively pool studies examining the relationship using a history of infections, or tried to explain the differences between the studies. The aim of this systematic review and meta-analysis was to assess the relationship between childhood infections, and the development childhood ALL by summarising the findings for an overall measure of infections, the frequency, severity, timing of infections, and examining specific infectious agents and syndromes.

MATERIALS AND METHODS

The Meta-analysis of Observational Studies in Epidemiology (MOOSE) was developed as a guideline for the reporting of meta-analyses of observational studies in epidemiology and was used for the current study (Stroup et al, 2000).

Data sources and searches. We performed electronic searches from inception to 21 February 2017 in Ovid MEDLINE, MEDLINE In-Process and Other Non-Indexed Citations, EMBASE, Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index for both Science and Social Science & Humanities), and Scopus. Supplementary Table 1 shows the search strategies used. Text words used included acute lymphoblastic leukaemia, acute leukaemia, infection, virus, and bacteria. We limited the search to subjects 0–19 years old, and did not restrict the search by language. References of the included studies were searched, and the first four pages of a Google search using the same keywords were used to search for grey literature.

Study selection. We defined the inclusion and exclusion criteria a priori as studies of any design excluding editorials, reviews, and case reports. Studies were included if: (1) the primary exposure of interest included a prior history of any infection before the diagnosis of childhood ALL; (2) the primary outcome of interest was defined as clinically diagnosed ALL in children aged ≤19 years; (3) comparisons were made against a control or comparison group; and (4) testing samples must have been collected and assessed prior to treatment, if laboratory investigations were used to determine past infections. Infections must have been reported by the parent or guardian, or obtained through other data sources such as medical records.

We excluded studies based on the following order: (1) definition for infections was not at the individual level, for example, at an ecological level that examines infections aggregated for a region; (2) definition for infections that examined population mixing; (3) infections were not explicitly infections during childhood (e.g., infections during pregnancy); (4) outcomes was not childhood ALL in children aged ≤19 years; (5) absence of a comparison group; (6) it was a review article; and (7) duplicate publication with the same study population. When more than one publication from a study was available, the most recent version, or the version with the exposure or outcome of interest that was closest to the objectives of this review was included. Studies were not restricted by publication status, and relevant studies in other languages were translated.

Two reviewers (JH and CT) independently evaluated the titles and abstracts of publications identified by the search strategy, and any publication thought to be potentially relevant by either reviewer was retrieved in full. Final inclusion of studies in the systematic review was determined by agreement of both reviewers. Agreement between reviewers was evaluated using the kappa statistic ($\kappa$). Strength of agreement was defined as slight ($\kappa=0.00–0.20$), fair ($\kappa=0.21–0.40$), moderate ($\kappa=0.41–0.60$), substantial ($\kappa=0.61–0.80$), or almost perfect ($\kappa=0.81–1.00$) (Landis and Koch, 1977).

Data extraction and quality assessment. Data extraction was conducted in duplicate (JH and CT) using a standard form, which collected information on: the primary exposure of ‘common infections’, defined as any infection occurring from birth to the diagnosis of ALL; secondary exposures of infection frequency, severity of infections; and study design, region, publication era, and source of controls. In studies that used laboratory investigations for identification of infectious agents, we extracted IgG antibody estimates that were not available, the polymerase chain reactions (PCR) method was extracted to assess for the presence of the agent. We extracted infections occurring in the first year of life or similar time windows in cases with multiple time windows, as we felt this best represented early exposure to infections. We extracted infection frequency levels for common infections, and defined severity based on admission to hospital. The adjusted models that incorporated the most confounders for our primary outcome ALL were extracted. Authors were contacted for further information regarding results that were not presented. Five authors were contacted (Nishi and Miyake, 1989; Schlehofer et al, 1996; Neglia et al, 2000; Rosenbaum et al, 2005; MacArthur et al, 2008), and three responded with no additional information (Nishi and Miyake, 1989; Neglia et al, 2000; Rosenbaum et al, 2005).
Study quality was assessed using the Meta Quality Appraisal Tool (MetaQAT) (Rosella et al, 2015) and the Critical Appraisal Skills Programme (CASP) for case–control (Programme CAS, 2014a), and cohort studies (Programme CAS, 2014b). Two reviewers (JH and CT) assessed each study. For case–control studies, we considered CASP scores of 1–3, 4–6, and 7–9 to be high, moderate, and low-risk of bias, respectively; for cohort studies, we considered CASP scores of 1–4, 5–8, and 9–11 to be high, moderate and low-risk of bias, respectively.

Data synthesis and analysis methods. Our analysis combined data at the study level. Our primary analysis sought to assess exposure to common infections vs no common infections (referent group) on the risk of developing ALL, relying on each study’s definition. The most frequent infection was used when studies did not report a common infection variable. We used the adjusted odds ratio (OR) or rate ratio (RR) to calculate a pooled overall effect, and assumed OR and RR were equivalent due to the rarity of the outcome (Greenland, 1987); ORs or RRs <1 suggest infections are protective against ALL. If a study presented multiple frequency categories, we used the lowest vs the highest category, a method commonly used in meta-analyses (Ba, 2016). The method described by Greenland was used to calculate the variance using the reported 95% confidence intervals (CI) (Greenland, 1987). We calculated a crude OR for studies not reporting one, and to facilitate the calculation we added 0.5 to all cells if one of the four cells reported a zero (Gart and Nam, 1988). In secondary analyses, we used the different exposure levels of infection to compute a regression slope (Greenland and Longnecker, 1992). If an exposure level was defined using a range, we used the midpoint of the range (e.g., 1–3 infections was 2), and if the level was ≥4, we assigned a frequency of 4. For infection severity, a dichotomous variable (yes vs no) was used to determine the relationship with ALL. Post hoc analyses examined the timing of infections in the first year of life compared to infections that occurred after the first year of life, and putative infectious agents was conducted if ≥3 studies reported the agent.

As we anticipated heterogeneity between the studies, we used an inverse variance weighted average, random-effects model where the Wald-type tests and confidence intervals were estimated under a normal distribution (DerSimonian and Laird, 1986). We investigated potential sources of heterogeneity using subgroup analyses and mixed-effects meta-regression. To examine the association of study-level characteristics and infection effect, we fitted mixed-effects meta-regression models to the natural logarithm of the OR. The natural logarithm of the OR was assumed to have a normal distribution (DerSimonian and Laird, 1986). We investigated potential sources of heterogeneity using subgroup analyses and mixed-effects meta-regression models to the natural logarithm of the OR. The natural logarithm of the OR was assumed to have a normal distribution (DerSimonian and Laird, 1986). We investigated potential sources of heterogeneity using subgroup analyses and mixed-effects meta-regression models to the natural logarithm of the OR. The natural logarithm of the OR was assumed to have a normal distribution (DerSimonian and Laird, 1986). We investigated potential sources of heterogeneity using subgroup analyses and mixed-effects meta-regression models to the natural logarithm of the OR.
studies that used administrative/medical record data, we found no association between infections and ALL (OR = 1.00, 95% CI: 0.61, 1.63, P = 0.994; $I^2 = 90.8\%$). Among studies that used laboratory data, we found infections to be associated with ALL (OR = 2.42, 95% CI: 1.54, 3.82, P < 0.001, $I^2 = 54.2\%$). The interaction effect showed no difference between self-reported and administrative/medical records data sources (OR = 0.89, 95% CI: 0.54, 1.48, $P = 0.656$). Infections identified through laboratory data increased the risk of ALL compared to infections captured through self-reported data (interaction effect OR = 2.73, 95% CI: 1.71, 4.36, P < 0.001), but not administrative/medical records data sources (interaction effect OR = 2.43, 95% CI: 1.24, 4.75, P = 0.009). Among studies that used self-reported data, every additional infection reduced the odds of ALL by 4% (OR = 0.96, 95% CI: 0.94, 0.98, P < 0.001), whereas among studies that used administrative/medical records data, every additional infection increased the odds of ALL by 11% (OR = 1.11, 95% CI: 1.07, 1.15; P < 0.001). We found self-reported and administrative/medical records data sources qualitatively differed in the frequency of infections (interaction effect OR = 0.86, 95% CI: 0.83, 0.90, P < 0.001). Severity of infections remained unchanged in studies with self-reported data (OR = 1.51, 95% CI: 0.86, 2.65; P = 0.158; $I^2 = 70.2\%$). Among self-reported studies, infections in the first year of life suggested a protective effect against ALL (OR = 0.88, 95% CI: 0.80, 0.98, P = 0.017). No association was found between infections in the first year of life and ALL among administrative/medical records data (OR = 0.93, 95% CI: 0.55, 1.56, P = 0.775), and did not differ from self-reported studies (interaction effect OR = 0.95, 95% CI: 0.56, 1.62, $P = 0.862$).

The results from our primary analysis remained unchanged when we restricted the analysis to B-cell precursor ALL or B-cell common ALL (OR = 0.87, 95% CI: 0.77, 0.98, P = 0.022). Meta-regression models that assessed study level characteristics included data source, region, publication era, source of controls, and risk of bias. Data source and region accounted for the largest proportion of heterogeneity between the studies ($R^2 = 47.2\%$, see Supplementary Table 4). Stratification by risk of bias indicated studies of low-risk of bias showed similar results to our main analysis (OR = 0.92, 95% CI: 0.76, 1.10, $P = 0.349$), whereas studies of moderate-to-high-risk of bias suggested infections increased the risk of ALL (OR = 1.45, 95% CI: 1.12–1.86, P = 0.005). Compared to studies of moderate-to-high-risk of bias, studies of low-risk of bias were more likely to suggest infections were protective against ALL (OR = 0.63, 95% CI: 0.46, 0.87, P = 0.004).

**DISCUSSION**

In this systematic review of 39 studies, we found no association between any common infections, frequency, severity of infections, and timing of infections and childhood ALL. We did however, find a qualitative difference in our subgroup analyses; infections increased the odds of developing ALL by 2.4-fold in studies with laboratory investigations. Further, infections identified through laboratory investigations increased the odds of ALL by 2.7-fold and 2.4-fold compared to infections identified through self-reported and administrative/medical records data, respectively. Among studies that used self-reported data, we found each additional infection reduced the odds of ALL by 4%, and this differed significantly from studies that used administrative/medical records data that suggested each additional infection increased the odds of ALL by 11%. The heterogeneity between the studies remained a challenge and could partly be explained by differences in the data sources.

We failed to demonstrate an association in our primary analysis, but found associations in our secondary and subgroup analyses by data source. There are three plausible explanations for the observed findings. First, the apparent results may be a chance finding from multiple testing. Second, the ascertainment of infections from parental recall has been shown to under-report childhood infections and may be inaccurate in both the timing and occurrence of infections, compared to medical records (McKinney et al, 1991; Simpson et al, 2007). Despite these potential issues, studies that confirmed the self-reported infections with medical records for accuracy and completeness still found an inverse association (Dockerty et al, 1999; Ajrouche et al, 2015). Although studies that used medical records were void of recall bias,
| Study design | Case ascertainment | Control selection | Data source and collection | Selected exposure definition | Matching variables |
|--------------|---------------------|-------------------|---------------------------|-----------------------------|-------------------|
| **Ateyah et al, 2017** | CC | 45 ALL cases Single hospital | 40 controls without cancer Same hospital as cases | Laboratory investigation | EBV anti-VCA IgG | 1:1 on age and sex |
| **da Conceicao Nunes et al, 2016** | CC | 60 ALL cases Single hospital | 120 controls without cancer Same hospital as cases | Laboratory investigation | EBV anti-VCA IgG | 1:2 on age and sex |
| **Ajrouche et al, 2015** | CC | 617 cases National cancer registry | 1225 controls without cancer Population controls | Self-report: interviews | Common infections | 1:1 on age and sex |
| **Lin et al, 2015** | Co | 62 ALL cases National cancer registry | 564,573 children without cancer from the national administrative database | Administrative database | Enterovirus infection | 1:1 on sex, age, urbanisation level, parental occupation, and index year of enterovirus infection |
| **Rudant et al, 2015** | CC | 4641 ALL cases National, clinical cancer, general physician registries, and hospitals | 797 controls without cancer Birth, general physician registries, hospitals, and population quotas | Self-report: interviews, or questionnaires | Common infections | — |
| **Ibrahim et al, 2014** | CC | 40 ALL cases Single hospital | 60 healthy controls from same region | Laboratory investigation | Parvovirus B19 IgG | Age and sex |
| **Vestergaard et al, 2013** | Co | 815 ALL cases National cancer registry | 1,777,314 children without cancer from the national database | Administrative data | Hospitalisation for infections | — |
| **Ahmed et al, 2012** | CC | 54 ALL cases Single hospital | 20 controls without leukaemia Single hospital | Laboratory investigation | EBV PCR | — |
| **Chang et al, 2012** | CC | 1039 ALL cases National cancer registry | 4,140 controls without cancer National administrative database | Administrative data | Common infections | 1:1 on date of birth, sex, and time of case diagnosis |
| **Mahjour et al, 2010** | CC | 90 ALL cases Single hospital | 90 controls without ongoing cancer from single hospital | Laboratory investigation | HSV IgG | 1:1 on age and sex |
| **Rudant et al, 2010** | CC | 634 ALL cases National cancer registry | 1,494 controls without cancer Population controls | Self-report: interviews | Common infections | 1:1 on age and sex |
| **Zaki and Ashray, 2010** | CC | 40 acute leukaemia Single hospital | 20 healthy controls from the same hospital | Laboratory investigation | Parvovirus B19 IgG | Age and sex |
| **Flores-Lujano et al, 2009** | CC | 45 ALL cases with Down syndrome from six select cancer institutions in Mexico City | 218 controls with Down syndrome without leukaemia Specialised institutions exclusively for Down syndrome | Self-report: interview | Common infections | — |
| **Tesse et al, 2009** | CC | 40 ALL cases from a single hospital | 40 healthy controls from the same hospital | Laboratory investigation | EBV IgG | 1:1 on ethnic origin and socioeconomic status |
| **Cardwell et al, 2008** | CC | 112 ALL cases National population-based medical records from general physician offices | 2125 controls without leukaemia Same database as cases | Medical records: Chart abstraction | Common infections | 1:1 on physician practice, sex, and date of birth |
| **MacArthur et al, 2008** | CC | 351 ALL cases Population-based cancer registries and oncology centres | 399 controls without cancer Provincial health insurance registration database | Self-report: interviews | Varicella | 1:1 on age, sex, and area of residence |
| Study design | Case ascertainment | Control selection | Data source and collection | Selected exposure definition | Matching variables |
|--------------|--------------------|-------------------|-----------------------------|-----------------------------|-------------------|
| **Roman et al, 2007** | 425 ALL cases National population-based medical records from general physician offices | 1031 controls without cancer Same database as cases | Medical records: Chart abstraction | Common infections | T : M on region of residence at diagnosis, sex, month, and year of birth |
| **Loutfy et al, 2006** | 68 ALL cases Single hospital | 20 controls Siblings of cases | Laboratory investigation | EBV anti-VCA IgG | — |
| **Zaki et al, 2006** | 20 acute leukaemia Single hospital | 20 healthy controls from the same hospital | Laboratory investigation | Parvovirus B19 IgG | Age and sex |
| **Ma et al, 2005** | 294 ALL cases Hospital-based network registry covering 35 counties in Northern and Central California | 376 controls without cancer Random selection from state-wide birth files | Self-report: interview | Stratified by non-Hispanic white and Hispanic; Common infections | 1 : 1 and 1 : 2 on child’s date of birth, sex, mother’s race, Hispanic status, and mother’s county of residence |
| **Rosenbaum et al, 2005** | 255 ALL cases Institutional cancer registry at 4 major centres serving 31 counties | 760 controls State live birth registry | Self-report: questionnaire | Colds | 1 : M on sex, year of birth, and race |
| **Surico and Muggeo, 2005** | 82 ALL cases Single hospital | 196 controls without cancer From the same hospital as cases | Laboratory investigation | EBV anti-VCA IgG and EBNA IgG latent infection | 1 : 2 on age, sex, and comparable socioeconomic status |
| **Jourdan-Da Silva et al, 2004** | 393 ALL cases National cancer registry | 530 controls without leukaemia or lymphoma Population controls | Self-report: questionnaire | Common infections | 1 : M on age, sex, and region of residence |
| **Canfield et al, 2004** | 97 ALL cases with Down syndrome Children’s Oncology Group registration files | 173 controls with Down syndrome without leukaemia From the same physician practice as the cases | Self-report: interview | Common infections | 1 : M on age |
| **Kerr et al, 2003** | 16 acute leukaemia | 23 controls with diseases requiring cerebral spinal fluid extraction | Laboratory investigation | Parvovirus B19 PCR | — |
| **Chan et al, 2002** | 80 ALL cases Clinical database | 228 controls without leukaemia Regional controls | Self-report: interviews | Common infections | — |
| **Perrillat et al, 2002** | 219 ALL cases Hospital records from four cities in France | 237 controls without cancer Controls were from the same hospital and from same catchment area of the hospital | Self-report: interview | Repeated common infections | 1 : M on sex, age, hospital, hospital catchment area, and ethnicity |
| **Salonen et al, 2002** | 40 acute leukaemia | 39 hospital controls | Laboratory investigation | HHV-6 IgG | 1 : 1 on age, sex, and season |
| **MacKenzie et al, 2001** | 27 ALL cases | 28 children with other cancers | Laboratory investigation | EBV PCR | — |
| **Petridou et al, 2001** | 94 ALL cases Clinical database of participating centres | 94 controls Hospital controls for non-infectious reason | Laboratory investigation | Parainfluenza 1, 2 and 3 IgG | 1 : 1 on sex, age, hospital, and time period |
they were often unable to include other important confounders, such as ethnicity, parental occupation, maternal age, birth weight, and parity (Dockerty et al, 2001; Hjalgrim et al, 2004; Ma et al, 2005; Lim et al, 2014). Finally, the findings from the laboratory studies must be interpreted with caution due to the study quality, and smaller sample sizes and larger effect sizes as shown by the asymmetry of the funnel plot.

The mutational mechanisms of ALL point to three potential pathways: (1) anomalies in lineage-specific factors (ETV6-RUNX1, IKZF1, and PAX5); (2) flaws in receptor protein tyrosine kinases and their down-stream pathways; and (3) epigenetic modifiers (Whitehead et al, 2016). Recent developments in genome and mouse model studies may change our initial understanding of the aetiology of ALL as new studies have generated new hypotheses with respect to identifying potential infectious candidates (Martin-Lorenzo et al, 2015; Swaminathan et al, 2015). The presence of parvovirus B19 IgG antibodies is associated with the presence of ETV6-RUNX1 (Ibrahem et al, 2014), and is associated with certain class II HLA alleles that are risk factors for the development of childhood ALL. Furthermore, parvovirus B19 has certain characteristics similar to other oncoviruses, that is, its DNA genome persists indefinitely in human tissues following acute infection, causing mild or no disease, and upregulates pro-inflammatory cytokines associated with ALL onset (Kerr and Mattey, 2015). The results from the small laboratory studies will require confirmation in larger population studies. Since half of 15-year-old adolescents have specific antiparvovirus B19 antibodies (Young and Brown, 2004), the measurement of the clinical syndromes caused by parvovirus B19 may be preferred to assess manifestations of the pathogen. Parvovirus B19 infection may provide only a subset of an oncogenic hit in a multistep carcinogenesis process.

The qualitative differences in our findings support the hypothesis of an alternative pathway for ALL development. Recent qualitative reviews have attempted to explain the positive association between infections and ALL and suggested studies that used medical records or administrative data may be capturing children with an earlier than expected altered immune system. These children may respond differently to infections, have a greater propensity to seek medical care when infections are contracted, and/or have a stronger immune response (Wiemels, 2012; Whitehead et al, 2016). The sensitivity to infections may be due to a lack of immunomodulation from lower levels of anti-inflammatory cytokine interleukin-10 in newborns who later go on to develop ALL (Chang et al, 2011).
As in previous reviews, there continues to be substantial heterogeneity among the studies; however, our review focuses on specific objectives and highlights the recent developments of the field (McNally and Eden, 2004; Buffler et al., 2005; Greaves, 2006; Ma et al., 2009; Maia Rda and Wunsch Filho, 2013). There are several limitations of this study. The heterogeneity between the

| Author, year | Odds ratio (95% CI) | Weights (%) |
|--------------|---------------------|-------------|
| **Primary analysis – any infection** | | |
| Ateyah et al., 2017 | 4.06 (1.20, 13.60) | 1.12 |
| da Conceloa Nunes et al., 2016 | 1.31 (0.54, 1.98) | 2.42 |
| Lin et al., 2015 | 0.43 (0.26, 0.69) | 3.18 |
| Ajrouche et al., 2015 | 0.75 (0.57, 0.99) | 4.20 |
| Rudert et al., 2015 | 0.95 (0.87, 1.04) | 4.86 |
| Ibrahim et al., 2014 | 3.62 (1.45, 8.78) | 1.75 |
| Vestergaard et al., 2013 | 0.92 (0.78, 1.07) | 4.68 |
| Ahmed et al., 2012 | 30.59 (1.76, 531.87) | 0.25 |
| Chang et al., 2012 | 3.18 (2.17, 4.66) | 3.69 |
| Mahjoub et al., 2010 | 3.87 (1.96, 7.65) | 2.38 |
| Rudert et al., 2010 | 0.70 (0.60, 0.90) | 4.52 |
| Zaki and Ashray, 2010 | 27.61 (1.56, 488.95) | 0.24 |
| Flores-Lujano et al., 2009 | 1.45 (0.64, 3.30) | 1.93 |
| Tesse et al., 2009 | 1.12 (0.70, 1.41) | 1.56 |
| Cardwell et al., 2008 | 1.05 (0.64, 1.74) | 3.13 |
| MacArthur et al., 2008 | 1.01 (0.69, 1.48) | 3.70 |
| Roman et al., 2017 | 1.30 (0.90, 1.80) | 3.87 |
| Loffty et al., 2006 | 0.27 (0.03, 2.25) | 0.43 |
| Zaki et al., 2016 | 27.88 (1.48, 526.12) | 0.23 |
| Ma et al., 2005 - Hispanic | 1.74 (0.80, 3.76) | 2.07 |
| Ma et al., 2005 - Non Hispanic White | 0.79 (0.40, 1.51) | 2.37 |
| Rosenbaum et al., 2006 | 0.97 (0.64, 1.51) | 3.56 |
| Surico and Muggeo, 2005 | 1.36 (0.75, 2.46) | 2.71 |
| Canfield et al., 2004 | 0.52 (0.28, 0.96) | 2.63 |
| Jourdan-Da Silva et al., 2004 | 0.80 (0.60, 1.00) | 4.00 |
| Kerr et al., 2003 | 16.92 (1.03, 17.18) | 0.42 |
| Chan et al., 2002 | 0.77 (0.38, 1.57) | 2.28 |
| Perrilat et al., 2002 | 0.60 (0.40, 1.00) | 3.32 |
| Salomen et al., 2002 | 3.25 (0.32, 32.68) | 0.37 |
| MacKenzie et al., 2001 | 0.83 (0.22, 3.14) | 0.97 |
| Petridou et al., 2001 | 1.90 (1.10, 3.20) | 2.97 |
| Neglia et al., 2000 | 0.71 (0.50, 1.01) | 3.84 |
| Dockerty et al., 1999 | 1.29 (0.68, 2.46) | 2.52 |
| McKinnon et al., 1999 | 0.49 (0.26, 0.95) | 2.50 |
| Schuz et al., 1999 | 1.00 (0.80, 1.20) | 4.52 |
| Schliehofer et al., 1996 | 1.63 (0.91, 2.92) | 2.76 |
| Nath and Miyake, 1989 | 1.71 (0.86, 3.37) | 2.37 |
| van Steensel-Moll et al., 1986 | 0.80 (0.60, 1.00) | 4.30 |
| Till et al., 1979 | 2.90 (0.84, 9.96) | 1.09 |
| Total (95% CI) | I² = 76%, Q (P < 0.001), Test for overall effect (P = 0.187) | 100 |

| Secondary analysis – Frequency of Infections | | |
| Ajrouche et al., 2015 | 0.93 (0.87, 0.99) | 12.48 |
| Rudert et al., 2015 | 0.97 (0.94, 0.99) | 14.37 |
| Chang et al., 2012 | 1.11 (1.07, 1.15) | 14.02 |
| Rudert et al., 2010 | 0.94 (0.88, 0.99) | 12.47 |
| Cardwell et al., 2008 | 1.19 (0.97, 1.45) | 5.03 |
| Ma et al., 2005 - Hispanic | 1.13 (0.97, 1.31) | 6.80 |
| Ma et al., 2005 - Non-Hispanic White | 0.95 (0.83, 1.09) | 7.70 |
| Canfield et al., 2004 | 0.81 (0.62, 1.08) | 3.01 |
| Jourdan-Da Silva et al., 2004 | 1.12 (0.89, 1.41) | 4.11 |
| Chan et al., 2002 | 0.79 (0.50, 1.26) | 1.27 |
| Perrilat et al., 2002 | 1.13 (0.81, 1.56) | 2.35 |
| Neglia et al., 2000 | 0.93 (0.89, 0.99) | 12.87 |
| Dockerty et al., 1999 | 1.09 (0.84, 1.40) | 3.51 |
| Total (95% CI) | I² = 80%, Q (P < 0.001), Test for overall effect (P = 0.977) | 100 |

| Secondary analysis – Severity of Infections | | |
| Ajrouche et al., 2015 | 1.48 (0.91, 2.41) | 23.11 |
| Vestergaard et al., 2013 | 0.92 (0.78, 1.07) | 36.69 |
| Flores-Lujano et al., 2009 | 3.45 (1.37, 8.66) | 11.16 |
| Van Steensel-Moll et al., 1986 | 1.00 (0.70, 1.40) | 29.04 |
| Total (95% CI) | I² = 71%, Q (P = 0.015), Test for overall effect (P = 0.286) | 100 |

| Secondary analysis - Parovirus B19 | | |
| da Conceloa Nunes et al., 2016 | 2.20 (1.02, 4.76) | 20.37 |
| Ibrahim et al., 2014 | 3.62 (1.49, 8.78) | 19.29 |
| Zaki and Ashray, 2010 | 27.61 (1.56, 488.95) | 6.31 |
| Zaki et al., 2006 | 27.88 (1.48, 526.12) | 6.11 |
| Kerr et al., 2003 | 16.92 (1.03, 77.18) | 9.34 |
| Petridou et al., 2001 | 1.10 (0.70, 1.90) | 22.63 |
| Schliehofer et al., 1996 | 0.48 (0.14, 1.69) | 15.94 |
| Total (95% CI) | I² = 72%, Q (P = 0.001), Test for overall effect (P = 0.020) | 100 |

![Figure 2. Random-effects model examining the association between common infections and risk of childhood acute lymphoblastic leukaemia. CI represents confidence interval. Common infections are reported as a two-class variable, or highest vs lowest in more than two categories. The secondary analysis for frequency of infections is a combined maximum likelihood effect estimate that estimates a trend from summarised dose-response data. The presence of parovirus B19 was measured as a dichotomous variable, presence of IgG antibodies vs no IgG antibodies for parovirus B19. All other studies, the reference was no infections.](image-url)

www.bjcancer.com | DOI:10.1038/bjc.2017.360
Childhood infections and risk of childhood ALL

Ahmed HG, Osman SI, Ashankyty IM (2012) Incidence of Epstein–Barr virus in pediatric leukemia in the Sudan. Clin Lymphoma Myeloma Leuk 12(2): 127–131.

Aţeyah ME, Hashem ME, Abedelsalam M (2017) Epstein-Barr virus and regulatory T cells in Egyptian paediatric patients with acute B lymphoblastic leukaemia. J Clin Pathol 70(5): 418–424.

Ateyah ME, Hashem ME, Abdelsalam M (2017) Epstein–Barr virus and regulatory T cells in Egyptian paediatric patients with acute B lymphoblastic leukaemia. J Clin Pathol 70(5): 219–223.

Brown RC, Dwyer T, Kasten C, Krotoski D, Li Z, Linet MS, Olsen J, Scheidt P, Winn DM (2007) Cohort Profile: The International Childhood Cancer Cohort Consortium (ICCC). Int J Epidemiol 36(A): 724–730.

Buffler PA, Kwan ML, Reynolds P, Urayama KY (2005) Environmental and genetic risk factors for childhood leukaemia: appraising the evidence. Cancer Epidemiol 29(3): 52–63.
concordance with general practitioner records. J Public Health Med 13(1): 13–22.

McKinney PA, Juszczak E, Findlay E, Smith K, Thomson CS (1999) Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study. Br J Cancer 80(11): 1844–1851.

McNally RJ, Eden TO (2004) An infectious aetiology for childhood acute leukaemia: a review of the evidence. Br J Haematol 127(3): 243–263.

Metayer C, Milne E, Clavel J, Infante-Rivard C, Petridou E, Taylor M, Schuz J, Spector LG, Dockerty JD, Magnani C, Pombo-de-Oliveira MS, Sinnett D, Murphy M, Roman E, Monge P, Ezzat S, Mueller BA, Scheurer ME, Armstrong BK, Birch J, Kaatsch P, Koliman S, Lightfoot T, Bhati P, Bondy ML, Rudant J, O’Neill K, Miligi L, Desyssyris N, Kang AY, Buffer PA (2013) The Childhood Leukemia International Consortium. Cancer Epidemiol 37(3): 336–347.

Neglia JP, Linet MS, Shu XO, Severson RK, Potter JD, Mertens AC, Wen W, Kersey JH, Robisson LL (2000) Patterns of infection and day care utilization and risk of childhood acute lymphoblastic leukaemia. Br J Cancer 82(1): 234–240.

Nishi M, Miyake H (1989) A case–control study of non-T cell acute lymphoblastic leukaemia of children in Hokkaido, Japan. J Epidemiol Community Health 43(4): 352–355.

Paltiel O, Laniado DE, Yanetz R, Deutsch L, Calderon-Margalit R, Harlap S, Friedlander Y (2006) The risk of cancer following hospitalization for infection in infancy: a Population-Based cohort study. Cancer Epidemiol Biomarkers Prev 15(10): 1964–1968.

Perrillat F, Clavel J, Auclerc MF, Baruchel A, Leverger G, Nelken B, Philippe N, Schaison G, Sommelet D, Vilmer E, Hemon D (2002) Day-care, early common infections and childhood acute leukaemia: a multicentre French case–control study. Br J Cancer 86(7): 1064–1069.

Peeters JL, Sutton AF, Jones DR, Abrams KR, Rushton L (2008) Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J Clin Epidemiol 61(10): 991–996.

Petridou E, Dalamaga M, Mentis A, Skalkidou A, Moustaki M, Karpathios T, Trichopoulos D (2001) Evidence on the infectious etiology of childhood leukaemia: the role of low herd immunity (Greece). Cancer Causes Control 12(7): 645–652.

Programme CAS (2014a) CASP Case Control Checklist. CASP. Oxford.

Programme CAS (2014b) CASP Cohort Study Checklist. CASP. Oxford.

Riley AW, Duncan GI (2016) Completing a national birth cohort in the united states. JAMA Pediatr 170(9): 829–830.

Roman E, Simpson J, Ansell P, Kinsey S, Smith A, Ansell P, Roman E (2007) Childhood leukaemia and primary prevention. Br J Cancer 86(7): 1064–1069.

Rossella L, Bowman C, Pach B, Morgan S, Fitzpatrick T, Goel V (2015) The development and validation of a meta-tool for quality appraisal of public health evidence: Meta Quality Appraisal Tool (MetaQAT). Public Health 136: 57–65.

Rosenbaum PF, Buck GM, Brecher ML (2005) Allergy and infectious disease histories and the risk of childhood acute lymphoblastic leukaemia. Paediatr Perinat Epidemiol 19(2): 152–164.

Rudant J, Lightfoot T, Urayama KY, Petridou E, Dockerty JD, Magnani C, Milne E, Spector LG, Ashton LJ, Desyssyris N, Kang AY, Miller M, Rondelli R, Simpson J, Stiakaki E, Orsi L, Roman E, Metayer C, Infante-Rivard C, Clavel J (2015) Childhood acute lymphoblastic leukaemia and indicators of early immune stimulation: a Childhood Leukemia International Consortium study. Am J Epidemiol 181(8): 549–562.

Rudant J, Orsi L, Menegais F, Petit A, Baruchel A, Bertrand Y, Lambilliotte A, Robert A, Michel G, Margueritte G, Tandonnet J, Mechinaud F, Bordigoni P, Hemon D, Clavel J (2010) Childhood acute leukaemia, early common infections, and allergy: The ESCALE Study. Am J Epidemiol 172(9): 1015–1027.

Salonen MJH, Siimes MA, Salonen EM, Vaheri A, Koskineni M (2002) Antibody status to HHV-6 in children with leukaemia. Leukemia 16(4): 716–719.

Schlehofer B, Blettner M, Geletneky K, Haaf HG, Kaatsch P, Michaelis J, MuellerLantzsch N, Niehoff D, Winkelspecht B, Wahrendorf J, Schlehofer JR (1996) Sero-epidemiological analysis of the risk of virus infections for childhood leukaemia. Int J Cancer 65(5): 584–590.

Schuz J, Kaletsch U, Meintert R, Kaatsch P, Michaelis J (1999) Association of childhood leukaemia with factors related to the immune system. Br J Cancer 80(3–4): 585–590.

Simpson J, Smith A, Ansell P, Roman E (2007) Childhood leukaemia and infectious exposure: a report from the United Kingdom Childhood Cancer Study (UKCCS). Eur J Cancer 43(16): 2396–2403.

Stroup DF, Berlin JA, Moher D, Olkin I, Ruiz-Wehrenberg W, Rennie D, Ross G, Rubin P, Tolleson J, Richardson W (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. JAMA 283(15): 2008–2012.

Surico G, Muggeo P (2005) Epstein–Barr virus infection at the onset of acute lymphoblastic leukaemia in children. Int J Cancer Austral Asia 4(1): 19–24.

Swaminathan S, Klemm L, Park E, Papatamianu E, Ford A, Kweon SM, Trageser D, Hasselfeld B, Henke N, Mooster J, Geng H, Schwarz K, Kogan SC, Casellas R, Schatz DG, Lieber MR, Greaves MF, Muschen M (2015) Mechanisms of clonal evolution in childhood acute lymphoblastic leukaemia. Nat Immunol 16(7): 766–774.

Tesse R, Santoro N, Giordano P, Cardinale F, De Mattia D, Armenio L (2009) Association between DEFB1 gene haplotype and herpes viruses seroprevalence in children with acute lymphoblastic leukaemia. Pediatr Hematol Oncol 26(8): 573–582.

Till M, Rapson N, Smith PG (1979) Family studies in acute leukaemia in childhood: a possible association with autoimmune disease. Br J Cancer 40(1): 62–71.

Urayama KY, Buffer PA, Gallagher ER, Ayooob JM, Ma X (2010) A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia. Int J Epidemiol 39(3): 718–732.

van Steensel-Moll HA, Valkenburg HA, van Zanen GE (1986) Childhood leukaemia and infectious diseases in the first year of life: a Register-Based Case–Control Study. Am J Epidemiol 124(4): 590–594.

Vestergaard TR, Rostgaard K, Grau K, Schmiegelow K, Hjalgrim H (2013) Hospitalisation for infection prior to diagnosis of acute lymphoblastic leukaemia in children. Pediatr Blood Cancer 60(3): 428–432.

Viechtbauer W (2010) Conducting meta-analyses in R with the metafor Package. J Stat Softw 36(4): 48.

Whitehead TP, Metayer C, Wiemels JL, Singer AW, Miller MD (2016) Childhood leukaemia and primary prevention. Curr Prob Pediatric Adolesc Health care 46(10): 317–352.

Wiemels J (2012) Perspectives on the causes of childhood leukaemia. Br J Cancer 89(6): 261–266.

Zaki ME, Hassan SA, Seleim T, Lateef RA (2006) Parvovirus B19 infection in children with acute leukaemia. Int J Lab Hematol 38(2): 159–166.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)