Cancer Progress and Priorities: Childhood Cancer
Philip J. Lupo1 and Logan G. Spector2

Background
It is estimated that 300,000 children 0–19 years of age are diagnosed with cancer worldwide each year (1). In high-income countries, cancer is the leading cause of death due to disease in children. The absolute risk of cancer in children is, however, quite low [183 cases per million in the United States (2)], and this rarity limits attainable sample sizes and types of studies. Childhood cancers are heterogeneous and display a markedly different range of tumor types than in adults, including several classes that are largely exclusive to children. Advances in diagnostics have further split tumors into molecularly-defined subtypes that inform prognosis, therapy, and increasingly etiology. Childhood cancer epidemiology has traditionally relied on interview-based case–control studies but in recent decades has added laboratory assessment of exposure, germline DNA analysis, and molecular classification of tumors to the research repertoire.

Descriptive Epidemiology
Worldwide incidence
Accurate estimates of worldwide childhood cancer incidence are important for characterizing the impact of these malignancies and informing policy decisions. However, many countries do not have cancer registries that quantify the incidence of childhood cancer. It is estimated that 300,000 children are diagnosed with cancer annually. Data from GLOBOCAN are commonly used as a primary source to estimate the global incidence of childhood cancer. Notably, the incidence of childhood cancer is highest in North America, parts of South and Central America, Europe, and Australia with an age-standardized incidence rate (ASR) of ≥15.4 per 100,000 person-years for those 0–19 years of age (Fig. 1A). These patterns largely hold true for leukemia diagnosed in those 0–19 years of age as well (Fig. 1B). However, there are limitations to using GLOBOCAN data. For instance, data are presented according to ICD site codes, which do not reflect the major childhood cancer diagnostic groups. It is also suspected that in middle- and low-income countries poor pathology, misdiagnosis, and unascertained cases contribute to underestimation of rates. Because of this, a recent report attempted to estimate the total incidence of global childhood cancer using a simulation-based approach. In this assessment, Ward and colleagues estimated that there were 397,000 children 0–14 years of age diagnosed with cancer worldwide in 2015, a number much higher than GLOBOCAN estimates (3).

International comparisons of rates between high-, middle-, and low-income countries using standard registry data should therefore be interpreted with the caveat that only diagnosed cancer is counted. Overall, leukemia is the most common cancer among children ages 0 to 14 years regardless of the geographic area. However, leukemias represent a slightly higher proportion of childhood cancers in Asia, Oceania, and Central and South America, while slightly lower on the African continent. In both North Africa and Sub-Saharan Africa, lymphomas are more common than in other regions, due primarily to the high rates of Burkitt lymphoma. Soft-tissue sarcomas are much more common in Sub-Saharan Africa, due to the high incidence of Kaposi sarcoma in the region. Other notable differences include a lower proportion of CNS tumors but higher proportion of renal tumors in Sub-Saharan Africa, and a higher proportion of germ cell tumors in Asia.

U.S. incidence
Approximately 16,000 children 0–19 years of age are diagnosed with cancer in the United States (11,000 cases among children 0–14 years of age and 5,000 cases among those 15–19 years of age; ref. 4). These numbers correspond to an ASR for all cancers of 16.4 cases per 100,000 person-years for 0–14 years and 23.3 per 100,000 person-years for 15–19 (2). Notably, the incidence of childhood cancer varies by year of life (Fig. 2). In addition, the distribution of cancer types shifts throughout childhood and adolescence. For example, non-CNS embryonal tumors are more common in early life compared with lymphomas, whereas lymphomas become relatively more common in adolescence (Fig. 2).

As with many adult cancers, the incidence of childhood cancer varies by race/ethnicity. However, while non-Hispanic black adults often have a higher incidence of several cancers, non-Hispanic white children often experience a higher incidence of cancer relative to non-Hispanic black and Hispanic children. One notable exception is Hispanic children have higher rates of both acute lymphoblastic lymphoma (ALL) and acute myeloid leukemia (AML) compared with non-Hispanic white and non-Hispanic black children (2). There is emerging evidence that some of these disparities may be due to underlying genetic ancestry (5). However, this information has not been fully exploited (e.g., through admixture mapping) to better understand the etiology of childhood cancer (6).

Worldwide survival
Worldwide, more than 100,000 children and adolescents younger than 20 years of age die from cancer per year (75,000 cancer deaths among children 0 to 14 years of age and 27,000 cancer-related deaths among 15- to 19-year-olds; ref. 7). Survival is generally higher in high-income countries (HIC). Specifically, survival has consistently increased in most of Europe, North America, Japan, and Oceania (8, 9). Whereas, several countries in Eastern Europe, Southeastern Asia, and Latin America have lagged behind. As with childhood cancer incidence, it is difficult to ascertain survival in countries without robust population-based cancer registries. International data presented in The Cancer Atlas (10), which is based on data from the CONCORD program, show that for roughly the same time period (1990s to early

1Department of Pediatrics, Section of Hematology-Oncology, Baylor College of Medicine, Houston, Texas. Department of Pediatrics, Division of Epidemiology & Clinical Research, University of Minnesota, Minneapolis, Minnesota.
Corresponding Author: Philip J. Lupo, Baylor College of Medicine, One Baylor Plaza MS: BCM622, Houston TX, 77030; Phone: 713-798-2960; Fax: 713-798-8711; E-mail: philip.lupo@bcm.edu
Cancer Epidemiol Biomarkers Prev 2020;29:1081-94
doi: 10.1158/1055-9965.EPI-19-0941
©2020 American Association for Cancer Research.
2000s), 5-year survival for childhood cancer overall was approximately 80% in high-income countries, roughly 55% in middle-income countries, and 40% in low-income countries (LIC). Furthermore, survival also differed by cancer type across those countries. Leukemia and lymphoma experience among the highest five-year survival in HICs (80% and 90%, respectively); however, in LICs only, 36% and 55% of children diagnosed with leukemia and lymphoma, respectively, survive five years after their diagnosis. The disparity is even greater for CNS tumors and neuroblastoma (which were considered together as a group in the Cancer Atlas Data). Five-year survival is reported at 71% in HICs and only 27% in LICs. These trends were also demonstrated in an assessment that used a simulation-based analysis to estimate global childhood cancer survival trends. Specifically, Ward and colleagues reported that global 5-year net childhood cancer survival is currently 37.4% (95% uncertainty interval, 34.7–39.8), with large variation by region, ranging from 8.1% (4.4–13.7) in eastern Africa to 83.0% (81.6–84.4) in North America (3).

U.S. survival
Childhood cancer remains the leading cause of disease-related mortality among children 1 to 14 years of age, with approximately 1,200 cancer-related deaths annually in the United States among
children younger than 15 years (4). The relative contribution of cancer to overall mortality for 15- to 19-year olds is lower than for the younger children, although approximately 600 deaths from cancer occur annually in this age group (11). Accordingly, death from cancer accounts for 12% of all deaths among children 1 to 14 years old, and 5% among adolescents (15–19 years old). However, survival rates for children 0–14 years of age have improved dramatically since the 1960s when the overall 5-year survival rate after a cancer diagnosis was estimated as 28% (12). Improvements in survival rates have continued into the mid-2000s in the United States, with the overall 5-year survival rate exceeding 80% for children and adolescents diagnosed during this period. In spite of this, survival does still lag for some cancer types (Fig. 3). For instance, as recent data from Surveillance, Epidemiology, and End Results (SEER) demonstrate, children diagnosed with CNS tumors, bone tumors, some types of soft tissue sarcoma, and hepatoblastoma have 5-year survival rates of <70%. While overall survival has improved in the United States, there remain differences based on several factors, including sex and race/ethnicity. In terms of sex, using data from SEER, investigators demonstrated that males had worse survival compared with females for several cancers, including ALL, ependymoma, neuroblastoma, and osteosarcoma (13). In addition, while non-Hispanic white children are more likely to be diagnosed with cancer, these children often have the best survival compared with other race/ethnicity groups. Specifically, data from SEER indicate that for most subtypes, non-Hispanic blacks and Hispanics have inferior survival compared with non-Hispanic whites (14–19). Differences in survival are likely to be complex, arising from several factors, including but not limited to socioeconomic status, adherence to therapy, differences in treatment, underlying tumor biology, and genetic susceptibility (20). Therefore, future studies evaluating disparities should employ a comprehensive approach to reducing the impact of these differences.

### Risk Factors

Knowledge on the etiology of childhood cancers comes mainly from case–control studies and must be interpreted within the limitations of the study design. Setting must also be considered as the vast majority of analytic epidemiology of childhood cancer have taken place in high income countries and may not generalize to lower income areas with different distributions of environmental and lifestyle risk factors. Childhood cancers are heterogeneous, and it must be recognized that each tumor has its individual risk factor profile. However there are some commonalities. For instance, the incidence is higher in males in nearly all cancers and many are thought to originate in utero. Some risk factors, like birth weight, cut across tumors but differ in the direction and magnitude of association, while others [e.g., hernias and Ewing sarcoma (21)] are unique to a particular cancer. It is also the case that the literature on causes is roughly proportional to incidence, so that we know far more about the etiology of ALL (36 cases per million) than hepatoblastoma (2.1 cases per million). On the basis of a review of the literature and overall impressions from the state of the field, we summarize the risk factors for childhood cancer below and in Table 1. In addition, as there have been some reviews specifically focused on particular childhood cancers or the role of specific risk factors on a range of childhood cancers, we have also provided a table (Table 2) that outlines some of these efforts.

### Demographics

Age, sex, and race/ethnicity each influence the risk of childhood cancer. Figure 2 shows the distribution of tumors across the pediatric age spectrum. Leukemias, primarily ALL, display a distinct peak in incidence from age 2 to 5 years, while CNS tumors have fairly steady incidence throughout childhood. Non-CNS embryonal tumors have

---

**Figure 2.**
Distribution of tumors across the pediatric age spectrum.
peak incidence in infancy, which then falls to near zero by 10 years of age. There is a peak in incidence of bone and soft-tissue sarcomas in mid-adolescence, as well as increasing incidence of lymphoma and other solid tumors that continues well into adulthood. Males have a higher risk for most childhood cancers, with rate ratios ranging from 6 for Burkitt lymphoma to 1.2 for AML; Wilms tumor, extragonadal GCT, thyroid carcinoma, and melanoma are notable for having slightly higher rates in females (22). The United States published cancer rates by race/ethnicity and so these data are often used to compare incidence. With few exceptions, the rate of childhood cancer is 25%–50% lower in black compared with white children, the rate of ALL is higher among Hispanic children but for most solid tumors lower than in white children, and incidence among Asian children is roughly similar to white children although some (i.e., neuroblastoma and Wilms tumor) have lower rates while germ cell tumors have dramatically higher rates (23).

### Gestational and perinatal

Because of the young age of onset, histologic resemblance of some pediatric cancer cells to embryonal cells, and pre- or perinatal detection of multiple types of cancer, most childhood cancers are thought to initiate in utero. Consequently, gestation is frequently examined as a critical window of risk. Birth weight, alone or with consideration of gestational age, has been most often studied (Table 1). Risk of most childhood cancers rises with increasing birth weight, from 5% per 500 g for astrocytomas to 17% per 500 g for Wilms tumor (24); hepatoblastoma, which is strongly associated with low birth weight, is the exception (25). Another consistent finding is that children with structural birth defects without a reported genetic syndrome have an elevated risk of pediatric cancer (26). Risk ratios for specific birth defect–cancer combinations range from 1.5 to 6.0 and the associations are apparent for most classes of birth defects and cancer (Table 1).

Parental age at birth of offspring is also consistently associated with most types of childhood cancer, with between 5% and 10% higher risk per 5 years of maternal age; paternal age may also be associated with higher risk of pediatric cancer, although the tight correlation with maternal age and the greater degree of missing data make it harder to assess (27–31). There is some suggestive evidence that pathologies of pregnancy, such as preeclampsia, gestational diabetes, and maternal obesity increase risk of childhood cancers but the evidence base is currently thin (32–37).

### Environmental

While key environmental exposures have been identified for adult cancers (e.g., smoking, benzene), much less is known in relation to childhood cancer. A notable difference between adult cancers and childhood cancers is the latency period associated with these conditions. For instance, smoking usually starts during adolescence or young adulthood, but associated malignancies do not become apparent until many decades later. However, several childhood cancers are predominantly diagnosed in infancy (e.g., embryonal neoplasms such as neuroblastoma) or in early childhood (2–5 years of age), such as ALL. Therefore, the disruption in molecular processes that may lead to childhood cancer are likely different from those of adult cancers; at the least, the carcinogenic process in children is necessarily much shorter in time. In addition, because of the age of onset, it is reasonable to surmise that many childhood cancers result from aberrations in early developmental processes. The current evidence to support a major etiologic role for environmental or other exogenous factors in childhood cancer is minimal (Table 1). While there is extensive evidence that high doses of ionizing radiation are associated with childhood cancer, the prevalence of this exposure is very low (38). Relatively common environmental exposures, including pesticides (39) and air

![Figure 3. Improvements in 5-year survival for childhood cancer, 1975-2010.](attachment:image-url)
pollution (40), have also been explored. While individual studies and meta-analyses have indicated associations between these exposures and some childhood cancers, effect sizes are relatively modest. One problem with previous studies of environmental exposures includes the limitations of using questionnaire data to estimate exposure or the use of proxies for exposure assessment (e.g., residential information). While residential information is arguably better than self-reported exposures, epidemiologic studies of childhood cancer must leverage novel approaches to better characterize the role of environment in etiology. For example, correlating information on environmental exposures to the landscape of somatic mutations could provide novel insights into the carcinogenic properties of these factors (41).

Genetic variation

Chromosomal abnormalities, subchromosomal structural variation, and pathogenic germline point mutations confer a sharply increased risk of cancer but underlie a minority of cases of childhood cancer (42–45). Next-generation sequencing studies have yielded precise estimates of the prevalence of germline pathogenic variation in most types of childhood cancer; for a few types such as adrenocortical carcinoma or hypodiploid ALL, half or more of cases have such variation, while in most cancers the prevalence is 5%–10% (46, 47). In fact, according to a large-scale effort by Zhang and colleagues, those childhood cancers where >10% can be attributed to highly penetrant pathogenic variants in known cancer predisposition genes include osteosarcoma, retinoblastoma, and adrenocortical carcinoma (47). Continued sequencing efforts have also yielded novel, rare, high-penetrance predisposition genes (48) and moderately rare, medium-penetrance variation (49–51). Importantly, there has been a recent systematic effort to outline the guidelines for pediatric cancer predisposition surveillance. While it is beyond the scope of this review to fully outline those recommendations, they can be found at https://clincancerres.aacrjournals.org/pediatricseres (52).

Table 1. Confirmed and suspected risk factors for selected childhood cancers.

<table>
<thead>
<tr>
<th>Topic</th>
<th>ALL</th>
<th>AML</th>
<th>NB</th>
<th>HB</th>
<th>RB</th>
<th>WT</th>
<th>MB</th>
<th>PNET</th>
<th>Epidermoidoma</th>
<th>Astrocytoma</th>
<th>Strength of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preconception/pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Vitamins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Occupational exposures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Residential exposures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Coffee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Ionizing radiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Birth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Maternal age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Paternal age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Chromosomal birth defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Non-chromosomal birth defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>High birth weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Low birth weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>C-section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Gestational age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Childhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Breastfeeding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Allergies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Residential chemical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Passive smoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Irradiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
</tbody>
</table>

Note: Taken from refs. 25, 30, 31,103-266. For strength of evidence: + epidemiologic evidence with little mechanistic support; ++ can cross placenta or has developmental consequences but epidemiologic evidence is equivocal; +++ strong epidemiologic and mechanistic evidence.

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; HB, hepatoblastoma; MB, medulloblastoma; NB, neuroblastoma; PNET, primitive neuroectodermal tumor; RB, retinoblastoma; WT, Wilms tumor.

Legend

Positive, effect estimate <1.5
Positive, effect estimate ≥1.5
No association
Negative, effect estimate >0.67
Negative, effect estimate 0.67
Inconclusive

Table 2. Selected reviews for childhood cancer and associated risk factors.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Authors</th>
<th>Year</th>
<th>PMID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood cancers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>Williams et al.</td>
<td>2019</td>
<td>30770347</td>
</tr>
<tr>
<td>AML</td>
<td>Puumala et al.</td>
<td>2013</td>
<td>23303597</td>
</tr>
<tr>
<td>Brain tumors</td>
<td>Johnson et al.</td>
<td>2014</td>
<td>25192704</td>
</tr>
<tr>
<td>Hepatoblastoma</td>
<td>Spector and Birch</td>
<td>2012</td>
<td>22692949</td>
</tr>
<tr>
<td>Rhabdomyosarcoma</td>
<td>Skapek et al.</td>
<td>2019</td>
<td>30617281</td>
</tr>
<tr>
<td>Wilms tumor</td>
<td>Chu et al.</td>
<td>2010</td>
<td>20670226</td>
</tr>
<tr>
<td>Risk factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pesticides</td>
<td>Chen et al.</td>
<td>2015</td>
<td>26371195</td>
</tr>
<tr>
<td>Ionizing radiation</td>
<td>Kendall et al.</td>
<td>2018</td>
<td>30151551</td>
</tr>
<tr>
<td>Genetic predisposition</td>
<td>Pilon and Lupo</td>
<td>2019</td>
<td>30822280</td>
</tr>
</tbody>
</table>
greater than in adults (23), which seems to be a general property of pediatric diseases (33). The genetic architectures of ALL and neuroblastoma are mature, with multiple validated loci, subtype-specific associations, transethnic replication, and ethnically-specific loci (49–51, 54–59). Two GWAS of Ewing sarcoma in European populations have also identified multiple loci (60, 61), but genetic risk in non-European populations has not been examined. Single GWAS of Wilms tumor (62) and osteosarcoma (63) have also identified a limited number of loci. More recently investigators have demonstrated associations of childhood cancers with trait-related variation, such as genetically-determined telomere length (64) and height (65). The most significant variants found to be associated with childhood cancers in GWAS are depicted in Fig. 4. While potential mechanisms for the genetic associations have been proposed (57–63), there have been few efforts to fully elucidate the functional consequences of variants identified in GWAS of childhood cancers. This is in part due to several variants being intronic or intergenic (63). However, there are some emerging efforts to elucidate the role of variants and genes identified in GWAS of childhood cancer, including Ewing sarcoma–related loci (66), IKZF1 variants and ALL (67), and BMI1 variants and ALL (68). Many SNPs do appear to be associated with lymphocyte development; however, additional work is needed to explore the biological underpinnings of these associations. Lastly, while small studies have examined transethnic replication of GWAS SNPs discovered in Europeans, genome-wide discovery has mostly not been performed in non-European populations despite many recent calls to diversify genomic research.

### Etiologic Heterogeneity

There is emerging evidence that etiologic heterogeneity within childhood cancer subtypes may have limited previous epidemiologic assessments of these conditions. Furthermore, our understanding of subtypes continues to emerge. For example, ALL has traditionally been classified as B-cell or T-cell, based on the cell type affected. However, advances in cytogenetics have led to the latest version of the WHO classification of ALL to include several subtypes defined by their translocations and other cytogenetic features: BCR-ABL1, MLL rearranged, TEL-AML1, hyperdiploidy, hypodiploidy, IL3-IGH, and E2A-PBX1 (69). Another example is rhabdomyosarcoma, which was originally classified by histologic type, for example, embryonal versus alveolar. However, through molecular advancements, further distinctions due to specific gene fusions between either PAX3 or PAX7 and FOXO1 that typically occur among the previously named alveolar types, are preferred risk stratification strategies compared with histology alone (70). Recent advances in genomics, epigenomics, and transcriptomics have allowed for molecular subtyping for a number of childhood malignancies. For example, at least four molecular subgroups of childhood medulloblastoma now exist (WNT, SHH, Group 3, and Group 4), each exhibiting different molecular and clinical features (71); more recent tumor phenotyping suggests even further subtypes (72). Likewise, for other brain tumors, distinct molecular subgroups have now been established for ependymoma (73), high-grade gliomas (74), low-grade gliomas (75), and AT/RT (76). Recently, four molecular subtypes have been suggested for diffuse large B-cell lymphoma: MCD (harboring the cooccurrence of MYD88L265P and CD279B mutations), BN2 (harboring BCL6 fusions and NOTCH2 mutations), N1 (harboring NOTCH1 mutations), and EZB (harboring EZH2 mutations and BCL2 translocations; ref. 77). Less often, molecular analyses have suggested “lumping” tumors previously thought to be dissimilar, as with Ewing sarcoma and primitive neuroectodermal tumors (PNET), which both frequently feature the EWS-FLI1 translocation (78). Also, GWAS of childhood cancers (including ALL and neuroblastoma) have pointed to differences in association between SNPs and molecularly defined subtypes (79, 80). These are just a few examples of the emerging landscape of tumor subtypes based on molecular features. Future epidemiologic studies must account for this information as etiologic factors could differ based on these characteristics.

### Screening and Prevention

The ultimate goal of etiologic research in childhood cancer is to enable risk prediction, early detection, and, eventually, prevention.
However, these goals remain distant for most children. Population-wide screening for pediatric cancer has, to our knowledge, only been attempted for neuroblastoma. The basis for screening was homovanillic acid, a catecholamine metabolite which serves as a biomarker for tumor burden in neuroblastoma and which is shed in urine. Thus, in the late 1980s and 1990s, screening for neuroblastoma in infants was attempted in three areas with the capacity for population-wide urine collection in infants: Quebec (81), Germany (82), and Japan (83). While these programs succeeded in identifying neuroblastoma earlier than clinical diagnosis, they did not improve mortality as they primarily detected favorable cases with a regressing phenotype, many of which would never have come to clinical attention. Hence, these programs were abandoned (84).

More recently, experts have issued recommendations in support of surveillance for tumor development in children with most genetic syndromes conferring high risk of cancer (85). Although there are few preventive measures to implement, there is consensus that surveillance can reduce morbidity and mortality through early detection. Screening for cancer predisposition, as opposed to screening for cancer in those with known predispositions, is more controversial. Newborn screening increasingly involves genetic in addition to metabolic testing (86), and thus could easily detect most types of pathogenic variation in cancer-associated genes. However, many screening programs prefer to include only conditions that are early-onset and for which there are interventions proven to improve outcome (87), which does not describe most pediatric cancers. Thus, to our knowledge only one area in the world has instituted newborn screening for pediatric cancer predisposition, in the Brazilian state of Paraná where the R337H founder mutation in TP53 has an especially high prevalence (88).

Prevention of pediatric cancer is not yet feasible for a number of reasons. The first is simply that for diseases as rare as these the number needed to “treat” with an intervention would be impractically large, possibly population-wide, and consequently would be economically unfavorable. A second reason is that, as discussed above, there are no modifiable risk factors for childhood cancer that are strong and prevalent enough to justify intervention. However, most of the modifiable risk factors for pediatric cancer (e.g., maternal smoking, obesity, air pollution) are also associated with far more common diseases, thus efforts to reduce exposure for other reasons may have the effect of reducing childhood cancer incidence.

**Future Directions**

**Molecular epidemiology**

Initial epidemiologic studies of childhood cancer gathered data mainly by parental interview and medical record abstraction. These assessments relied on the case–control study design. However, the focus of the past decade has largely been on the molecular epidemiology of childhood cancers. This has been facilitated in part through the case–parent trio study design. Case–parent trios allow the estimation of inherited genetic effects, maternal genetic effects (which can be used as a proxy or mediator of the intrauterine environment), and gene–environment interactions. In addition, the case–parent trio approach does not require the inclusion of a control group, which is a practical advantage, as control selection on the national scale has become increasingly difficult in recent years. This is, in part, due to the reliance on random digit dialing for control selection (89). Another option for control selection is utilizing birth certificate controls, which have been leveraged for two Children’s Oncology Group (COG) studies (90, 91), as well as studies of other pediatric and perinatal outcomes (92). This could be a feasible approach for epidemiologic studies of childhood cancer that require a comparison group. However, the scientific and practical appeal of the case–parent trio design for molecular epidemiology studies remains compelling. Because of this, several recent COG epidemiology studies have relied on this approach. This includes studies of osteosarcoma (93), neuroblastoma (94), Wilms tumor, Ewing sarcoma, germ cell tumors (95), and rhabdomyosarcoma. However, beyond genetic susceptibility to childhood cancer, few studies have explored using biological markers of exposure in studies of childhood cancer, which is in part due to the limited availability of samples collected prior to diagnosis. An emerging and important population-based resource for molecular epidemiology of childhood cancer is the use of dried blood spots (DBS) collected and archived as part of newborn screening efforts. DBS have been used in genetic epidemiology studies of childhood cancer (67), as well as using metabolomics to reveal novel ALL phenotypes (96). In addition, DBS can be used to estimate prenatal exposures, including cotinine from tobacco smoke (97) and benzene (98).

There are also methods for leveraging genetic data to address questions not necessarily related to inherited genetic susceptibility. Two primary examples are (i) evaluating maternal genetic effects (as described earlier), which can be done using parental genetic data from case–parent trios (99), and (ii) Mendelian randomization (100), which is a method of using genetic variation to examine the effect of an exposure (or another trait like birth weight) on disease in observational studies. These methods are more recently being leveraged in epidemiologic studies of childhood cancer. For instance, there has been an exome-wide association study of maternal genetic effects on ALL (101). In addition, Mendelian randomization has been used to characterize the role of height on osteosarcoma risk (65) and telomere length on neuroblastoma risk (64).

A final underexplored area in the molecular epidemiology of childhood cancer is leveraging epigenetics, especially as these modifications relate to germline DNA. Notably, environmental exposures can lead to epigenetic modifications that influence gene expression and can modulate disease risk associated with genetic variation (101). For example, there is emerging evidence that air pollution exposure (a suspected risk factor for several childhood cancers) is associated with changes in DNA methylation (102). Therefore, a novel approach in better ascertaining the association between air pollution and particular childhood cancers could be evaluating DNA methylation marks associated with this exposure. Epigenetics therefore holds substantial promise for identifying mechanisms through which genetic and environmental factors jointly contribute to childhood cancer risk and outcome. As the underlying etiologies of the vast majority of childhood cancers appear multifactorial, including both genetic and environmental risk factors, molecular epidemiology will continue to be an important component in the assessment of these conditions.

**Tumor biology**

As discussed, there is a growing awareness of the molecular heterogeneity within childhood cancer subtypes. As this molecular heterogeneity could point to etiologic heterogeneity, it will be vital to incorporate information on somatic mutations in future epidemiologic studies of childhood cancer. Furthermore, information on somatic mutations could be leveraged to better understand biological processes underlying etiology. For example, in an assessment by Alexandrov and colleagues of over 7,000 tumors yielded more than...
20 distinct mutational signatures, which were associated with various features including age, mutagenic exposures, and defects in DNA maintenance (41). In addition, several of these mutational signatures are of “cryptic” origin. Epidemiologic assessments that characterize the exposures associated with these signatures could yield novel insights into the mutational processes underlying the development of cancer with potential implications for prevention and therapy.

Global epidemiology

It should be noted that the overwhelming majority of etiologic studies of ALL have been conducted in high-income countries, especially the United States and countries in Europe. It is critical that future studies include populations in middle- and low-income countries as exposures as well as genetic variation are likely to differ in these populations and etiologic features may differ.

Conclusions

While there have been tremendous strides in improving outcomes for children with cancer, there is still a great deal of work related to disentangling the etiologic origins of these conditions. Future studies should incorporate novel exposure methodologies, molecular features of tumors, and a more complete assessment of gene–environment interactions. Through these efforts, it is hoped that our understanding of the causes of childhood cancer can be better ascertained, leading to novel surveillance or prevention strategies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: P.J. Lupo, L.G. Spector
Development of methodology: P.J. Lupo, L.G. Spector
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.J. Lupo, L.G. Spector
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P.J. Lupo, L.G. Spector
Writing, review, and/or revision of the manuscript: P.J. Lupo, L.G. Spector
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P.J. Lupo, L.G. Spector
Study supervision: P.J. Lupo

Acknowledgments

This work was funded, in part, through the National Cancer Institute (U19CA180886), Cancer Prevention and Research Institute of Texas (CPRIT RP170071 and RP180755), and St. Baldrick’s Foundation (522277). We would like to acknowledge the work of the Children’s Oncology Group Epidemiology Committee in supporting this work.

Received August 7, 2019; revised December 18, 2019; accepted March 9, 2020, published first June 1, 2020.

References


221. Parodi S, Merlo DF, Ranucci A, Miligi L, Benvenuti A, Rondelli R, et al. Risk of
220. Pang D, McNally R, Birch JM. Parental smoking and childhood cancer:
215. Omidakhsh N, Bunin GR, Ganguly A, Ritz B, Kennedy N, Von Ehrenstein OS,
211. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Kheifets L. Birth weight and
210. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Kheifets L. Birth weight and
214. Olshan AF, Smith JC, Bondy ML, Neglia JP, Pollock BH. Maternal vitamin use
2003;88:373
Childhood Leukemia International Consortium. Cancer Epidemiol 2019;59:
Parental age and the risk of childhood acute myeloid leukemia: results from the
78.
147.
111.
94.
43.
42.
41.
32.
31.
30.
11.
5.
2.
1.