A first-born twin has a higher risk of acute leukemia in a population-based assessment of cancer in twins in California, and a lower than anticipated rate of twin concordance

Investigation of twins has enabled estimations of the germline and somatic genetic contribution to many pediatric cancers.[1] Twin concordance for childhood acute leukemias is estimated at approximately 10%[2] for lymphoblastic leukemia (ALL) and myeloid leukemia (AML) phenotypes. While purpose-built twin cohorts allow for large enrollments of twin cases with leukemia, these studies are at risk of inaccuracies resulting from ascertainment bias.[3] In contrast, records-based population-based studies allow accrual in an unbiased fashion. Here, we identified substantially lower than expected pediatric and adolescent/young adult (AYA) twin leukemia concordance rates through a population-based registry assessment in California from 1982-2022. We further demonstrated a unique association showing first-born twins are more likely to develop ALL than second-born. We show the difference in risk of ALL development by birth order may be reflected in epigenome-wide alterations in DNA methylation, in which first-born twins show significant hypomethylation compared to second-born siblings. Furthermore, we show significant alterations in DNA methylation at sites with known association to birth order and cortisol responsiveness, implying a potential biological explanation for this finding.

This research was performed under review by a California State Institutional Review Board which permits health research on de-identified Biobank specimens. We merged registry data from the California Birth Master Statistical File (BMSF) from 1982 to 2017 and the California Cancer Registry (CCR) from 1988 to 2022 to identify a total of 1,213 twins in which at least one sibling was diagnosed with leukemia from 0-39 years of age. Of these, there were 255 twin pairs in which at least one sibling in the pair had a diagnosis of leukemia, along with 265 distinct cancer diagnoses and 258 leukemia diagnoses within these 255 twin pairs (Table 1). Concordant cases were defined as both siblings in a twin pair having diagnoses in the same subgroup, and cross-cancer concordance as both having diagnoses in separate subgroups. Two concordant twin pairs were identified for both lymphoid leukemia and AML (Table 2). Concordance rates were defined as the total number of concordant twin pairs divided by the sum of concordant and discordant twin pairs. The twin concordance rate was 0.8% (2 of 255 twin pairs) for all leukemias, 0.5% for lymphoid leukemias (1 of 199), and 2.9% for AML (1 of 34). The percentage of monozygotic (MZ) twins was estimated assuming dizygotic twins are equally likely to be represented as same sex or opposite sex twins.[3] The estimated MZ concordance rate was 2.6% for all leukemias, 1.4% for lymphoid leukemias, and 16.7% for AML. By diagnosis age, the rate of lymphoid leukemia concordance for 1-10 years was 0.7% (MZ: 2.1%). The rate of AML development for ages <1 year was 25% (MZ: 100%). Standardized incidence ratios (SIR) were calculated based on previously published methods.[4] The SIR for developing a leukemia of the same subgroup as the proband in twin siblings was 15.6 (95% Confidence Interval [CI]: 1.9-56.4), and 2.9 for development of any pediatric/AYA cancer (0.4-10.5) (*Online Supplementary Table S1*).

We next assessed associations between twin plurality order (being born first or second within the twin pair) and development of leukemia. Across the full twin cohort, there was a significant increased risk of leukemia development associated with being the first-born twin (Fisher's exact test, Odds Ratio [OR]: 1.80; *P*=0.001) (Figure 1A). Lymphoblastic leukemias were the most highly associated with being first-born (OR: 2.26; *P*<0.001) and this did not vary by age at cancer diagnosis or mode of birth (*data not shown*), while AML showed a non-significant decreased risk (OR: 0.55; *P*=0.127). To evaluate how plurality order may be reflected by DNA methylation at birth, we investigated 41 pediatric ALL-discordant twin pairs who had previously undergone genome-wide DNA methylation profiling at birth[5] for differences in DNA methylation based on twin plurality order. Deconvolution analysis showed no significant differences in nucleated cell proportions by plurality order. We used linear regression controling for ALL cases' status, birthweight, sex, twin pair number, nucleated cell proportions, and array batch effect to identify 1,394 significant differentially methylated probes (False Discovery Rate [FDR]: <0.05: 710,010 total CpG assessed), with a moderately elevated genomic inflation value of 2.21 (Figure 1B, C and *Online Supplementary Table S2*). There was a total of 534 significant differentially methylated regions assessed through Comb-P (Šidák, *P*<0.05) (Figure 1D and *Online Supplementary Table S3*).

We then evaluated CpG with known association to birth order[6] across subsequent gestations. Of 341 birth-order CpG, 305 overlapped with the 710,010 probes from the EPIC array ALL-discordant twin cohort. Mean delta-beta values (the first minus second-born absolute DNA methylation value) were significantly lower (hypomethylated) across the 24

twin pairs in which the ALL case was first-born compared to the 17 in which the case was second-born (Wilcoxon rank-sum, *P*=0.037) (Figure 1E). We further investigated 22,100 CpG sites nominally associated with cortisol levels,[7] of which 19,193 CpG overlapped with the EPIC array ALL twin cohort. Mean delta-beta values are significantly different in instances in which the ALL case sibling was born first compared to the case sibling being born second (*P*=0.011). We identified an overall concordance rate of 0.8% for twins with leukemia, which is notably lower than prior studies which range from 5-25%.[2] Our estimation of concordance in MZ twins for lymphoid leukemias remains low at 1.4%, while that of AML is within the previously reported range at 16.7%. Accounting for approximately 70% of MZ being monochorionic,[8] concordance increases to just 2.4% for ALL twins, and 3.4% specifically for the 1-10 year old age group. Remarkably, none of the 6 instances of lymphoid leukemia in twins <1 year of age demonstrated sibling concordance, despite prior estimated rates nearing 100%.[2] Twin studies previously helped to identify that the initiating somatic genetic translocation event in childhood leukemia development can occur in the intrauterine period and subsequently pass between siblings through shared placental circulation[9,10] in instances of MZ, monochorionic twins. The low rate of twin concordance in this study indicates that acquired genetic, epigenetic and discordant responses to environmental exposure events may contribute more heavily to pediatric and AYA cancer etiology than previously appreciated in pediatric acute leukemias,[11] or at least among those in recent decades in California. We identified a twin pair concordant for lymphoid leukemias with specific diagnoses of B-cell lymphoblastic leukemia/lymphoma and Burkitt's cell leukemia which do not share known driver mutations. This combination suggests an alternative disease mechanism of concordance than that of shared fetal circulation and may point toward shared genetic predisposition to hematologic malignancies, which we were not able to evaluate.

These results show that concordant cases of twin acute leukemia are rarer than previously reported, which may result from biases in active twin recruitment leading to oversampling of rare concordant twins relative to the true population background rate.[2] It is also notable that this study examined twins in a multiethnic birth population in which Latino births represent the majority population; most prior population-based studies had been performed in non-Latino Whites. While these results do not directly alter clinical recommendations,[12] additional nuance should be considered in how concordance risk is discussed with families after an initial diagnosis of ALL in a twin sibling.

This study was limited by the lack of zygosity information available in the evaluated registry data; however, our estimation of MZ twin pairs allows for broad comparison to previously identified concordance rates. In a subset of twins with ALL from this cohort investigated with a DNA single nucleotide polymorphism array, approximately 50% of same sex pairs were MZ,[5] which is consistent with the estimated results displayed here. Furthermore, we are limited by the right censoring of registry data (to the year 2022), raising potential for lead-time bias in twin siblings

Table 1. Twin leukemia cohort characteristics.

Characteristic	Value
Total N of twin pairs	255*
Opposite sex pairs, N (%)	89 (34.9)
Same sex pairs, N (%)	166 (65.1)
Same sex: Male, N (%)	101 (60.8)
Same sex: Female, N (%)	65 (39.2)
Birth characteristics	
Gestational age in days, median, IQR, range	260, 21, 183-309
Age of mother in years, median, IQR, range	31, 10, 16-49
Age of father in years, median, IQR, range	33, 10, 18-60
Cesarian delivery, N (%)	183 (71.8)
Vaginal delivery, N (%)	72 (28.2)
Use of assisted reproductive technology, N (%)	8 (3.1)
Birthweight in grams (SD, *P* value) Cases Controls	 2,501 (514, *P*=1) 2,473 (524)
Cancer diagnosis characteristics	
Total cancer diagnoses, N	265
Total leukemia diagnoses, N	258
Leukemia diagnoses by category* Lymphoid leukemias Acute myeloid leukemia Chronic myeloproliferative diseases MDS and other myeloproliferative diseases Unspecified and other specified leukemias	 200 35 11 8 4
Age of leukemia diagnosis in years, median, IQR, range	6, 10, 0-30
N of diagnoses by age category, N <1 year 1-9 years 10-19 years >20 years	 13 160 52 33

*Diagnoses were categorized by International Classification of Disease (ICD)-O-3 histology codes following the International Classification of Childhood Cancer, 3rd Edition, into five leukemia subgroups. IQR: interquartile range; MDS: myelodysplastic syndrome; N: number; SD: standard deviation.

born in the latest birth years assessed, though the inclusion of over three decades of cancer and birth registry data limits this impact.

These results indicate first-born twins are more likely to

Table 2. Twin pairs with concordant and cross-cancer concordant leukemia diagnoses.

Pair#		Cancer subgroup, ICCC-3 Recode	ICD-O3	Diagnosis	Sex	Age range, in years
Leukemia-concordant						
1	Twin A	AML	9872/3	AML, minimal differentiation	M	<1
	Twin B	AML	9872/3	AML, minimal differentiation	M	<1
2	Twin A	Lymphoid leukemias	9826/3	Burkitt's cell leukemia	M	1-10
	Twin B	Lymphoid leukemias	9811/3	B lymphoblastic leukemia/lymphoma, NOS	M	1-10
Cross-cancer concordant						
1	Twin A	Lymphoid leukemias	9811/3	B lymphoblastic leukemia/lymphoma, NOS	F	1-10
	Twin B	Rhabdomyosarcomas	8910/3	Embryonal rhabdomyosarcoma, NOS	F	1-10

Category	All			Same sex			Estimated MZ
	Conc	Disc	Rate, %	Conc	Disc	Rate, %	Rate, %
All leukemias	2	253	0.8	2	164	1.2	2.6
Lymphoid leukemias All ages <1 year old at diagnosis 1-10 years old at diagnosis >10 years old at diagnosis	 1 0 1 0	198 6 148 44	 0.5 0 0.7 0	 1 0 1 0	 133 5 97 31	 0.7 0 1.0 0	 1.4 0 2.0 0
AML All ages <1 year old at diagnosis 1-10 years old at diagnosis >10 years old at diagnosis	 1 1 0 0	 33 3 10 20	 2.9 25 0 0	 1 1 0 0	 19 2 7 10	 5.0 33 0 0	 16.7 100 0 0
Chronic myeloproliferative diseases	0	11	0	0	8	0	0
MDS and myeloproliferative diseases	0	8	0	0	3	0	0
Unspecified and other specified leukemias	0	4	0	0	1	0	0

Concordant cases were defined as twin siblings sharing leukemia diagnoses in the same ICCC-3 Recode subgroup. Cross-cancer concordant cases were defined as twin siblings with cancer diagnoses in separate recode subgroups, with at least one having a leukemia diagnosis. Bottom panel shows rates of concordance in twin cancer cases across all leukemia diagnoses, by leukemia subgroup and by age of diagnosis. The number of monozygotic (MZ) twin pairs was estimated based on the distribution of same sex and opposite sex twin pairs, assuming all concordant cases within the same sex group are from MZ twins. AML: acute myeloid leukemias; Conc: concordant; Disc: discordant; F: female; ICD: International Classification of Diseases; M: male; MDS: myelodysplastic syndrome; NOS: not otherwise specified.

develop ALL compared to second-born twins, a finding which had not previously been described. We conducted linear regression modeling in twins discordant for ALL to investigate how this relationship might be reflected in DNA methylation at birth in which we found evidence of genomic inflation indicative of higher-than-expected *P* values across most CpG probes on the EPIC array. This finding suggests the underlying biological mechanism responsible for the difference in leukemia risk between first- and second-born twins is reflected in a broad alteration (generally in the hypo-direction) of DNA methylation across the epigenome. Interestingly, first-born singleton births (compared to subsequent gestations) are at increased risk of ALL,[13] while the inverse is seen for pediatric AML,[14] which mirrors our

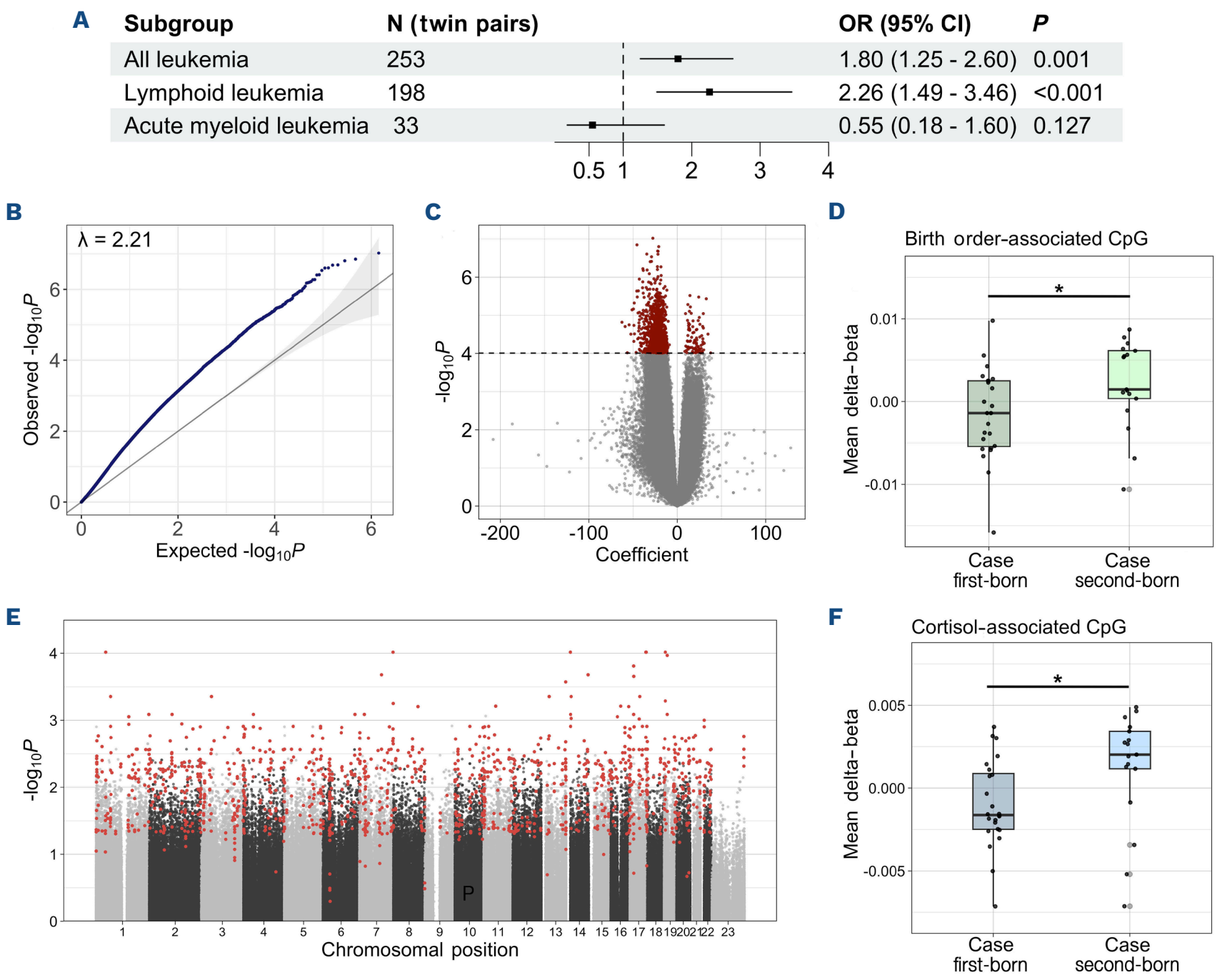

Figure 1. Birth plurality order is significantly associated with development of leukemia in leukemia-discordant twin pairs. (A) Forest plot demonstrating Odds Ratio (OR) from Fisher's exact test for the association between twin plurality order (first- or second-born of a set of twin siblings) and pediatric leukemia in leukemia-discordant twin pairs. A significantly higher odds of developing leukemia of any type is seen in first-born twins. By disease group, twins discordant for the development of lymphoid leukemia had significantly higher odds of being first-born. Twins discordant for development of acute myeloid leukemia, however, had lower odds of being first-born, but this difference was not statistically significant. (B) Quantile-quantile plot of linear regression model for the relationship between DNA methylation status and twin plurality order. Lambda value represents genomic inflation. (C) Volcano plot from the linear regression model showing significance ($-\log_{10}P$) by regression coefficient. The 1,394 significant differentially methylated probes (False Discovery Rate <0.05) are shown in red. (D) Volcano plot showing distribution of $-\log_{10}P$ values by genomic position. Probes found within the 534 significant differentially methylated regions are highlighted in red. (E) Distribution of mean delta-beta (first-born sibling DNA methylation beta value minus second-born within a twin pair) values across 305 CpG with known association to birth order found within the discordant acute lymphoblastic leukemia (ALL) twin DNA methylation array dataset. Mean delta-beta values are significantly different in instances in which the ALL case sibling was born first (N=24) compared to the case sibling being born second (N=17; Wilcoxon rank-sum, *P*=0.037). (F) Distribution of mean delta-beta values across 19,193 CpG with known association to post-partum cortisol levels found within the discordant ALL twin DNA methylation array dataset. Mean delta-beta values are significantly different in instances in which the ALL case sibling was born first (N=24) compared to the case sibling being born second (*P*=0.011). CI: Confidence interval; N: number.

results here with twins. The relationship of birth order with DNA methylation[6] also mirrors our results, suggesting the biological mechanism leading to an increase in ALL risk in first-born singletons may have similarities to that of first-born twins. We further found a significant association between the development of ALL and DNA methylation at cortisol-associated sites, indicating lower exposure to cortisol during birth may contribute to the significantly increased risk of ALL in first-born twins. Increased, early exposure to cortisol is hypothesized to reduce risk of ALL.[15] Further investigation in a broader cohort of twins with acute leukemia is warranted.

Authors

Eric M. Nickels,[1,2] Naying Zhou[2] and Joseph L. Wiemels[2]

[1]Cancer and Blood Disease Institute, Children's Hospital Los Angeles and [2]Center for Genetic Epidemiology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

Correspondence:
J.L. WIEMELS - wiemels@usc.edu

https://doi.org/10.3324/haematol.2024.286309

Received: August 6, 2024.
Accepted: December 31, 2024.
Early view: January 16, 2025.

Disclosures

No conflicts of interest to disclose.

Contributions

EMN is responsible for data curation, methodology, statistical analysis, and data visualization. NZ is responsible for statistical analysis. JLW is responsible for the study methodology. All authors wrote the manuscript.

Acknowledgments

The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement U58DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the author(s), and no endorsement by the California Department of Public Health, the National Cancer Institute, or the Centers for Disease Control and Prevention or their contractors and subcontractors is intended or should be inferred.

Funding

Funding was provided by the V Foundation (FP067172), the USC Norris Comprehensive Cancer Center SPORE development award (to JLW), and the Shirley McKernan Courage Foundation (CHLA) (to JLW).

Data-sharing statement

This study used biospecimens from the California Biobank Program. Per California Health and Safety Code Sections 124980(j), 124991(b), (g), (h), and 103850 (a) and (d), which protects the confidentiality of data obtained from biospecimens, we are respectfully unable to share raw, individual level genomic and genome-wide DNA methylation data reported in this study, which are the property of the State of California. Should we be contacted regarding individual level data contributing to the findings reported in this study, inquiries will be directed to the California Department of Public Health Institutional Review Board to establish an approved protocol to utilize the data, which cannot otherwise be shared peer-to-peer.

References

1. Kadan-Lottick NS, Kawashima T, Tomlinson G, et al. The risk of cancer in twins: a report from the childhood cancer survivor study. Pediatr Blood Cancer. 2006;46(4):476-481.
2. Greaves MF, Maia AT, Wiemels JL, Ford AM. Leukemia in twins: lessons in natural history. Blood. 2003;102(7):2321-2333.
3. Buckley JD, Buckley CM, Breslow NE, Draper GJ, Roberson PK, Mack TM. Concordance for childhood cancer in twins. Med Pediatr Oncol. 1996;26(4):223-229.
4. Feng Q, Nickels E, Muskens IS, et al. Increased burden of familial-associated early-onset cancer risk among minority Americans compared to non-Latino Whites. Elife. 2021;10:e64793.
5. Nickels EM, Li S, Myint SS, et al. DNA methylation at birth in monozygotic twins discordant for pediatric acute lymphoblastic leukemia. Nat Commun. 2022;13(1):6077.
6. Li S, Spitz N, Ghantous A, et al. A Pregnancy and Childhood Epigenetics Consortium (PACE) meta-analysis highlights potential relationships between birth order and neonatal blood DNA methylation. Commun Biol. 2024;7(1):66.
7. Houtepen LC, Vinkers CH, Carrillo-Roa T, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. Nat Commun. 2016;7:10967.
8. Hall JG. Twinning. Lancet. 2003;362(9385):735-743.
9. Greaves M. In utero origins of childhood leukaemia. Early Hum Dev. 2005;81(1):123-129.
10. Wiemels JL, Xiao Z, Buffler PA, et al. In utero origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia.

Blood. 2002;99(10):3801-3805.
11. Hong D, Gupta R, Ancliff P, et al. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. Science. 2008;319(5861):336-339.
12. Ford AM, Colman S, Greaves M. Covert pre-leukaemic clones in healthy co-twins of patients with childhood acute lymphoblastic leukaemia. Leukemia. 2023;37(1):47-52.
13. Paltiel O, Lemeshow S, Phillips GS, et al. The association between birth order and childhood leukemia may be modified by paternal age and birth weight. Pooled results from the International Childhood Cancer Cohort Consortium (I4C). Int J Cancer. 2019;144(1):26-33.
14. Von Behren J, Spector LG, Mueller BA, et al. Birth order and risk of childhood cancer: a pooled analysis from five US States. Int J Cancer. 2011;128(11):2709-2716.
15. Schmiegelow K, Vestergaard T, Nielsen SM, Hjalgrim H. Etiology of common childhood acute lymphoblastic leukemia: the adrenal hypothesis. Leukemia. 2008;22(12):2137-2141.