Infection, immune responses and the aetiology of childhood leukaemia

Mel Greaves

Abstract | Childhood leukaemia is the principal subtype of paediatric cancer and, despite success in treatment, its causes remain enigmatic. A plethora of candidate environmental exposures have been proposed, but most lack a biological rationale or consistent epidemiological evidence. Although there might not be a single or exclusive cause, an abnormal immune response to common infection(s) has emerged as a plausible aetiological mechanism.

Acute leukaemia is the principal subtype of paediatric cancer in affluent societies, where its incidence rate is within the range of 30–45 per 10^6 children per year[1]. This translates to an accumulative risk of ~1 in 2000 up to the age of 15 and an annual incidence in the United Kingdom of around 500 cases. In most developing or less-affluent countries, incidence data is much less reliable, with few population-based registries. However, where data is available, the rate of childhood leukaemia seems to be several-fold lower — for example, in parts of India and the Caribbean[1,2] — and perhaps up to tenfold lower for the common variant of acute lymphoblastic leukaemia (ALL) in black South Africans[3].

Controlled clinical trials of chemotherapy in the United States and Europe have steadily improved the survival of children with acute leukaemia to the current figure of around 85% (REFS 4–6). This incremental success depended in part on the recognition of prognostic variables and biological subtypes of the disease, which merited differential treatment[6].

However, much still remains to be achieved. Not all children worldwide have access to optimal treatment, and success, when achieved, might come with a price[7,8]. Additionally, cure rates remain modest or poor in high-risk groups — for example, children with particular chromosomal abnormalities such as *BCR–ABL* or *MLL–AF4* fusions or advanced disease (high white-blood-cell count)[6,9]. Efficacious therapy that is biologically targeted but less toxic is, as in other cancers, the current goal. In this context, the question of what the causes of childhood leukaemia might be deserves some attention.

Sleuthing 'the cause'

Much speculation and epidemiological endeavour in identifying the underlying causal mechanisms in childhood leukaemia has, in retrospect at least, been naive and simplistic. The idea that there might be a singular and exclusive cause of all cases — that is, 'the cause' — is patently absurd. Leukaemia, in common with most other cancer types, is not like malaria or tuberculosis. Similarly, the implicit premise that 'cause' should be attributable solely to 'an exposure' is to ignore biology. Leukaemia, most if not all cancers and, indeed, most human ailments are the result of combinatorial impacts of crucial exposures, modifying influences, inherited susceptibility and chance. Current genetic epidemiology takes this unavoidable complexity on board. An additional dimension that, ultimately, will be a requisite of any comprehensive solution is that of evolutionary biology. The central tenet of Darwinian or evolutionary medicine is that the causation of human illness or disease cannot be fully comprehended outside the context of the ancient and more recent historical contingencies that have shaped our genetics and physiology[10]. This principle applies to the most prevalent cancers and leukaemia[11] as well as it does to infectious diseases, diabetes and obesity[10–12].

Over the past few decades or so, the vacuum of definitive data on the causation of childhood leukaemia has, by default, helped spawn a plethora of explanations and theories[13,14] (BOX 1). Some accord with widespread and entrenched mythology embellished by an, at times, alarmist and uncritical media. Others remain plausible but unsubstantiated, whereas a few (for example, vitamin K injection at birth) are now ruled out.

The case for radiation. The only established causal exposure for childhood leukaemia is ionizing radiation. This unambiguous conclusion derives from data on Japanese atomic bomb survivors from 1945 who were acutely exposed to up to ~200 mSv (REF. 15) and, at a much lower dose level (~10 mSv), from historical

The Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, United Kingdom.
e-mail: mel.greaves@icr.ac.uk
doi:10.1038/nrc1816
Published online
9 February 2006

At a glance

- Childhood leukaemia is the most common paediatric cancer in developed societies.
- The disease is biologically heterogeneous and no single or exclusive causal mechanism is likely.
- The natural history of paediatric leukaemia usually involves pre-natal initiation of pre-leukaemic clones (frequently by chromosome translocation) followed by postnatal promotion, secondary mutation and overt disease. Latency after initiation can be very variable (a few months to 15 years).
- Ionizing radiation is an accepted cause of leukaemia but not a significant cause. Non-ionizing radiation (for example, electromagnetic field radiation) seems to be a very weak or negligible cause.
- Large, case–control epidemiology studies are required that incorporate biological subtypes of disease and inherited alleles associated with susceptibility. These studies also need to be driven by plausible biological hypotheses.
- Two infection-based hypotheses have been proposed and assessed: Kinlen's 'population-mixing' hypothesis and Greaves' 'delayed-infection' hypothesis.
- The body of epidemiological evidence now available is consistent with the view that many childhood leukaemias arise as a consequence of an abnormal immune response to common infection(s), but the mechanisms remain to be determined.

data on diagnostic exposure of the fetus from X-ray pelvimetry during pregnancy[16]. But does it not follow that ionizing radiation is a principal or significant cause of *de novo* leukaemia?

Natural, background ionizing radiation delivers exposures of some 2–3 mSv per year. This is both cosmic (γ-rays) and terrestrial in origin, and most of the latter is from gaseous radon, which delivers primarily to the bronchial epithelium. Based on the known dose–response relationship from the Japanese 1945 exposure and the lower-level diagnostic exposures, the International Commission on Radiological Protection has adopted a theoretical model of risk estimation. This estimation extrapolates to very low and protracted levels of exposure with the double premise that dose–response for cancer risk is linear at very low doses, and that there is no threshold dose of radiation for mutation and leukaemia[17–19]. If this were to be correct then, given the relative rarity of childhood leukaemia but the very large number of individuals exposed, natural or background radiation could account for 20–25% of all cases[18,19]. The biophysical rationale for no threshold is based on the argument that single radiation tracks will impact stochastically and independently. But it also assumes that the relevant biological response — mutation and leukaemogenesis — is autonomous or non-cooperative. It is not clear if this assumption can accommodate chromosome translocations and multi-step leukaemogenesis.

Box 1 | **Postulated causal exposures**

- Car exhaust fumes
- Pesticides
- Ionizing radiation
- Non-ionizing electromagnetic fields
- Electric fields
- Vitamin K injection at birth
- Hot dogs or hamburgers (depending on whether the consumer (patient) was in California or Colorado)
- Domestic animals
- Organic dust from cotton, wool or synthetic fibres
- Natural light deprivation through melatonin disruption
- Artificial, fluorescent light exposure in hospital neonatal care units
- Parental cigarette smoking
- Maternal medicinal drug taking (during pregnancy)
- Maternal alcohol consumption (during pregnancy)
- Chemical contamination in drinking water
- Infections

Epidemiologically assessing the risks of very-low-dose radiation is very difficult, but the UK Children's Cancer Study Group (UKCCS) found no association between variable, measured levels of natural γ-radiation or radon in the home and risk of childhood leukaemia[20,21]. These considerations and uncertainties indicate that the suggestion that a substantial fraction of childhood leukaemias might be caused by natural ionizing radiation should be treated with some considerable circumspection.

More severe problems of biological plausibility plague interpretation of the available data on non-ionizing electromagnetic field radiation (EMF). The pooled epidemiological data indicates that exposure to fields of >0.4 μT, through proximity to pylons or from household wiring and electrical appliances, might double the risk of childhood leukaemia[22]. The statistics are not themselves impressively robust, and the findings could be due to chance or to selection bias in controls. Proving that exposure to EMF never causes any case of leukaemia is impossible, but the inherent weakness of the data, coupled with the absence of any verified biological mechanism by which such low-energy fields might contribute to cell transformation or disease promotion, relegates EMF to unimpressive-candidate status. As should consideration of the sobering fact that the Earth's natural magnetic field delivers a dose of 30–70 μT (albeit static, not oscillating). The worst-case scenario in the United Kingdom, given that so few people are actually exposed to >0.4 μT, is that somewhere between 1 or 2 extra cases a year (out of a total of 500) could derive from artificial EMF exposure[23]. Others argue that it is not electromagnetic fields but low-frequency alternating currents of electrical fields in the houses of children that are relevant to leukaemogenesis[24]. Or, alternatively, exposure downwind to pollutant aerosols generated by the corona ions emanating from high-voltage overhead power lines[25]. There is no consistent epidemiological data to endorse these contentions and no supportive biological evidence. The UKCCS found no association between the risk of leukaemia and measured electrical fields in homes or proximity to major power supplies[26,27].

What makes a good causal candidate? Any candidate causal mechanism for childhood leukaemia should satisfy a number of criteria. These should include

Box 2 | **Human blood cell cancers and infection***

- Epstein–Barr virus (EBV) is the causative agent in Burkitt lymphoma (B-cell) and some Hodgkin lymphomas.
- Human herpesvirus 8 (HHV8) is the causative agent in primary effusion lymphoma and the plasma-cell variant of multicentric Castleman disease. These herpesviruses can transform lymphoid cells and are resident as episomal virions. The disease risk is increased by immunosuppression.
- Human T-cell lymphotropic virus 1 (HTLV1) retrovirus is the causative agent in adult T-cell leukaemia/lymphoma (and tropical spastic paraparesis) in adults. The virus transforms cells and can integrate into the host DNA. Person to person infection through cellular transfer of virus (for example, through blood or milk) might occur.
- *Helicobacter pylori* is a cause of gastric lymphoma (B-cell) in adults[‡], which is mediated by an influx of activated T cells or other inflammatory cells[§]. Hepatitis C might cause some non-Hodgkin lymphoma through indirect mechanisms (possibly by antigen-driven proliferation).

*See REFS 58,129 for reviews on this subject. [‡]Other types of specific bacterial infections seem to be associated with lymphomas of mucosa-associated lymphoid tissue in different sites, including chlamydia (in the eye), campylobacter (in the small intestine) and borrelia (in the skin)[130]. [§]Treatment of cases with antibiotic regimens before extensive clonal evolution has occurred results in resolution of disease — that is, the driver has been removed[131].

an accommodation of the biological heterogeneity of disease[28]. There is already evidence that different aetiological mechanisms are linked to subtypes of haematological cancer in both adults and children. This includes infection-associated B- or T-cell leukaemia/lymphoma (BOX 2), acute leukaemias that arise from an inherited or constitutive predisposition[29], acute myeloid leukaemia (AML) linked to benzene exposure[30], and secondary AML elicited by previous therapeutic exposure to alkylating agents or topoisomerase-II-inhibitory drugs[31]. *A priori* therefore, it is highly unlikely that all childhood blood-cell cancers will share a common or exclusive cause. However, it is not intuitively obvious which subtypes or level of subclassification should be regarded as important in this respect. Only very-large-scale studies, as in the current UKCCS, have been designed to accommodate the hierarchy of biological subsets, which includes cell type (by morphology and immunophenotype), chromosome karyotype and molecular abnormalities[32]. In terms of previous hypotheses with biological plausibility, the most unambiguous classification separates out infant acute leukaemias with *MLL* gene fusions from common B-cell precursor ALL (cALL) in the peak incidence range of 2–5 years as biologically distinct subsets for which there are grounds for suspecting distinctive aetiologies[33]. A body of genetic epidemiological evidence now indicates that infant leukaemia might be caused by transplacental chemical carcinogenesis[33–38].

Any plausible hypothesis should also explain the appearance of the very marked peak of incidence of cALL at 2–5 years in the Western world between 1920 and 1945 (but later in Japan and China) and the slow but significant increase since (~1% per annum) in affluent societies[39,40]. And, crucially, it should have a biological rationale that is consistent with current understanding of the natural history of the disease and molecular pathways of leukaemogenesis and that is testable by current genetic epidemiological methods.

The underlying biology and natural history

In the context of the subtypes of leukaemia and their possibly distinctive aetiologies, the natural history of the disease — the timing and number of rate-limiting mutations — is crucial. This time-ordered framework has now been unravelled by molecular scrutiny of leukaemia fusion genes in identical twins with concordant leukaemia[41,42], in retrospective analyses of archived neonatal blood spots[43,44], and from large-scale cord-blood screening[45]. These studies have provided distinctive natural histories for infant leukaemia versus childhood ALL and AML (reviewed in REFS 46,47). These profiles have an impact on considerations of aetiological mechanisms and their timing.

For infant ALL, relevant exposures and complementary genetic susceptibility factors impact exclusively *in utero* during fetal haematopoiesis[46]. For more typical childhood ALL and AML a minimal and spaced two-step model is relevant, with two distinct temporal 'exposure' windows: one *in utero*, when leukaemia is commonly initiated through chromosomal rearrangements, and a second, postnatal window that is linked to emergence of overt disease through secondary genetic changes (FIG. 1). Pertinent to this evolutionary sequence is the finding that initiation of disease *in utero* by chromosomal translocation seems to exceed the rate of overt ALL by some 100-fold (REF. 45). A similar proportionality might apply to other childhood tumours, including Wilms tumours and neuroblastoma[45], indicating that prenatal initiation of paediatric cancers might be very common but that transition to overt disease is rare. This must impact on considerations of what kinds of exposures are likely to be relevant.

The most frequent genetic lesions in cALL are hyperdiploidy and *TEL–AML1* (*ETV6–RUNX1*) fusions.

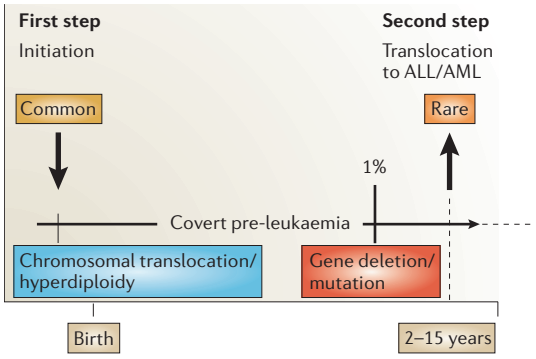


Figure 1 | **A 'minimal' model illustrating the crucial sequential events in the development of childhood leukaemia.** 1% refers to the frequency of transition between covert pre-leukaemia and overt clinical leukaemia[45,47]. A similar sequence of events occurs in Down syndrome-associated acute megakaryoblastic leukaemia, in which *GATA1* mutations arise in fetal liver and, in the context of constitutive trisomy 21, establish a pre-leukaemic state [132]. Note that infant acute lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML) (<1 year of age) has a much-abbreviated natural history in which all the necessary genetic events (including *MLL* fusion) are thought to occur prenatally[46].

One percent of normal newborn infants have a *TEL–AML1* fusion gene that seems to be functional — that is, it is in frame, in the correct cell type and is present as an expanded clone[45]. The mechanism of generating this and other, similar chimeric fusion genes is DNA double-strand breakage followed by error-prone non-homologous end-joining repair[47]. These mechanisms are in themselves effectively blind to gene function and cell type as well as to whether resultant fusions do or do not contribute to leukaemogenesis. Therefore, the 1% figure of functional fusions must be distilled or clonally selected from a much greater frequency of non-functional or inconsequential breaks and fusions. We see this with *BCR–ABL* fusion in which high-sensitivity screening for non-selected fusions indicates that almost everyone generates this molecular lesion[47]. A reasonable inference is that whatever causes chromosome translocations *in utero* to initiate leukaemia is possibly ubiquitous. Whatever that might be, it follows that the second, postnatal 'hit' is a crucial bottleneck in the development of clinical leukaemia.

We assume, without direct evidence, that the relevant promoting or 'precipitating' postnatal exposures and crucial secondary genetic events occur proximal to diagnosis itself, perhaps preceding it by something like 3 to 12 months. However, as most childhood leukaemias present with no overt pre-clinical phase, the timing of these key events is uncertain.

Animal modelling with the *TEL–AML1* fusion gene has endorsed this two-step model. Transfection of murine stem cells with *TEL–AML1*-expressing vectors leads to sustained expansion of B-cell precursors and a partial differentiation block, but no overt leukaemia[48–50]. Leukaemia might, however, emerge when mice expressing a *TEL–AML1* transgene are crossed with mice bearing other single oncogenic lesions — for example, INK4A-null animals[51]. The same dual requirement is observed with the myeloid leukaemia fusion gene *AML1–ETO*. In this instance, overt myeloid leukaemia can be promoted by exposure to low-dose chemical carcinogens[52].

An infectious aetiology?

Earlier in the past century, infections were regarded as the most likely cause of childhood leukaemia[53]. The prime rationale for this view was that the age distribution of the disease was similar to that of common childhood infectious diseases such as measles, and many patients had a record of infections (mostly respiratory) before or coincident with diagnosis. Some clinicians even had the audacity to suggest what kind of infectious insult might be involved[54,55]. Consider for example the statement "We incline on our evidence to the belief that the solution of the problem of leukaemia lies rather in some peculiar reaction to infection than in the existence of some specific infective agent" (J. Poynton, H. Thursfield and D. Paterson, Great Ormond Street Hospital for Sick Children, London, 1922)[54], contrasting with "The remaining hypothesis is one which postulates a specific infection" (C. E. Kellet, Royal Infirmary, Newcastle, 1937)[55]. Although the observational basis on which their speculations were based was fragile, there is a striking echo of these ideas in current hypotheses (see below). Infection as a cause fell out of favour, albeit for the wrong reason, when it was recognized that the disease was not, in the clinical sense, contagious[56].

Later, however, the discovery that the retrovirus human T-cell lymphotropic virus 1 (HTLV1) was a causative factor for adult T-cell leukaemia/lymphoma[57], that leukaemia in domestic cattle, cats and chickens was viral in origin[58], and evidence from other haematological cancers (BOX 2) helped sustain the idea that childhood leukaemia might have an infectious, possibly viral, cause. In 1988, two specific hypotheses were proposed that implicated infection as an important causal factor in childhood leukaemia.

In the early 1980s, apparent clusters of childhood leukaemia and non-Hodgkin lymphoma (NHL) around the nuclear reprocessing plants in Sellafield (or Seascale village) in Cumbria, England, and at Dounreay (Thurso town) on the extreme north coast of Scotland had attracted considerable public and media attention[59]. The popular assumption was that it had to be either environmental radioactive contamination from the plants or transmission of germline mutations by exposed parents working at the plants. These ideas were fuelled by revelations of environmental leaks of radioactivity[59] and by preliminary epidemiological data on parental occupational exposures[60]. However, Leo Kinlen noted, as others had, that the evidence implicating radiation, particularly from recorded exposure levels, was hardly credible, and he argued that other possibilities should be entertained. He suggested that the increased incidence of childhood leukaemia and NHL might have infectious origins that were due to the unusual population mixing that occurred when the isolated communities of Seascale and Thurso were enlarged to accommodate the influx of migrant professional workers[61]. He assessed the credibility of this explanation by analysing a parallel 'rural new town' situation of Glenrothes in Scotland and documented a transient increase (around threefold) in the incidence rate of childhood leukaemia (measured by mortality statistics) contemporaneous with the relatively abrupt influx of substantial numbers of newcomers[61]. This and subsequent studies throughout the United Kingdom and elsewhere (see below) provided a firm basis for suggesting that some childhood leukaemia clusters might be an unusual outcome of a common but relatively non-pathological infection arising in individuals who were non-immune, and following contact or 'population mixing' with carriers. Kinlen favoured the view, by analogy with animal leukaemias, that this might involve transformation with a hitherto unidentified virus. In Kinlen's studies, most patients have the common form of disease, ALL, but the data have been interpreted to indicate that all subtypes of childhood and infant leukaemia, as well as NHL, can share a common aetiology.

Also in 1988, I proposed a 'delayed-infection' hypothesis for childhood leukaemia[62]. This model sought to explain the peak incidence of cALL at 2–5 years that, in geographical comparisons and time-trend data, seemed to be a feature of modern or more affluent societies[3]. The proposed explanation had its roots in the biology

Non-homologous end-joining repair
Predominant, cellular mechanism for the repair of double-stranded DNA breaks; error-prone as no template sequence is used.

Time–space clustering
Increased incidence of cases in one place in a defined time period. This can occur by chance. A vivid illustration of this is the finding of statistically significant clusters in the United Kingdom in proximity to 'military sites' that turn out to be derelict medieval forts (R. Cartwright, personal communication).

of the disease and in immunology. It proposes a role for infection in the context of a minimal 'two-hit' model of the natural history of the disease, which at the time (1988) was speculative but is now substantiated (see above). Infection was postulated to have a crucial role in promoting, through the immune response, the crucial second or postnatal genetic error or 'hit'. The essential element of this proposal was based on an evolutionary argument[33]. This was that the immune cell network (that is, T-cell subsets and cytokine signalling) and repertoire had been programmed by historical events (lethal infections and genetic selection) to anticipate infectious exposure neonatally and in infancy, and that this exposure was essential for adaptive organisation of the immune network for future compliant and efficient response. This basic concept enjoys substantial experimental support[63,64]. Absence or diminution of infections early in life is a feature of more affluent 'hygienic' societies that, overall, has produced substantial benefits in terms of reduced infant mortality. Paradoxically, however, such infectious insulation might predispose the immune system to aberrant or pathological responses following subsequent or 'delayed' exposure, which in this context is predicted to precipitate ALL through proliferative or apoptotic stress (see below). The risk of leukaemia under such circumstances would also be contingent on the existence of a prenatally generated pre-leukaemic clone, and a genetic background that influences the strength of signals in particular components of the immune network.

An essentially similar immunological argument is embodied in the 'hygiene hypothesis' for childhood allergies[65,66] and also applied to some autoimmune diseases in children and young adults, including type I diabetes[67] and multiple sclerosis[68]. It is possible that these divergent pathologies reflect the same underlying mismatch between evolutionary programming of the immune system and contemporary patterns of infection in early life[11]. Delayed exposure to common infection(s) has also been proposed for the aetiology of Hodgkin lymphoma in young adults[69]. In the latter case, the precedent or model was with poliovirus and the timing of virus infection in relation to the development of the nervous system[70].

Whereas the Kinlen model assumes that defective immunosurveillance or protective immunity against virus X is most probably at fault, my model predicted that an abnormal (and positive) immune response to one or more common infections (bacterial or viral) was responsible, albeit indirectly, for promoting leukaemia. Although the Kinlen 'population-mixing' hypothesis and the 'delayed-infection' hypothesis differ in detail and hypothetical mechanism, they do share common ground. Both posit that childhood leukaemia is a rare response to one or more common infections acquired by personal contact under particular 'modern' socio-demographic circumstances.

Epidemiological evidence for infection

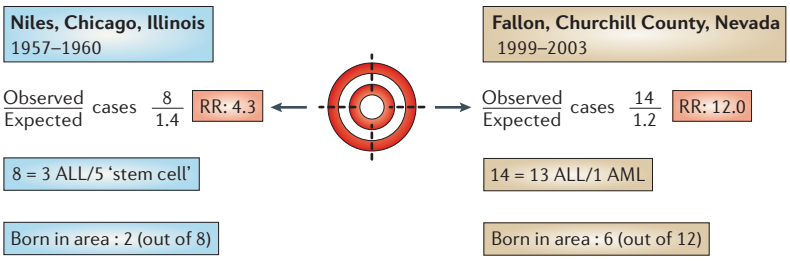
Over the past 15 years or so, a substantial body of epidemiological data has accumulated from studies that have sought to address this issue, although few have been designed to specifically address the two prime hypotheses. Some are descriptive studies, others are case versus control comparisons. They vary significantly in several crucial parameters that influence quality, likelihood of confounding and robustness of statistics (including control selection, size and methodology) and accuracy of recorded exposures. Two extensive audits and commentaries on these data were published in 1999 (REF. 71) and in 2004 (REF. 72). The most recent of these surveys concludes that although conflicting data exists, the weight of evidence overall is supportive of both population mixing and delayed infection in infancy as being significant causal factors.

Clusters, population mixing and leukaemia

The first recorded cluster of childhood leukaemias was in Niles, a suburb of Chicago in Cook County, Illinois in 1957–1960 (REF. 73) (BOX 3). The cluster had some striking characteristics. All eight cases occurred in the single parish of St John Brebeuf, and all patients and/or their older siblings attended the same school. The presentation of the cases was paralleled by what was described as a 'rheumatic-like' illness. The authors drew the prescient conclusion that the cause was most likely a consequence of infection possibly tied in with the coincident rapid population expansion within the parish[73]. Other clusters of varying magnitude have been reported[71,72], with the most recent and dramatic being that in the small desert town of Fallon, Nevada, where proximity to a very large naval air-base has led to the view that leakage of carcinogenic jet fuel is responsible[74]. Litigation lawyers and moviemakers are there in droves, but more prosaic explanations are possible. Overall, childhood leukaemia shows relatively modest evidence for time–space clustering[71,72,75,76].

Box 3 | **Leukaemia cluster busting; then and now**

Almost all cases in these two examples of leukaemia clusters are acute lymphoblastic leukaemia (ALL). The first recorded cluster of childhood leukaemias was in Niles, a suburb of Chicago in Cook County, Illinois in 1957–1960. All eight cases occurred in the same parish and all patients and/or their older siblings attended the same school. In the most recent case — the small desert town of Fallon, Nevada — analysed cases have been predominantly common B-cell precursor ALL (cALL) with the usual spectrum of chromosomal karyotypes (J. Wiemels, personal communication), indicating that they are typical ALL. This makes the postulated jet fuel (polycyclic hydrocarbons and benzene) an unlikely aetiological agent. In the case of Niles, 6 of the 8 cases were born outside the cluster area and moved in between 1 and 10 years of age. Similarly, in Fallon, 6 of the first 12 cases were born outside the cluster zone, indicating that the crucial infectious exposure time is unlikely to be around either the time of birth or *in utero*. AML, acute myeloid leukaemia; RR, relative risk (the ratio of observed over expected cases, with the latter coming from population-based estimates).



Subsequent to his findings at Seascale, Dounreay and Glenrothes[61], Kinlen has recorded a series of unusual space–time population mixing events associated with occupational or military activities, and each seems to be associated with a temporarily increased risk of childhood leukaemia[71,72,77–81]. Independent studies have provided confirmatory data[71,72], including a systematic study of population mixing and leukaemia/NHL incidence throughout the area of Cumbria in north-east England[82] and an increased risk of ALL in the new territories of Hong Kong coincident with population increases and mixing[83]. Although a few negative results have been reported[71,72], the overall consistency of these data strongly indicate that population mixing is a significant causal factor in childhood leukaemia. This explanation could well apply to the currently contentious cluster in Fallon, Nevada[74,84].

It is difficult to envisage that population mixing provides any exposure other than infection. What is less obvious is the extent to which this particular explanation applies to childhood leukaemia in general. Additionally, these descriptive epidemiological studies, convincing though they may be, provide little insight into mechanisms or timing of the putative infectious exposures in relation to the natural history of the disease. Some authors have interpreted their data on population mixing to suggest that the crucial infectious exposure time might be around the time of birth[82]. Others have suggested that it might occur *in utero*[85]. The data from the Niles and Fallon clusters make this timing improbable. In the case of Niles, 6 of the 8 cases were born outside the cluster area and moved in between 1 and 10 years of age. Similarly, in Fallon, 6 of the first 12 cases were born outside the cluster zone (BOX 3). Assuming that the cases of ALL in these clusters have the 'conventional' natural history with prenatal initiation, then these features indicate that for many, if not all, cases the relevant exposure was most probably a postnatal promotional event for leukaemogenesis.

Delayed infection in infancy and leukaemia

Time–space clustering is also compatible with the delayed-infection hypothesis, but more-specific epidemiological predictions of this hypothesis relate to the timing of infections or opportunities for infection early in life[33,62]. As no individual infectious species is nominated as a candidate, only known common infections of infancy or proxies for infections can be used as markers of exposure. In practice, the patterns of exposure, timing of infections in the first year of life and the immunological response to such challenges will have multifactorial determinants. These will include breastfeeding practices and mothers' immune status, family sibship size and age range, social factors such as hygiene conditions and interactions with other children of the same or older age, community factors such as population density, mobility, age structure and infectious history. Additionally, genetic factors that govern the immune response are likely to be relevant[143]. Currently, we have no algorithm to compute the overall likelihood of infection and immunological response under these highly variable circumstances. Investigators have therefore analysed individual parameters that might impact on risk.

The simplest prediction of the hypothesis is that patients with childhood leukaemia might be expected, when compared with controls, to have, on average, fewer recorded common infections in the first year of life, and less social contact and potential for infectious exposure outside the home — for example, through day-care attendance. With respect to the first criterion, there is a difficulty of disparities between mothers' recall and medical records (UKCCS, unpublished data). Several studies have documented a decreased incidence in common upper-respiratory-tract infections, including otitis media (ear infection), in cases compared with controls, but a few studies also failed to find such a difference or found an association in the opposite direction[71,72]. One obvious difficulty here, apart from the variable design and size of these studies and accuracy of data, is that the relevant infections, with respect to leukaemia, may or may not prompt symptoms and so may or may not result in a documented diagnosis. In this unfavourable circumstance, well-recognized proxies for infectious exposure might be preferable.

The likelihood of acquiring common infections is well-recognized epidemiologically as being highly dependent on social and physical contacts, and the numbers or frequency of such contacts. A substantial body of evidence indicates that playgroup or day-care attendance in early life increases the likelihood of acquiring infections such as *Haemophilus influenzae B*, cytomegalovirus and common infections of the upper respiratory tract, and that the probability of infection relates to the level of contact in terms of numbers of children in the social group and frequency of attendance[86–89]. Attendance at day-care facilities in infancy is reported to reduce the risk of allergies[90] and type I diabetes[91] in accord with the predictions of the hygiene hypothesis. The UKCCS was designed to evaluate these parameters (as well as several other exposures that are potentially relevant to leukaemia and other cancers). It was the largest, most systematic study of its kind, involving 3,838 children with cancer (1,736 with leukaemia) and 7,629 unaffected controls[32]. The Northern California case–control study of childhood leukaemia adopted the same questionnaire and will eventually accrue a similar large number of cases. The UKCCS have now reported their findings[92], as have the Northern California study group with their preliminary but still extensive data set[93,94]. Both of these studies record that attendance at playgroup during the first year of life provides significant protection against childhood ALL (TABLE 1). The Northern California study also provides preliminary evidence that the extent of protection from ALL that is derived from day-care attendance is quantitatively proportional to the number and frequency of social contacts[93]. If this association holds true when the full cohort of ~1,000 patients are analysed, then it will provide compelling evidence.

There are some caveats to these findings. In the UKCCS, there also seemed to be some (lesser) degree of protection against other paediatric cancers (in particular,

Table 1 | **Day-care attendance in infancy and risk of childhood ALL**

Country	Number of cases	Period	Odds ratio* (confidence interval‡)	References
Greece	136	0–2 years	0.28 (0.09–0.88)	97
New Zealand	121	0–1 year	0.65 (0.36–1.17)§	133
Quebec	491	0–1 year	0.49 (0.31–0.77)	98
Hong Kong	98	0–1 year	0.63 (0.38–1.07)§	134
France	240	From birth onwards	0.6 (0.4–1.0)	99
France	408	0–3 months	0.6 (0.4–0.8)	100
California (**a**)	140	0–1 year	0.6 (0.45–0.95)	93
California (**b**) ‖	294	0–1 year	0.42 (0.18–0.99)	94
United Kingdom**	1286	0–1 year	0.48 (0.37–0.62)¶ 0.69 (0.51–0.93)#	92
United States of America‡‡	1744	0–6 months	0.91 (0.72–1.15)§	101

*Compared with 1.0 (no day-care attendance). ‡95% confidence intervals of odds ratio. §Not significant. ‖Analysis includes the 140 cases in **a**. ¶ALL cases versus healthy controls. #ALL cases versus non-ALL malignancies. **Carried out by the UK Children's Cancer Study Group. ‡‡Carried out by the USA-based Children's Cancer Group and the National Cancer Institute. ALL, acute lymphoblastic leukaemia.

central nervous system tumours (UKCCS investigators, unpublished data)) that was conferred by early playgroup attendance[92]. Some investigators regard this as an indication of some unknown bias or confounding, and at present this cannot be ruled out. Alternatively, an impact on other cancers might be real and therefore a finding of considerable interest that is worthy of further exploration. Intriguingly, there is some independent epidemiological support for the idea that some central nervous system tumours might have an infectious aetiology[95,96]. The latest data from the Northern California study group also includes the finding that day-care attendance is protective for non-Hispanics but not Hispanics[94]. The investigators rationalize this finding with the observation that the social arrangements of Hispanic families facilitate multiple social contacts (and infectious opportunities) within the family setting, and this might obviate any increase in risk owing to lack of formal day-care attendance.

Whilst the UKCCS was in progress, several other groups published data on day-care attendance and leukaemia risk. Some of these studies are flawed by failing to accommodate the crucial aspect of timing of exposure (that is, the first year of life) and/or were too small in size to provide statistically robust results. Nevertheless, four such studies found a significant protective effect of day-care attendance in early life[97–100] (TABLE 1). None of the studies have reported an increased risk of leukaemia associated with day-care attendance in infancy. Overall, these studies are concordant and supportive of the delayed-infection hypothesis. However, one large study stands out as providing conflicting data. The USA-based Children's Cancer Group (CCG) and the National Cancer Institute (NCI) analysed day-care attendance, retrospectively, in a large case–control study and found no association with overall pre-diagnosis day-care attendance or day-care attendance during the first year of life[101] (TABLE 1). It is not easy to reconcile these conflicting data. One possibility, reflected in both the Northern California and UKCCS data sets, is that levels or numbers of contacts are crucial and that this was not sufficiently accommodated in the design of the CCG study. Another potential difficulty is bias in control selection. This is an extraordinarily difficult aspect of case–control epidemiological studies and no currently available method is without bias, including that used by the UKCCS. Nevertheless, many epidemiologists in Europe and the USA now regard the random digit telephone-dialling selection method used by the CCG study group as flawed and it is possible that this resulted in a selection bias.

Other epidemiological data

Birth order or parity status might be anticipated to influence the risk of childhood ALL, as those born with older siblings should be more exposed to common infections in infancy than first-borns. Studies of parity versus risk of childhood leukaemia have produced inconsistent results overall[72], although the largest of these studies found a significantly decreasing risk with birth order for ALL but not AML — that is, the risk was greatest for firstborn[102].

Other data on seasonal variation in timing of birth or diagnosis of patients has provided limited and indirect evidence to support an infectious aetiology for childhood ALL[71,72]. A low level of protection against ALL is observed with protracted breastfeeding in several, but not all, studies[72]. Vaccination in infancy might be expected to have some impact on risk either by general modulation of the immune system or through specific immunological sensitization to particular microorganisms or groups of microorganisms. Overall, analyses of vaccination histories have produced conflicting and inconsistent results[71,72,103]. An important exception might be with *Haemophilus influenzae B*, which seems to be protective against childhood ALL[103–105].

Worldwide incidence data indicates a significant correspondence between rates of childhood allergies, type I diabetes and ALL[106,107]. This indicates that there might well be shared infectious or immunological risk factors. At the level of individuals, however, there seems to be a striking reciprocity between the risk of allergies versus childhood ALL[108,109]. This provides at least a hint that opposing immunological pathways could be involved along with distinctive genetic susceptibility profiles.

A search for mechanistic evidence

The classical epidemiological approach to identifying causal exposures for childhood leukaemia through case–control scrutiny of exposures might soon exhaust its potential, although multinational cooperative studies on this relatively rare disease could prove valuable, particularly when incorporating genetic analysis. In this context, the availability of high-throughput single-nucleotide polymorphism arrays that can be customized to genes in pertinent pathways relevant to prenatal exposures (for example, immune response or cytokine

Representative difference analysis
This method assesses whether there are any sequences in the DNA of patients' leukaemic cells that are not of endogenous origin and is based on subtracting leukaemic ('tester') DNA from constitutive or normal, non-leukaemic ('driver') DNA of the same individual.

T-helper (T_H)1 and T_H2 response
A T_H1 cell-mediated immune response is mediated by pro-inflammatory cytokines such as interferon-γ, interleukin-1β and tumour-necrosis factor. It promotes cellular immune responses against intracellular infections and malignancy. A T_H2 response involves production of cytokines, such as interleukin-4, that stimulate antibody production. T_H2 cytokines promote secretory immune responses of mucosal surfaces to extracellular pathogens, and allergic reactions.

genes, or chemical detoxification) should prove valuable. Infection has emerged as arguably the most plausible candidate exposure that is involved in the aetiology of most cases of childhood ALL. What is required now is further biological and genetic evidence that is pertinent to mechanisms by which infection might impact on risk. Answering the following questions might be helpful.

Is there any evidence for exogenous viral sequences in leukaemic cells? If childhood leukaemia mirrors leukaemia in cats, cattle and chickens, HTLV1-associated adult T-cell leukaemia or Epstein–Barr virus-associated Burkitt lymphoma, there follows an expectation that viral sequences should be present and detectable in leukaemic cells from patients. Molecular screening for candidate viruses or virus families has, to date, produced negative results (TABLE 2); however, the list of candidates is by no means exhausted, and the relevant virus could have been missed. A potentially more productive experiment is to screen for anonymous, non-human sequences by representative difference analysis. This method has recently been applied to DNA samples from 11 cases of childhood cALL[110]. No foreign sequences were detected in a high-sensitivity screen that could detect single-copy sequences of ≥9 kb (REF. 110). These data seem to provide compelling evidence against direct viral transformation. Representative difference analysis is not, however, 100% foolproof. It is conceivable, though perhaps unlikely, that a very small virus or proviral component might be involved, or that a transforming virus has operated through a hit-and-run mechanism.

Can infection-derived proliferative stress of pre-leukaemic cells be modelled? The infection hypothesis would gain more support if evidence for the second hit could be modelled and specific cellular and biochemical mechanisms implicated.

A candidate mechanism (FIG. 2) is that a deregulated T-cell response to infection results in a potent inflammatory response and a cytokine/chemokine release profile that, at least transiently, suppresses haematopoiesis and/or induces apoptosis. In this context, the pre-leukaemic clone could have a distinctive survival and/or proliferative advantage. A regenerative wave of expansion of these cells might then result in secondary mutations, either by chance or through the genotoxic influence of an inflammatory milieu of oxidative stress. Inflammatory lesions are well-recognized promoters of carcinogenesis[111]. It might be feasible to test this idea *in vivo*, exploiting the now available *TEL–AML1*-driven pre-leukaemia murine models (see above), but a more tractable approach could come from *in vitro* systems. My laboratory has now developed an inducible system for the expression of *TEL–AML1* in an appropriate cellular context that should provide a test bed for assessing the possible biochemical basis of selection of pre-leukaemic clones[112]. Intriguingly, when TEL–AML1 protein expression is induced, cell growth is slowed down and the expression levels of some 250 genes are altered. As anticipated from functional studies of transcriptional deregulation by TEL–AML1 (REFS 113,114), most altered genes (~200 out of 250) are downregulated, and this includes some known direct targets of AML1 and several pro-apoptotic genes. The prediction to be tested is that in this model system, cells that express TEL–AML1 will have a demonstrable survival and proliferative advantage when challenged with one or more apoptotic signals that are known to be activated in T-cell responses to infection, such as interferon-γ, tumour-necrosis factor (TNF) or transforming growth factor-β (TGFβ). If one or more such cytokines, or particular pathways, are clearly implicated in selection, then *in vivo* modelling would be viable.

Does genetic variation in the signal strength of the immune-response pathways influence risk of leukaemia? Further insight into mechanisms of infection-based aetiology will almost certainly derive from evaluation of genetic variation in risk. Although strong inherited predisposition only has a role in a small minority (≤5%) of cases of childhood leukaemias[29], some degree of inherited susceptibility through normal allelic variation is very likely to exist. A recent study provides evidence that inherited variation in genes interacts with infectious environmental exposures in infancy to determine the patterns (for example, cytokine profiles) of immune responses[143]. The obvious candidate genes, in the context of the infection hypotheses, include those that are recognized as encoding crucial components of the immune network, including highly polymorphic human leukocyte antigen (HLA) genes in the major histocompatability complex (MHC) loci, cytokines, chemokines and their receptors. Knowledge of immune system molecules and genetic screening technologies are now mature enough for this test to be applied with a prospect of meaningful data. An encouraging precedent is with childhood diseases that are thought to have an infectious aetiology operating through the hygiene hypothesis route, with delayed and/or abnormal immune responses[115]. Interleukin-12 (IL-12) has a crucial role in regulating the balance of T-helper (T_H)1 and T_H2 responses, with its activity favouring T_H1 responses[116]. It is therefore intriguing that allelic variations in the *IL12B* gene that are associated with low or high expression of IL-12

Table 2 | **Molecular screening for viral sequences in ALL**

Virus screened for	Screening method	References
Polyoma viruses JC and BK	Specific PCR	135–137*
Parvovirus B19	Specific PCR	138‡
Human herpesvirus family (HHV4, 5, 6, 7 and 8)	PCR using degenerate primers	139‡,140
Bovine leukaemia virus	Southern blotting	141
TT virus¶	Specific PCR§	141,142
Exogenous microbial sequences	Representative difference analysis	110

*In diagnostic leukaemic samples (blood or marrow) or ‡in archived neonatal blood spots (Guthrie cards) of patients. Note that the absence of viral sequences in Guthrie cards constitutes no negative evidence if the critical-exposure window is postnatal. ¶TT virus, named after a patient (TT), is a novel but common DNA virus. Its taxonomy is uncertain, but it might be a circovirus. §Some positivity for TT virus was recorded in relapse samples but not in the initial diagnostic samples[142]. ALL, acute lymphoblastic leukaemia.

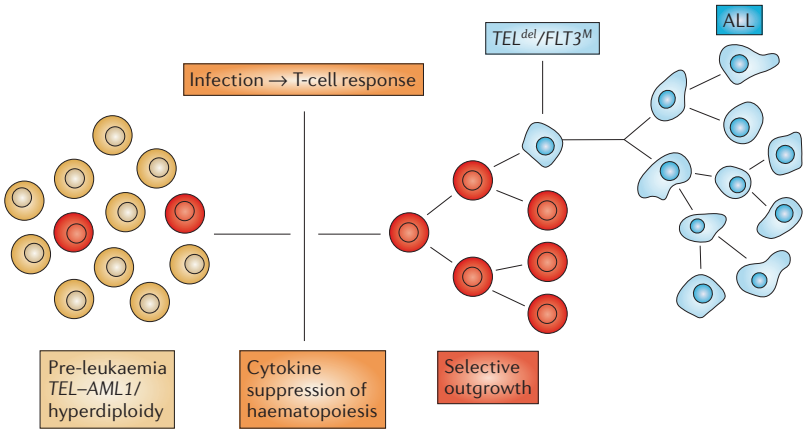


Figure 2 | **A model for infection-derived proliferative stress in the selection of pre-leukaemic cells.** This figure shows a speculative model of how infection-mediated immunological/inflammatory responses might provide a selective milieu for pre-leukaemic clones (red). A deregulated T-cell response to infection results in a potent inflammatory response and a cytokine/chemokine release profile that, at least transiently, suppresses haematopoiesis and/or induces apoptosis. In this context, the pre-leukaemic clone could have a distinctive survival and/or proliferative advantage. A regenerative wave of expansion of these cells might then result in secondary mutations, either by chance or through the genotoxic influence of an inflammatory milieu of oxidative stress, and facilitate the emergence of overt leukaemic cells (blue). *TEL*^{del} is a deletion of the normal *TEL* allele and is a very common occurrence in *TEL–AML1*-associated acute lymphoblastic leukaemia (ALL)[47]. *FLT3*^{M} mutations that activate FLT3 kinase are a relatively frequent (~30%) alteration in infant leukaemia and in childhood ALL with hyperdiploidy.

Aplasia
Defective production of red and white blood cells by the bone marrow.

seem to be associated with an increased risk of asthma or type I diabetes, respectively[117,118]. A possible evolutionary rationale for these genetic associations is that previous genetic selection for protection against lethal pandemics or epidemics of infectious diseases now, paradoxically, endows the descendents of this Darwinian largesse with vulnerability to disease if the expected infections do not occur[119,120]. The UKCCS has provided evidence for an association between the risk of childhood cALL and certain HLA alleles, specifically an increased risk with *HLA-DPB1*0201* (REF. 121). Others have previously reported associations between childhood ALL and HLA-DR alleles[122]. These findings would benefit from independent confirmation and extension. However, when taken at face value, these findings provide evidence that the immune system, and most probably its response to infection, is indeed involved in childhood ALL.

Can particular infections be implicated? These endeavours would still leave us short of knowing the identity of the putative infections that might promote childhood leukaemia, and ultimately this information will be required. At present, there is no evidence incriminating individual species of viruses or bacteria; either or both could be involved, and any assumption of an exclusive offending microbe, as yet to be discovered, might, as in most autoimmune diseases, prove illusory.

For childhood leukaemia, further and more-detailed HLA analysis might assist screening of candidate microbes through the structural constraints of associated peptide-binding regions[123]. There might also be clinical situations that provide rare windows of opportunity. One such opportunity exists with a small subgroup of children with ALL who have a previous pre-clinical phase of aplasia[124,125] that is almost always associated with infection (albeit either cause or effect) and which can involve a variety of viral or bacterial species[124,126]. The aplasia is transient and either spontaneously remits or is very responsive to corticosteroids[124,125], and precedes ALL by some 2 to 9 months. Immunoglobulin or T-cell receptor gene analysis in the few cases that have been studied confirm, as expected, the presence of the leukaemic, or pre-leukaemic, cells in the aplastic phase[127,128]. Could these patients and their aplasias signal the tip of the iceberg that needs to be uncovered — the infection-triggered selection of pre-leukaemic clones?

Ultimately, these mechanistic insights will be important both for endorsing the credibility of the aetiological hypotheses and for encouraging the prospect of some kind of prophylactic vaccination that might prevent a substantial proportion of childhood leukaemias.

1. *International Incidence of Childhood Cancer* (eds Parkin, D. M. *et al.*) (IARC Scientific Publications, Lyon, 1988).
2. *International Incidence of Childhood Cancer Volume II* (eds Parkin, D. M. *et al.*) (IARC Scientific Publications No. 144, Lyon, 1998).
3. Greaves, M. F. *et al.* Geographical distribution of acute lymphoblastic leukaemia subtypes: second report of the collaborative group study. *Leukemia* **7**, 27–34 (1993).
4. Pinkel, D. Lessons from 20 years of curative therapy of childhood acute leukaemia. *Br. J. Cancer* **65**, 148–153 (1992).
5. Kersey, J. H. Fifty years of studies of the biology and therapy of childhood leukemia. *Blood* **90**, 4243–4251 (1997).
6. Pui, C.-H., Relling, M. V. & Downing, J. R. Acute lymphoblastic leukemia. *N. Engl. J. Med.* **350**, 1535–1548 (2004).
Comprehensive, recent summary of current understanding of the biological and clinical heterogeneity of childhood leukaemia.
7. Neglia, J. P. *et al.* Second neoplasms after acute lymphoblastic leukemia in childhood. *N. Engl. J. Med.* **325**, 1330–1336 (1991).
8. Hudson, M. in *Childhood Leukemias* (ed. Pui, C.-H.) 463–481 (Cambridge University Press, Cambridge, 1999).
9. *Childhood Leukemias* (ed. Pui, C.-H.) (Cambridge University Press, Cambridge, 1999).
10. Nesse, R. M. & Williams, G. C. *Evolution and Healing. The New Science of Darwinian Medicine* (Weidenfeld & Nicolson, 1995).
11. Greaves, M. *Cancer. The Evolutionary Legacy* (Oxford University Press, Oxford, 2000).
12. Stearns, S. C. & Ebert, D. Evolution in health and disease. *Q. Rev. Biol.* **76**, 417–432 (2001).
13. Linet, M. S. & Devesa, S. S. in *Leukemia* (eds Henderson, E. S., Lister, T. A. & Greaves, M. F.) 131–151 (Saunders, Philadelphia, 2002).
14. Ross, J. A., Davies, S. M., Potter, J. D. & Robison, L. L. Epidemiology of childhood leukemia, with a focus on infants. *Epidemiol. Rev.* **16**, 243–272 (1994).
15. Preston, D. L. *et al.* Cancer incidence in atomic bomb survivors. Part III: leukemia, lymphoma and multiple myeloma, 1950–1987. *Radiat. Res.* **137** (Suppl.), S68–S97 (1994).
16. Doll, R. & Wakeford, R. Risk of childhood cancer from fetal irradiation. *Br. J. Radiol.* **70**, 130–139 (1997).
17. Brenner, D. J. *et al.* Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proc. Natl Acad. Sci. USA* **100**, 13761–13766 (2003).
Critical analysis of the contentious issue of risks from low-dose ionizing radiation.
18. Wakeford, R. The cancer epidemiology of radiation. *Oncogene* **23**, 6404–6428 (2004).
19. Wakeford, R. & Tawn, E. J. The risk to health from low doses of ionising radiation. *Nuclear Future* **1**, 107–114 (2005).
20. UK Childhood Cancer Study Investigators. The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation. 2: γ radiation. *Br. J. Cancer* **86**, 1727–1731 (2002).
21. UK Childhood Cancer Study Investigators. The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation. I: radon gas. *Br. J. Cancer* **86**, 1721–1726 (2002).
22. Ahlbom, A. *et al.* A pooled analysis of magnetic fields and childhood leukaemia. *Br. J. Cancer* **83**, 692–698 (2000).
23. UK Childhood Cancer Study Investigators. Exposure to power-frequency magnetic fields and the risk of childhood cancer. *Lancet* **354**, 1925–1931 (1999).

24. Coghill, R. W., Steward, J. & Philips, A. Extra low frequency electric and magnetic fields in the bedplace of children diagnosed with leukaemia: a case–control study. *Eur. J. Cancer Prev.* **5**, 153–158 (1996).
25. Fews, A. P., Henshaw, D. L., Wilding, R. J. & Keitch, P. A. Corona ions from powerlines and increased exposure to pollutant aerosols. *Int. J. Radiat. Biol.* **75**, 1523–1531 (1999).
26. UK Childhood Cancer Study Investigators. Exposure to power frequency electric fields and the risk of childhood cancer in the UK. *Br. J. Cancer* **87**, 1257–1266 (2002).
27. UK Childhood Cancer Study Investigators. Childhood cancer and residential proximity to power lines. *Br. J. Cancer* **83**, 1573–1580 (2000).
28. Greaves, M. Molecular genetics, natural history and the demise of childhood leukaemia. *Eur. J. Cancer* **35**, 173–185 (1999).
29. Taylor, G. M. & Birch, J. M. in *Leukemia* (eds Henderson, E. S., Lister, T. A. & Greaves, M. F.) 210–245 (WB Saunders, Philadelphia, 1996).
30. Savitz, D. A. & Andrews, K. W. Review of epidemiologic evidence on benzene and lymphatic and hematopoietic cancers. *Am. J. Ind. Med.* **31**, 287–295 (1997).
31. Smith, M. A., McCaffrey, R. P. & Karp, J. E. The secondary leukemias: challenges and research directions. *J. Natl Cancer Inst.* **88**, 407–418 (1996).
32. UK Childhood Cancer Study Investigators. The United Kingdom Childhood Cancer Study: objectives, materials and methods. *Br. J. Cancer* **82**, 1073–1102 (2000).
Detailed description of the design of the largest case–control epidemiological study of possible causes of childhood cancer.
33. Greaves, M. F. Aetiology of acute leukaemia. *Lancet* **349**, 344–349 (1997).
34. Ross, J. A., Potter, J. D. & Robison, L. L. Infant leukemia, topoisomerase II inhibitors, and the *MLL* gene. *J. Natl Cancer Inst.* **86**, 1678–1680 (1994).
35. Alexander, F. E. *et al.* Transplacental chemical exposure and risk of infant leukaemia with *MLL* gene fusion. *Cancer Res.* **61**, 2542–2546 (2001).
36. Strick, R., Strissel, P. L., Borgers, S., Smith, S. L. & Rowley, J. D. Dietary bioflavonoids induce cleavage in the *MLL* gene and may contribute to infant leukemia. *Proc. Natl Acad. Sci. USA* **97**, 4790–4795 (2000).
37. Wiemels, J. L. *et al.* A lack of a functional NAD(P)H: quinone oxidoreductase allele is selectively associated with pediatric leukemias that have *MLL* fusions. *Cancer Res.* **59**, 4095–4099 (1999).
38. Smith, M. T. *et al.* Low NAD(P)H:quinone oxidoreductase activity is associated with increased risk of leukemia with *MLL* translocations in infants and children. *Blood* **100**, 4590–4593 (2002).
39. Draper, G. J., Kroll, M. E. & Stiller, C. A. Childhood cancer. *Cancer Surv.* **19–20**, 493–517 (1994).
40. Greaves, M. F. & Alexander, F. E. An infectious etiology for common acute lymphoblastic leukemia in childhood? *Leukemia* **7**, 349–360 (1993).
41. Ford, A. M. *et al. In utero* rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature* **363**, 358–360 (1993).
42. Ford, A. M. *et al.* Fetal origins of the *TEL–AML1* fusion gene in identical twins with leukemia. *Proc. Natl Acad. Sci. USA* **95**, 4584–4588 (1998).
43. Gale, K. B. *et al.* Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc. Natl Acad. Sci. USA* **94**, 13950–13954 (1997).
44. Wiemels, J. L. *et al.* Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet* **354**, 1499–1503 (1999).
45. Mori, H. *et al.* Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc. Natl Acad. Sci. USA* **99**, 8242–8247 (2002).
Provides unambiguous evidence that pre-leukaemic clones with chromosomal translocations are generated prenatally at approximately 100-times the rate of overt leukaemia.
46. Greaves, M. F., Maia, A. T., Wiemels, J. L. & Ford, A. M. Leukemia in twins: lessons in natural history. *Blood* **102**, 2321–2333 (2003).
Details the insights into the timing and development of childhood leukaemia — these insights were derived from molecular studies of concordant leukaemia in identical twins.
47. Greaves, M. F. & Wiemels, J. Origins of chromosome translocations in childhood leukaemia. *Nature Rev. Cancer* **3**, 639–649 (2003).
48. Tsuzuki, S., Seto, M., Greaves, M. & Enver, T. Modelling first-hit functions of the t(12;21) *TEL–AML1* translocation in mice. *Proc. Natl Acad. Sci. USA* **101**, 8443–8448 (2004).
***In vivo* mimicry in mice of the pre-leukaemic status initiated by TEL–AML1.**
49. Fischer, M. *et al.* Defining the oncogenic function of the TEL–AML1 (ETV6–RUNX1) fusion protein in a mouse model. *Oncogene* **24**, 7579–7591 (2005).
50. Morrow, M., Horton, S., Kioussis, D., Brady, H. J. M. & Williams, O. TEL–AML1 promotes development of specific hematopoietic lineages consistent with preleukemic activity. *Blood* **103**, 3890–3896 (2004).
***In vivo* mimicry in murine cells of the pre-leukaemic status initiated by TEL–AML1.**
51. Bernardin, F. *et al.* TEL–AML1, expressed from t(12;21) in human acute lymphocytic leukemia, induces acute leukemia in mice. *Cancer Res.* **62**, 3904–3908 (2002).
52. Yuan, Y. *et al.* AML1–ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc. Natl Acad. Sci. USA* **98**, 10398–10403 (2001).
53. Cooke, J. V. The incidence of acute leukemia in children. *JAMA* **119**, 547–550 (1942).
54. Poynton, F. J., Thursfield, H. & Paterson, D. The severe blood diseases of childhood: a series of observations from the Hospital for Sick Children, Great Ormond Street. *Br. J. Child. Dis.* **XIX**, 128–144 (1922).
55. Kellett, C. E. Acute myeloid leukaemia in one of identical twins. *Arch. Dis. Childhood* **12**, 239–252 (1937).
56. Ward, G. The infective theory of acute leukaemia. *Br. J. Child. Dis.* **14**, 10–20 (1917).
57. *Human T-Cell Leukemia/Lymphoma Virus. The Family of Human T-Lymphotropic Retroviruses: Their Role in Malignancies and Association With AIDS* (eds Gallo, R. C., Essex, M. E. & Gross, L.) (Cold Spring Harbor Laboratory Press, New York, 1984).
58. Schulz, T. F. & Neil, J. C. in *Leukemia* (eds Henderson, E. S., Lister, T. A. & Greaves, M. F.) 200–225 (Saunders, Philadelphia, 2002).
59. Committee on Medical Aspects of Radiation in the Environment (COMARE). *First Report. The Implications of the New Data on the Releases From Sellafield in the 1950s for the Conclusions of the Report on the Investigation of the Possible Increased Incidence of Cancer in West Cumbria* (Department of Health, 1986).
60. Gardner, M. J. *et al.* Results of case–control study of leukaemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria. *Br. Med. J.* **300**, 423–429 (1990).
61. Kinlen, L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish New Town with nuclear reprocessing sites in Britain. *Lancet* **ii**, 1323–1327 (1988).
First description of the 'population-mixing' hypothesis for childhood leukaemia.
62. Greaves, M. F. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* **2**, 120–125 (1988).
First description of the 'delayed-infection' hypothesis for childhood leukaemia.
63. Wills-Karp, M., Santeliz, J. & Karp, C. L. The germless theory of allergic disease: revisiting the hygiene hypothesis. *Nature Rev. Immunol.* **1**, 69–75 (2001).
64. Dunne, D. W. & Cooke, A. A worm's eye view of the immune system: consequences for evolution of human autoimmune disease. *Nature Rev. Immunol.* **5**, 420–426 (2005).
65. Strachan, D. P. Family size, infection and atopy: the first decade of the 'hygiene hypothesis'. *Thorax* **55**, S2–S10 (2000).
66. Yazdanbakhsh, M., Kremsner, P. G. & van Ree, R. Allergy, parasites and the hygiene hypothesis. *Science* **296**, 490–494 (2002).
67. Kolb, H. & Elliott, R. B. Increasing incidence of IDDM a consequence of improved hygiene? *Diabetologia* **37**, 729–731 (1994).
68. Alvord, E. C. Jr *et al.* The multiple causes of multiple sclerosis: The importance of age of infections in childhood. *J. Child Neurol.* **2**, 313–321 (1987).
69. Gutensohn, N. & Cole, P. Childhood social environment and Hodgkin's disease. *N. Engl. J. Med.* **304**, 135–140 (1981).
70. Backett, E. M. Social patterns of antibody to poliovirus. *Lancet* **i**, 779–783 (1957).
71. Little, J. *Epidemiology of Childhood Cancer* (IARC Scientific Publications, Lyon, 1999).
72. McNally, R. J. Q. & Eden, T. O. B. An infectious aetiology for childhood acute leukaemia: a review of the evidence. *Br. J. Haematol.* **127**, 243–263 (2004).
Most comprehensive, recent audit of epidemiological data on possible causal mechanisms involving infection in childhood leukaemia.
73. Heath, C. W. Jr & Hasterlik, R. J. Leukemia among children in a suburban community. *Am. J. Med.* **34**, 796–812 (1963).
First description of a time–space cluster for childhood leukaemia.
74. Steinmaus, C., Lu, M., Todd, R. L. & Smith, A. H. Probability estimates for the unique childhood leukemia cluster in Fallon, Nevada, and risks near other U. S. military aviation facilities. *Environ. Health Perspect.* **112**, 766–771 (2004).
75. Alexander, F. E. Space-time clustering of childhood acute lymphoblastic leukaemia: indirect evidence for a transmissible agents. *Br. J. Cancer* **65**, 589–592 (1992).
76. Petridou, E. *et al.* Space–time clustering of childhood leukaemia in Greece: evidence supporting a viral aetiology. *Br. J. Cancer* **73**, 1278–1283 (1996).
77. Kinlen, L. J. Epidemiological evidence for an infective basis in childhood leukaemia. *Br. J. Cancer* **71**, 1–5 (1995).
78. Kinlen, L. J., Clarke, K. & Hudson, C. Evidence from population mixing in British New Towns 1946–85 on an infective basis for childhood leukaemia. *Lancet* **336**, 577–582 (1990).
79. Kinlen, L. J. & John, S. M. Wartime evacuation of children and mortality from childhood leukaemia in England and Wales in 1945–49. *Br. Med. J.* **309**, 1197–1202 (1994).
80. Kinlen, L. J. & Hudson, C. Childhood leukaemia and polio-myelitis in relation to military encampments in England and Wales in the period of national military service, 1950–63. *Br. Med. J.* **303**, 1357–1362 (1991).
81. Kinlen, L. J. & Balkwill, A. Infective cause of childhood leukaemia and wartime population mixing in Orkney and Shetland, UK. *Lancet* **357**, 858 (2001).
82. Dickinson, H. O. & Parker, L. Quantifying the effect of population mixing on childhood leukaemia risk: the Seascale cluster. *Br. J. Cancer* **81**, 144–151 (1999).
83. Alexander, F. E. *et al.* Clustering of childhood leukaemia in Hong Kong: association with the childhood peak and common acute lymphoblastic leukaemia and with population mixing. *Br. J. Cancer* **75**, 457–463 (1997).
84. Kinlen, L. & Doll, R. Population mixing and childhood leukaemia: Fallon and other US clusters. *Br. J. Cancer* **91**, 1–3 (2004).
85. Smith, M. Considerations on a possible viral etiology for B-precursor acute lymphoblastic leukemia of childhood. *J. Immunother.* **20**, 89–100 (1997).
86. Istre, G. R., Conner, J. S., Broome, C. V., Hightower, A. & Hopkins, R. S. Risk factors for primary invasive Haemophilus influenzae disease: increased risk from day care attendance and school-aged household members. *J. Pediatr.* **106**, 190–195 (1985).
87. Adler, S. P. Molecular epidemiology of cytomegalovirus: viral transmission among children attending a day care center, their parents, and caretakers. *J. Pediatr.* **112**, 366–372 (1988).
88. Fleming, D. W., Cochi, S. L., Hightower, A. W. & Broome, C. V. Childhood upper respiratory tract infections: to what degree is incidence affected by day-care attendance? *Pediatr.* **79**, 55–60 (1987).
89. Goodman, R. A., Osterholm, M. T., Granoff, D. M. & Pickering, L. K. Infectious diseases and child day care. *Pediatr.* **74**, 134–139 (1984).
90. Krämer, U., Heinrich, J., Wjst, M. & Wichmann, H.-E. Age of entry to day nursery and allergy in later childhood. *Lancet* **352**, 450–454 (1998).
91. McKinney, P. A. *et al.* Early social mixing and childhood Type 1 diabetes mellitus: a case–control study in Yorkshire, UK. *Diabet. Med.* **17**, 236–242 (2000).
92. Gilham, C. *et al.* Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from a UK case–control study. *Br. Med. J.* **330**, 1294–1297 (2005).
Largest study to date providing evidence that social contacts in infancy can reduce risk of childhood ALL.
93. Ma, X. *et al.* Daycare attendance and risk of childhood acute lymphoblastic leukaemia. *Br. J. Cancer* **86**, 1419–1424 (2002).
Provides evidence that increasing levels of social contact in early life (through day-care) proportionally reduce the risk of childhood ALL.

94. Ma, X. *et al.* Ethnic difference in daycare attendance, early infections, and risk of childhood acute lymphoblastic leukemia. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1928–1934 (2005).
95. Krynska, B. *et al.* Detection of human neurotropic JC virus DNA sequence and expression of the viral oncogenic protein in pediatric medulloblastomas. *Proc. Natl Acad. Sci. USA* **96**, 11519–11524 (1999).
96. McNally, R. J. Q. *et al.* An infectious aetiology for childhood brain tumours? Evidence from space–time clustering and seasonality analyses. *Br. J. Cancer* **86**, 1070–1077 (2002).
97. Petridou, E. *et al.* Age of exposure to infections and risk of childhood leukaemia. *Br. Med. J.* **307**, 774 (1993).
98. Infante-Rivard, C., Fortier, I. & Olsen, E. Markers of infection, breast-feeding and childhood acute lymphoblastic leukaemia. *Br. J. Cancer* **83**, 1559–1564 (2000).
99. Perrillat, F. *et al.* Day-care, early common infections and childhood acute leukaemia: a multicentre French case–control study. *Br. J. Cancer* **86**, 1064–1069 (2002).
100. Jourdan-Da Silva, N. *et al.* Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. *Br. J. Cancer* **90**, 139–145 (2004).
101. Neglia, J. P. *et al.* Patterns of infection and day care utilization and risk factors of childhood acute lymphoblastic leukemia. *Br. J. Cancer* **82**, 234–240 (2000).
102. Dockerty, J. D., Draper, G., Vincent, T., Rowan, S. D. & Bunch, K. J. Case–control study of parental age, parity and socioeconomic level in relation to childhood cancers. *Int. J. Epidemiol.* **30**, 1428–1437 (2001).
103. Ma, X. *et al.* Vaccination history and risk of childhood leukaemia. *Int. J. Epidemiol.* **34**, 1100–1109 (2005).
104. Groves, F. D., Sinha, D., Kayhty, H., Goedert, J. J. & Levine, P. H. Haemophilus influenzae type b serology in childhood leukaemia: a case–control study. *Br. J. Cancer* **85**, 337–340 (2001).
105. Auvinen, A., Hakulinen, T. & Groves, F. Haemophilus influenzae type B vaccination and risk of childhood leukaemia in a vaccine trial in Finland. *Br. J. Cancer* **83**, 956–958 (2000).
106. Stene, L. C. & Nafstad, P. Relation between occurrence of type 1 diabetes and asthma. *Lancet* **357**, 607–608 (2001).
107. Feltbower, R. G., McKinney, P. A., Greaves, M. F., Parslow, R. C. & Bodansky, H. J. International parallels in leukaemia and diabetes epidemiology. *Arch. Dis. Childhood* **89**, 54–56 (2004).
108. Schüz, J., Morgan, G., Böhler, E., Kaatsch, P. & Michaelis, J. Atopic disease and childhood acute lymphoblastic leukemia. *Int. J. Cancer* **105**, 255–260 (2003).
109. Wen, W. *et al.* Allergic disorders and the risk of childhood acute lymphoblastic leukemia (United States). *Cancer Causes Control* **11**, 303–307 (2000).
110. MacKenzie, J. *et al.* Infectious agents and childhood acute lymphoblastic leukemia: no evidence of molecular footprints of a transforming virus. *Haematologica* (in the press).
111. Balkwill, F., Charles, K. A. & Mantovani, A. Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell* **7**, 211–217 (2005).
112. Ford, A. M., Cardus, P. & Greaves, M. F. Modelling molecular consequences of leukaemia initiation by TEL–AML1 fusion. *Blood* **104**, 566 (2004).
113. Hiebert, S. W. *et al.* The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. *Mol. Cell. Biol.* **16**, 1349–1355 (1996).
114. Guidez, F. *et al.* Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL–AML1 oncoprotein. *Blood* **96**, 2557–2561 (2000).
115. Cookson, W. The alliance of genes and environment in asthma and allergy. *Nature* **402** (Suppl.), B5–B11 (1999).
116. Trinchieri, G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nature Rev. Immunol.* **3**, 133–146 (2003).
117. Morahan, G. *et al.* Association of *IL12B* promoter polymorphism with severity of atopic and non-atopic asthma in children. *Lancet* **360**, 455–459 (2002).
118. Morahan, G. *et al.* Linkage disequilibrium of a type 1 diabetes susceptibility locus with a regulatory *IL12B* allele. *Nature Genet.* **27**, 218–221 (2001).
119. Le Souëf, P. N., Goldblatt, J. & Lynch, N. R. Evolutionary adaptation of inflammatory immune responses in human beings. *Lancet* **356**, 242–244 (2000).
120. Greaves, M. F. Evolution, immune response, and cancer. *Lancet* **356**, 1034 (2000).
121. Taylor, G. M. *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.* **11**, 1585–1597 (2002).
122. Dorak, M. T. *et al.* Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. *Blood* **94**, 694–700 (1999).
123. Sturniolo, T. *et al.* Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nature Biotechnol.* **17**, 555–561 (1999).
124. Breatnach, F., Chessells, J. M. & Greaves, M. F. The aplastic presentation of childhood leukaemia: a feature of common-ALL. *Br. J. Haematol.* **49**, 387–393 (1981).
125. Hasle, H. *et al.* Transient pancytopenia preceding acute lymphoblastic leukemia (pre-ALL). *Leukemia* **9**, 605–608 (1995).
126. Heegaard, E. D., Madsen, H. O. & Schmiegelow, K. Transient pancytopenia preceding acute lymphoblastic leukaemia (pre-ALL) precipitated by parvovirus B19. *Br. J. Haematol.* **114**, 810–813 (2001).
127. Liang, R., Cheng, G., Wat, M. S., Ha, S. Y. & Chan, L. C. Childhood acute lymphoblastic leukaemia presenting with relapsing hypoplastic anaemia: progression of the same abnormal clone. *Br. J. Haematol.* **83**, 340–342 (1993).
128. Ishikawa, K. *et al.* Detection of neoplastic clone in the hypoplastic and recovery phases preceding acute lymphoblastic leukemia by *in vitro* amplification of rearranged T-cell receptor δ chain gene. *J. Pediatr. Hematol. Oncol.* **17**, 270–275 (1995).
129. *Infectious Causes of Cancer* (ed. Goedert, J. J.) (Humana Press, New Jersey, 2000).
130. Isaacson, P. G. & Du, M.-Q. MALT lymphoma: from morphology to molecules. *Nature Rev. Cancer* **4**, 644–653 (2004).
131. Wotherspoon, A. C. *et al.* Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of *Helicobacter pylori*. *Lancet* **342**, 575–577 (1993).
132. Hitzler, J. K. & Zipursky, A. Origins of leukaemia in children with Down syndrome. *Nature Rev. Cancer* **5**, 11–20 (2005).
133. Dockerty, J. D. *et al.* Infections, vaccinations and the risk of childhood leukaemia. *Br. J. Cancer* **80**, 1483–1489 (1999).
134. Chan, L. C. *et al.* Is the timing of exposure to infection a major determinant of acute lymphoblastic leukaemia in Hong Kong? *Paediatr. Perinatal Epidemiol.* **16**, 154–165 (2002).
135. MacKenzie, J., Perry, J., Ford, A. M., Jarrett, R. F. & Greaves, M. JC and BK virus sequences are not detectable in leukaemic samples from children with common acute lymphoblastic leukaemia. *Br. J. Cancer* **81**, 898–899 (1999).
136. Smith, M. A. *et al.* Investigation of leukemia cells from children with common acute lymphoblastic leukemia for genomic sequences of the primate polyomaviruses JC virus, BK virus, and simian virus 40. *Med. Pediatr. Oncol.* **33**, 441–443 (1999).
137. Priftakis, P. *et al.* Human polyomavirus DNA is not detected in Guthrie cards (dried blood spots) from children who developed acute lymphoblastic leukemia. *Med. Pediatr. Oncol.* **40**, 219–223 (2003).
138. Isa, A., Priftakis, P., Broliden, K. & Gustafsson, B. Human parvovirus B19 DNA is not detected in Guthrie cards from children who have developed acute lymphoblastic leukemia. *Pediatr. Blood Cancer* **42**, 357–360 (2004).
139. Bogdanovic, G., Jernberg, A. G., Priftakis, P., Grillner, L. & Gustafsson, B. Human herpes virus 6 or Epstein–Barr virus were not detected in Guthrie cards from children who later developed leukaemia. *Br. J. Cancer* **91**, 913–915 (2004).
140. MacKenzie, J. *et al.* Screening for herpesvirus genomes in common acute lymphoblastic leukemia. *Leukemia* **15**, 415–421 (2001).
141. Bender, A. P. *et al.* No involvement of bovine leukemia virus in childhood acute lymphoblastic leukemia and non-Hodgkin's lymphoma. *Cancer Res.* **48**, 2919–2922 (1988).
142. Shiramizu, B., Yu, Q., Hu, N., Yanagihara, R. & Nerurkar, V. R. Investigation of TT virus in the etiology of pediatric acute lymphoblastic leukemia. *Pediatr. Hematol. Oncol.* **19**, 543–551 (2002).
143. Hoffjan, S. J. *et al.* Gene–environment interaction effects on the development of immune responses in the 1st year of life. *Am. J. Hum. Genet.* **76**, 696–704 (2005).

Acknowledgements

This review is dedicated to the memory of Professor Sir Richard Doll who chaired the UKCCS Management Committee and was a powerful advocate for epidemiological studies on childhood leukaemia. The author's research is supported by a specialist programme grant from the Leukaemia Research Fund and by The Kay Kendall Leukaemia Fund.

Competing interests statement

The author declares no competing financial interests.

DATABASES

The following terms in this article are linked online to:

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene

AML1 | *TEL*

National Cancer Institute: http://www.cancer.gov

acute lymphoblastic leukaemia | acute myeloid leukaemia | Hodgkin lymphoma | non-Hodgkin lymphoma

Access to this interactive links box is free online.