HHS Public Access

Author manuscript

*JCO Oncol Pract*. Author manuscript; available in PMC 2025 May 01.

Published in final edited form as:

*JCO Oncol Pract*. 2024 May ; 20(5): 631–642. doi:10.1200/OP.23.00325.

# Population-Based Data Linkage Describing Patterns of Cancer Clinical Trial Enrollment Among Children and Adolescents

**David A. Siegel, MD, MPH**[1], **Eric B. Durbin, DrPH, MS**[2], **Brad H. Pollock, PhD, MPH**[3], **Allison Grimes, MD**[4], **Lingyun Ji, PhD**[5,6], **Todd A. Alonzo, PhD**[5,6], **Sarah L. Vargas, PhD**[5], **Bin Huang, DrPH, MS**[2], **Jaclyn R. McDowell, DrPH**[2], **Ellen Lycan, DrPH**[2], **Peter Ransdell, MS**[2], **Eric Tai, MD, MS**[1], **Michael E. Roth, MD**[7], **David R. Freyer, DO, MS**[6,8]

[1]Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA

[2]Kentucky Cancer Registry, Markey Cancer Center, University of Kentucky, Lexington, KY

[3]University of California Davis Comprehensive Cancer Center, Sacramento, CA

[4]University of Texas Health Science Center at San Antonio, San Antonio, TX

**CORRESPONDING AUTHOR** Eric B. Durbin, DrPH, MS, Kentucky Cancer Registry, University of Kentucky, 2365 Harrodsburg Rd, Ste A230, Lexington, KY 40504-3381; ericd@kcr.uky.edu.

AUTHOR CONTRIBUTIONS

**Conception and design:** David A. Siegel, Eric B. Durbin, Brad H. Pollock, Allison Grimes, Eric Tai, Michael E. Roth, David R. Freyer

**Provision of study materials or patients:** Eric B. Durbin

**Collection and assembly of data:** David A. Siegel, Eric B. Durbin, Lingyun Ji, Todd A. Alonzo, Sarah L. Vargas, Ellen Lycan, Peter Ransdell

**Data analysis and interpretation:** David A. Siegel, Eric B. Durbin, Brad H. Pollock, Todd A. Alonzo, Bin Huang, Jaclyn R. McDowell, Eric Tai, Michael E. Roth, David R. Freyer

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI https://doi.org/10.1200/OP.23.00325.

DISCLAIMER

The findings and conclusions are those of the authors and do not necessarily represent the official positions of the Centers for Disease Control and Prevention or of the National Institutes of Health.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/op/authors/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Eric B. Durbin

**Travel, Accommodations, Expenses:** Caris Life Sciences, Inc

Brad H. Pollock

**Consulting or Advisory Role:** Astellas Pharma, LEO Pharma

Allison Grimes

**Speakers' Bureau:** Servier

**Expert Testimony:** Covenant Children's Hospital

Lingyun Ji

**Consulting or Advisory Role:** Takeda, Pfizer

Michael E. Roth

**Research Funding:** Eisai, Pfizer

No other potential conflicts of interest were reported.

[5]Children's Oncology Group, Monrovia, CA

[6]University of Southern California, Los Angeles, CA

[7]Division of Pediatrics, University of Texas MD Anderson Cancer Center, Houston, TX

[8]Children's Hospital Los Angeles, Los Angeles, CA

## Abstract

**PURPOSE—**Database linkage between cancer registries and clinical trial consortia has the potential to elucidate referral patterns of children and adolescents with newly diagnosed cancer, including enrollment into cancer clinical trials. This study's primary objective was to assess the feasibility of this linkage approach.

**METHODS—**Patients younger than 20 years diagnosed with incident cancer during 2012-2017 in the Kentucky Cancer Registry (KCR) were linked with patients enrolled in a Children's Oncology Group (COG) study. Matched patients between databases were described by sex, age, race and ethnicity, geographical location when diagnosed, and cancer type. Logistic regression modeling identified factors associated with COG study enrollment. Timeliness of patient identification by KCR was reported through the Centers for Disease Control and Prevention's Early Case Capture (ECC) program.

**RESULTS—**Of 1,357 patients reported to KCR, 47% were determined by matching to be enrolled in a COG study. Patients had greater odds of enrollment if they were age 0-4 years (ν 15-19 years), reported from a COG-affiliated institution, and had renal cancer, neuroblastoma, or leukemia. Patients had lower odds of enrollment if Hispanic (ν non-Hispanic White) or had epithelial (eg, thyroid, melanoma) cancer. Most (59%) patients were reported to KCR within 10 days of pathologic diagnosis.

**CONCLUSION—**Linkage of clinical trial data with cancer registries is a feasible approach for tracking patient referral and clinical trial enrollment patterns. Adolescents had lower enrollment compared with younger age groups, independent of cancer type. Population-based early case capture could guide interventions designed to increase cancer clinical trial enrollment.

## INTRODUCTION

Approximately 15,000 children and adolescents younger than 20 years are diagnosed with cancer each year in the United States.[1] However, most of these patients are not enrolled in clinical trials, especially among adolescents.[2] Clinical trial enrollment of pediatric patients with cancer is associated with lower mortality, better supportive care and psychological outcomes, and higher quality of life in survivorship.[3,4] Several barriers to clinical trial enrollment exist, including low referral of patients to treatment centers offering clinical trials, limited clinical trial availability, and geographic distance to clinical trials.[5,6] Some barriers disproportionally affect adolescents and young adults (AYAs), age 15-39 years at diagnosis, who are often treated at adult-focused cancer centers and less likely than children to be enrolled in clinical trials.[6] Understanding referral and clinical trial enrollment patterns can guide efforts to reduce clinical trial enrollment barriers faced by leading clinical cancer

trial organizations, such as the Children's Oncology Group (COG), and by other health professionals.

Central (eg, state and District of Columbia [DC]) cancer registries serve as potential data sources to identify gaps in clinical trial enrollment and to describe patient referral patterns. However, while central cancer registries identify >95% of patients with cancer, national cancer registry data are unavailable for analysis until 24-36 months postdiagnosis.[7] To increase the speed of pediatric cancer case ascertainment by cancer registries in the United States, the Centers for Disease Control and Prevention piloted the pediatric Early Case Capture (ECC) program to test the feasibility of hospitals, clinics, and laboratories reporting new cases of cancer to central cancer registries within 30 days of diagnosis instead of the required 6 months.[8,9] While the nine registries that participated in the ECC program, including the Kentucky Cancer Registry (KCR), provided reliable data with high completeness,[9] further assessment can evaluate whether these data could be used to more quickly and completely identify patients for potential clinical trial enrollment.

The objectives of this study were to assess the feasibility of linking data between a state cancer registry (KCR) and a database of a large clinical trials consortium (COG) to describe gaps in clinical trial enrollment and to assess the timeliness of case ascertainment of linked patients. Assessing patients by demographic and cancer characteristics who are missing from the COG registry could be useful for elucidating referral patterns and gaps in clinical trial enrollment, with the goal of building a data linkage framework that could be expanded to the broader US childhood and adolescent population with cancer.

## METHODS

This is a retrospective cohort study of data from the COG and KCR databases for patients younger than 20 years diagnosed during 2012-2017. In the KCR database, all patients were Kentucky residents diagnosed with malignant cancer (behavior code = 3),[10] including patients diagnosed out-of-state. KCR is charged with collecting data for 100% of Kentucky residents diagnosed with cancer. Data from the COG database represented patients diagnosed in every US state and DC, which allowed linkage to Kentucky residents diagnosed both in and out of state.[11] Patients in the COG database had both malignant and nonmalignant tumors. Patients in the COG database provided informed consent and were enrolled into one or more COG studies, including therapeutic clinical trials, registry studies, supportive care trials, and biology and specimen banking studies. Patients treated at a COG-affiliated hospital but did not enroll in a COG study were not included in the COG database. Patients were included in the linkage if they had either a first primary cancer or a second malignant neoplasm; for patients with two or more cancers, only the first cancer diagnosis was included in the results.

### Database Linkage

We securely transferred COG data to KCR staff, who performed database linkage on the basis of a probabilistic approach that included name (first, middle, and last names were available for KCR, and name initials were available from COG), date of birth, sex, postal zip code, cancer site (anatomic code[10]), histologic type, and diagnosis date. To compare

two records, blocking variables were used to increase the efficiency of probabilistic linkage and define at least one data element that must match. Blocking variables included the Soundex phonetic of first and last names, initials of first and last names, date of birth, date of diagnosis, cancer site, histologic type, and zip code. Match*Pro 1.6.5 (National Cancer Institute) was used to perform the probabilistic linkage. Two independent reviewers from KCR (E.B.D., J.R.M.) assessed matched patients, patients where matching was uncertain, and a proportion of unmatched patients as identified by Match*Pro. Reviewers used additional information to assist in the review of uncertain matches, including addresses, procedure/enrollment dates, and treatment/enrollment facilities. A Match*Pro cutoff score was used to determine the total volume of uncertain matches requiring review; the score was set very low to ensure the opportunity to review many potential pairs. Using this approach, when only patient initials were available for matching from COG, even minor discrepancies in other fields usually resulted in nonmatch. Approximately 1,000 record pairs were manually reviewed to ensure accuracy of the linkage match. The few (<five) disagreements between initial reviews were discussed by reviewers to arrive at consensus.

## Data Elements

Using a linked database file, we described matched and unmatched patient data after linkage. Matching of a patient indicated enrollment in a COG study. We described patient demographics including age, sex, race and ethnicity, Appalachian region,[12] if the patient was reported from a COG-affiliated hospital (including the two in Kentucky and those out-of-state), and the number of days after pathologic diagnosis the patient was reported to KCR. Race and ethnicity were defined as non-Hispanic Black, Hispanic, non-Hispanic Other (which included Asian/Pacific Islander, American Indian/Alaska Native, and individuals with multiple races documented), non-Hispanic White, and Unknown race and ethnicity. Appalachian geography was included as a variable because this region of Kentucky experiences a high burden of cancer incidence and poor cancer outcomes.[13] Clinical characteristics were stratified by cancer type defined by International Classification of Childhood Cancer (ICCC) categories.[14]

## Statistical Analysis

Multivariable logistic regression analysis was performed to identify factors significantly associated with cases being matched in the COG database. Adjusted odds ratios (ORs) and 95% CIs were used to reflect the magnitude of association between a factor and the odds of matching. The backward selection approach was used to identify the final model, which included only statistically significant variables except sex. Goodness-of-fit test statistics, including the Hosmer and Lemeshow test and several R2 measures, were examined for the model fit. All statistical tests were two-sided with a statistical significance level of <0.05, and all analyses were performed using SAS 9.4 (SAS Institute Inc, Cary, NC).

The study was conducted with approval from the Institutional Review Board (IRB) of the University of Kentucky. Data use agreements were signed by the Children’s Hospital of Philadelphia on behalf of COG, as per the COG research protocol, and the University of Kentucky. Existing COG written informed consents allow for this sharing of data.

# RESULTS

## KCR Case Demographics

During 2012-2017, KCR indicated 1,370 cancer cases among children and adolescents younger than 20 years; 1,357 patients were diagnosed with a new cancer (Table 1), and 13 patients had a second recorded malignant neoplasm. Of the 1,357 patients, 32% were diagnosed at age 0-4 years, 30% were diagnosed at age 15-19 years, 86% were non-Hispanic White, 504 (37%) patients were seen by an out-of-state facility, 359 (26%) were reported solely from hospitals outside of Kentucky, and 1,183 (87%) were reported from a COG-affiliated institution. Most patients seen outside of Kentucky were seen at Cincinnati Children's Hospital Medical Center (OH) and Vanderbilt-Ingram Cancer Center (TN). Both are COG-associated facilities in states adjacent to Kentucky. Of patients diagnosed between age 0 and 4 years, 402 of 430 (93%) were reported by a COG-affiliated institution, whereas 224 of 238 (94%), 250 of 277 (90%), and 307 of 412 (75%) were reported by a COG-affiliated institution for patients diagnosed between age 5 and 9, 10 and 14, and 15 and 19 years, respectively.

## COG Database Patient Demographics

In the COG database, 43,298 patients were identified during 2012-2017 (Table 1); 40,336 (93%) had a malignant behavior code, and 2,278 (5%) had a borderline behavior code. In the COG database, 41% were diagnosed at age 0-4 years, 18% were diagnosed at 15-19 years, and 57% were non-Hispanic White. Of the 43,298 patients, 38,508 (89%) were enrolled in registry studies, 18,209 (42%) in therapeutic trials, 6,875 (16%) in biology and specimen banking studies, and 3,910 (9%) in supportive care trials.

## Results of Patient Matching

Of 1,357 patients in the KCR database, as indicated by matching, 641 (47%) were enrolled in a COG study and 716 (53%) were not enrolled (Table 2). Two potential matches were determined not to match because of inconsistencies in the COG-provided first and last initials. One Kentucky resident was present in the COG database but not in the KCR database. Of 641 matched patients, 623 (97%) had tumors with a malignant behavior code and 18 (3%) had tumors with a borderline behavior code in the COG database, respectively. The percentage of patients enrolled was 64% among patients age 0-4 years at diagnosis and 29% among patients age 15-19 years. By race and ethnicity, non-Hispanic White patients had the highest enrollment (49%), followed by non-Hispanic Black (39%), non-Hispanic Other (37%), and Hispanic (31%). By cancer type, the highest percentage of enrolled patients were those with renal tumors (89%), neuroblastoma (86%), leukemias (83%), and bone tumors (69%); the lowest had epithelial malignancies and melanoma (4%). Patients reported from COG-affiliated hospitals had 52% enrollment compared with 14% at non-COG-affiliated hospitals.

Enrollment in a COG study varied by age of diagnosis and cancer type (Fig 1). For those diagnosed at a COG facility, the percentage enrolled was highest among patients age 0-4 years (64%) and lowest for patients age 15-19 years (37%; Fig 2). For those not diagnosed at

a COG facility, 14% of patients age 0-19 years were enrolled, including 48% of patients age 0-4 years.

### Regression Analysis

Multivariable logistic regression modeling showed that children age 0-4 years had greater odds of being enrolled in a COG study compared with patients age 15-19 years (adjusted OR, 1.75 [95% CI, 1.17 to 2.62]; Table 3). Hispanic patients had lower odds of enrolling compared with non-Hispanic White children. Patients with renal tumors (OR, 21.27 [95% CI, 7.11 to 63.65]), neuroblastoma (OR, 17.80 [95% CI, 8.02 to 39.55]), and leukemia (OR, 16.12 [95% CI, 10.41 to 24.99]) had the highest odds of enrollment compared with patients with CNS neoplasms. By year, patients diagnosed in 2012 or 2013 had greater odds of enrollment ($\nu$ 2017). Patients from non-Appalachian versus Appalachian regions had lower odds of enrollment (OR, 0.56 [95% CI, 0.41 to 0.77]).

### ECC Results

Of 1,357 patients, 892 (66%) were reported to KCR in ≤30 days (Table 2). Of these 892 patients, 53% matched with the COG database. Only 36% of patients matched among the 465 patients reported in >30 days. Of the 1,357 patients, 802 (59%) were reported to KCR in ≤10 days (Fig 3) and differences were seen by matching status, age, and cancer type. Of the cases reported from a Kentucky facility, 85% were reported within 30 days. Of the cases reported from outside of Kentucky, only 8% were reported within 30 days.

## DISCUSSION

This study demonstrated the feasibility of linking central cancer registry data with data from a pediatric cancer clinical trial consortium. By linking COG data with KCR data, we determined that of Kentucky residents younger than 20 years diagnosed with cancer during 2012-2017, 47% enrolled in a COG study. This linkage strategy allowed us to describe COG enrollment patterns by patient demographics and cancer type. As part of a program focusing on early reporting, two thirds of patients were reported to KCR in ≤30 days and 59% were reported in ≤10 days.

Previous studies have successfully linked central cancer registry data with electronic health record data.[15-17] Studies of adult populations used linkage of clinical trial data with central cancer registry data to identify clinical trial enrollment gaps.[18,19] Similar to previous studies,[16,17] we found that several logistical elements were essential for completing this linkage. First, we identified variables with high completeness for linkage. Second, manual review was needed for a subset of potential matches; staff time needed for this step was not a major barrier because of the efficient and accessible interface available in the Match*Pro linkage software. Third, we identified variables for analysis (eg, patient demographics and tumor characteristics) that were standardized across databases or could be recoded; COG and KCR used similar coding for demographic and histology variables, and discrepancies were straightforward to resolve. Finally, evaluation of regulatory requirements included review of data sharing agreements, IRB exemption criteria, and confirmation of use of data allowed by COG patient consent agreements. Our experience suggests that these

considerations can be satisfactorily addressed, making this approach feasible. Future efforts could focus on how to increase the speed of the linkage steps to facilitate rapid identification of patients.

The finding that younger patients had greater odds of enrollment in COG studies is similar to previous reports.[20-22] A 2003 report of registration rates in the Children's Cancer Group and the Pediatric Oncology Group (which merged in 2000 to become COG) stated that the percentage of patients enrolled into clinical trials among those younger than 15 years was three times higher than those age 15-19 years.[20,23] High enrollment rates for patients age 0-4 years likely reflect available trials for the most common diagnoses, specifically acute lymphoblastic leukemia.[24] This study found higher enrollment rates among patients age 0-4 years even among patients not reported from a COG-affiliated institution; future investigation of this referral pattern could be used to develop interventions for patient groups with lower enrollment not diagnosed at COG-affiliated institutions. Our results are also consistent with previous studies documenting enrollment in COG registry studies, such as the Childhood Cancer Research Network (CCRN)[25,26]; an estimated 36% of pediatric cancer cases enrolled in CCRN during 2008-2015.[25] A similar study during 2004-2015 estimated that 19.9% of patients younger than 20 years enrolled in up-front COG therapeutic trials.[22]

Observed differences in matching by cancer type are consistent with the published literature: patients with leukemia and kidney cancer had greater odds of enrollment in a COG study, whereas patients with germ cell tumors and melanoma had lower odds of enrollment.[20,25] ICCC type XI epithelial tumors and melanomas, which contain thyroid cancers, are more common among adults than children. Low enrollment of patients with these cancer types in our predominantly pediatric sample is likely driven by their low incidence, non-availability of COG clinical trials for these cancer types, and probable referral of these patients to nonpediatric oncology providers. Enrollment differed by year. The odds of matching in any one particular year might depend on time-related factors including the number or type of open clinical trials, incidence trends by cancer type, and changes in regional referral patterns. Future studies could further examine differences between diagnosis and enrollment year (Appendix Table A1, online only) and the number or type of open clinical trials, which may require detailed examination by cancer type, age, and local incidence. Findings in this study showing lower enrollment among Hispanic children and adolescents are consistent with findings by Aristizabal et al,[21] who cited language and cultural barriers as potential causes. Interventions addressing social determinants of health, including race and ethnicity, in pediatric cancer clinical trials are feasible[27]; use of linked data could help inform these interventions.

There are several potential explanations for low enrollment in COG studies: patients and families might choose to not participate in available clinical trials; they might be ineligible, they might not be offered enrollment, or there might not be an available clinical trial for that cancer type or stage.[21,28] Enrollment on registry studies might be lower if a therapeutic trial is not available for that patient.[25] Specifically, at non-COG hospitals, patients might not be aware of available studies, might not be referred by their providers, or might not have geographic access.[6,21] Moreover, AYA patients have lower rates of clinical trial enrollment when treated by medical rather than pediatric hematologists/oncologists; lower enrollment

was also noted by AYA at non-COG sites compared with COG-affiliated hospitals.[29,30] Patients are more likely to receive care in a community setting as they enter AYA range,[31,32] which may lead to lower enrollment.

Of patients with CNS neoplasms, most (51%) were reported to KCR after 30 days, potentially because of some cases needing multiple external pathology consultations or possibly longer times needed for unbiopsied tumors, which might take longer to report to central cancer registries (eg, radiographic diagnosis). However, some cancer types with low enrollment percentages (eg, germ cell cancer and epithelial tumors and melanoma) had most patients reported to KCR in the first 30 days. Patients diagnosed outside of Kentucky (eg, OH) were less likely to be reported to KCR in the first 30 days; most of these cases were reported to KCR from an out-of-state cancer registry. For cancer types not frequently reported in the first few days (eg, retinoblastoma, CNS neoplasms), targeted interventions with clinical providers may be indicated to reduce reporting time.[20]

Rapid description of enrollment patterns could potentially help clinical trial organizations or central cancer registries design interventions to increase clinical trial enrollment. Clinical trial organizations such as COG could routinely and frequently link data with cancer registries, providing registries a list of patients not enrolled. In this case, these lists could be used to identify either COG or non–COG-affiliated institutions with lower enrollment and gaps in enrollment by demographic characteristic or cancer type. This information could be used to develop institution- or physician-specific outreach interventions to increase awareness of and referral to potential clinical trials. In addition, cancer registry data have been successfully used to identify patients for clinical interventions after the end of treatment.[33] A similar process could be possible for pediatric patients with cancer. Although the timeline would likely not be rapid enough to influence enrollment in initial therapeutic trials, rapidly linked data could be used to identify patients for downstream enrollment into relapse, supportive care, survivorship and late outcome, and registry or biology studies. Potentially, with consent obtained by the central cancer registry, hospitals or registries could use this information for specific patient outreach. However, feasibility studies would be needed to assess the timeliness and logistical elements of this process.

A strength of this study is the use of a central cancer registry to link with a clinical database using a scalable methodology. However, this study has some limitations. Because linkage was limited to Kentucky residents, findings might not be generalizable nationally; the linkage methodology did not address state-by-state variation in data and regulatory requirements. Absence of full name or other identifiers in the COG database might have resulted in a small number of negative mismatches. Although there are no large-scale, competing clinical trial organizations for pediatric cancer, COG enrollment might underestimate total clinical trial enrollment because some patients in Kentucky might have enrolled in non-COG, institution-specific clinical trials (ie, out-of-state trials for patients with multiple-relapsed disease) or, for age 18-19 years, some patients might have enrolled in trials designed for adult patients.[4,34] Finally, we were not able to assess reasons for a patient not enrolling on a COG study.

COG data linkage with multiple central cancer registries could help clinicians describe enrollment patterns across a large geographic area, and linkage could expand to age >19 years to assess enrollment patterns for young adults with cancer. A future national linkage of clinical trial and cancer registry data could be used for real-time identification of patients and could contribute to data sharing initiatives. A national linkage could enhance programs' ability, such as the National Childhood Cancer Registry (NCCR), to be used as a tool for data sharing, patient identification, and assessment of gaps of coverage.[35] National registries (eg, NCCR) linked with clinical trial data could be invaluable for governmental initiatives such as the Childhood Cancer Data Initiative and CC-DIRECT, which aim to increase the power of randomized clinical trials, promote access to valuable research databases, and ensure that families of children with cancer are connected with clinical trial and research opportunities.[36-39]

## ACKNOWLEDGMENT

The authors would like to thank State health departments and cancer registrars and Cancer Research Informatics Shared Resource Facility and Biostatistics and Bioinformatics Shared Resource Facility of the Markey Cancer Center.

### SUPPORT

Supported by Children's Oncology Group NCORP Research Base grant, UG1 CA189955; NCTN Operations Center Grant U10 CA180886; NCTN Statistics & Data Center Grant U10 CA180899; Centers for Disease Control and Prevention NPCR NU58DP007144; National Cancer Institute SEER HHSN26100001; National Cancer Institute P30 CA177558; and Commonwealth of Kentucky.

## DATA SHARING STATEMENT

Data from the Kentucky Cancer Registry and Children's Oncology Group that support the findings of this study are available in deidentified format on request. These data are not publicly available with personal identifiable information due to privacy and legal restrictions.

## APPENDIX

**TABLE A1.**

Differences Between Diagnosis and Enrollment Year Among Patients in COG Database

| | Enrollment Year | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Diagnosis Year** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **Total** |
| All COG cases | | | | | | | | | | |
| Total | 6,672 | 7,383 | 7,281 | 7,337 | 6,915 | 6,728 | 717 | 208 | 57 | 43,298 |
| 2012 | 6,672 | 984 | 155 | 84 | 28 | 24 | 11 | 15 | 4 | 7,977 |
| 2013 | 0 | 6,399 | 831 | 160 | 59 | 27 | 16 | 17 | 6 | 7,515 |
| 2014 | 0 | 0 | 6,295 | 877 | 139 | 58 | 28 | 20 | 6 | 7,423 |
| 2015 | 0 | 0 | 0 | 6,216 | 743 | 162 | 46 | 34 | 6 | 7,207 |
| 2016 | 0 | 0 | 0 | 0 | 5,946 | 797 | 79 | 41 | 15 | 6,878 |
| 2017 | 0 | 0 | 0 | 0 | 0 | 5,660 | 537 | 81 | 20 | 6,298 |

| | Enrollment Year | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Diagnosis Year** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **Total** |
| Matched cases between COG and KCR | | | | | | | | | | |
| Total | 104 | 103 | 118 | 93 | 100 | 112 | 10 | 0 | 1 | 641 |
| 2012 | 104 | 18 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 124 |
| 2013 | 0 | 85 | 16 | 1 | 0 | 0 | 0 | 0 | 0 | 102 |
| 2014 | 0 | 0 | 101 | 5 | 0 | 0 | 0 | 0 | 0 | 106 |
| 2015 | 0 | 0 | 0 | 86 | 7 | 3 | 0 | 0 | 0 | 96 |
| 2016 | 0 | 0 | 0 | 0 | 93 | 12 | 0 | 0 | 0 | 105 |
| 2017 | 0 | 0 | 0 | 0 | 0 | 97 | 10 | 0 | 1 | 108 |

Abbreviations: COG, Children's Oncology Group; KCR, Kentucky Cancer Registry.

## REFERENCES


1. U.S. Cancer Statistics Working Group: U.S. Cancer Statistics Data Visualizations Tool, based on 2021 submission data (1999-2019): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. www.cdc.gov/cancer/dataviz
2. Thomas SM, Malvar J, Tran HH, et al. : A prospective comparison of cancer clinical trial availability and enrollment among adolescents/young adults treated at an adult cancer hospital or affiliated children's hospital. Cancer 124:4064–4071, 2018 [PubMed: 30291804]
3. Abdelhadi OA, Joseph J, Pollock BH, et al. : Additional medical costs of chronic conditions among adolescent and young adult cancer survivors. J Cancer Surviv 16:487–496, 2022 [PubMed: 33899161]
4. Bleyer A, Montello M, Budd T, et al. : National survival trends of young adults with sarcoma: Lack of progress is associated with lack of clinical trial participation. Cancer 103:1891–1897, 2005 [PubMed: 15795902]
5. Russo C, Stout L, House T, et al. : Barriers and facilitators of clinical trial enrollment in a network of community-based pediatric oncology clinics. Pediatr Blood Cancer 67:e28023, 2020 [PubMed: 31556250]
6. Tai E, Buchanan N, Eliman D, et al. : Understanding and addressing the lack of clinical trial enrollment among adolescents with cancer. Pediatrics 133:S98–S103, 2014 (suppl 3) [PubMed: 24918214]
7. NPCR standards. Centers for Disease Control and Prevention. https://www.cdc.gov/cancer/npcr/standards.htm
8. H.R. 1553, 110th Congress, 2nd Session (2008). https://www.congress.gov/bill/110th-congress/house-bill/1553/text
9. Puckett M, Neri A, Rohan E, et al. : Evaluating early case capture of pediatric cancers in seven central cancer registries in the United States, 2013. Public Health Rep 131:126–136, 2016 [PubMed: 26843678]
10. Fritz A, Percy C, Jack A, et al.: International Classification of Diseases for Oncology (ed 3). Geneva, World Health Organization, 2000
11. About us. Children's Oncology Group. https://childrensoncologygroup.org/about
12. Appalachian Regional Commission. https://www.arc.gov/
13. Yao N, Alcalá HE, Anderson R, et al. : Cancer disparities in rural Appalachia: Incidence, early detection, and survivorship. J Rural Health 33:375–381, 2017 [PubMed: 27602545]
14. Steliarova-Foucher E, Stiller C, Lacour B, et al. : International Classification of Childhood Cancer, third edition. Cancer 103:1457–1467, 2005 [PubMed: 15712273]

15. Ruppert LP, He J, Martin J, et al. : Linkage of Indiana State Cancer Registry and Indiana Network for Patient Care data. J Registry Manag 43:174–178, 2016 [PubMed: 29595920]
16. Hernandez MN, Voti L, Feldman JD, et al. : Cancer registry enrichment via linkage with hospital-based electronic medical records: A pilot investigation. J Registry Manag 40:40–47, 2013 [PubMed: 23778697]
17. Thompson CA, Jin A, Luft HS, et al. : Population-based registry linkages to improve validity of electronic health record-based cancer research. Cancer Epidemiol Biomarkers Prev 29:796–806, 2020 [PubMed: 32066621]
18. Green AK, Tabatabai SM, Bai X, et al. : Validation of a population-based data source to examine national cancer clinical trial participation. JAMA Netw Open 5:e223687, 2022 [PubMed: 35315914]
19. Lund JL, Meyer AM, Deal AM, et al. : Data linkage to improve geriatric oncology research: A feasibility study. Oncologist 22:1002–1005, 2017 [PubMed: 28408619]
20. Liu L, Krailo M, Reaman GH, et al. : Childhood cancer patients' access to cooperative group cancer programs: A population-based study. Cancer 97:1339–1345, 2003 [PubMed: 12599243]
21. Aristizabal P, Singer J, Cooper R, et al. : Participation in pediatric oncology research protocols: Racial/ethnic, language and age-based disparities. Pediatr Blood Cancer 62:1337–1344, 2015 [PubMed: 25755225]
22. Faulk KE, Anderson-Mellies A, Cockburn M, et al. : Assessment of enrollment characteristics for Children's Oncology Group (COG) upfront therapeutic clinical trials 2004-2015. PLoS One 15:e0230824, 2020 [PubMed: 32324751]
23. Our history. Children's Oncology Group. https://childrensoncologygroup.org/history
24. Siegel DA, King JB, Lupo PJ, et al. : Counts, incidence rates, and trends of pediatric cancer in the United States, 2003-2019. J Natl Cancer Inst 115:1337–1354, 2023 [PubMed: 37433078]
25. Brown AL, Sok P, Scheurer ME, et al. : An updated assessment of 43,110 patients enrolled in the Childhood Cancer Research Network: A Children's Oncology Group report. Cancer 128:2760–2767, 2022 [PubMed: 35482017]
26. Musselman JR, Spector LG, Krailo MD, et al. : The Children's Oncology Group Childhood Cancer Research Network (CCRN): Case catchment in the United States. Cancer 120:3007–3015, 2014 [PubMed: 24889136]
27. Aziz-Bose R, Zheng DJ, Umaretiya PJ, et al. : Feasibility of oncology clinical trial-embedded evaluation of social determinants of health. Pediatr Blood Cancer 69:e29933, 2022 [PubMed: 36069432]
28. Howerton MW, Gibbons MC, Baffi CR, et al. : Provider roles in the recruitment of underrepresented populations to cancer clinical trials. Cancer 109:465–476, 2007 [PubMed: 17200964]
29. Parsons HM, Penn DC, Li Q, et al. : Increased clinical trial enrollment among adolescent and young adult cancer patients between 2006 and 2012-2013 in the United States. Pediatr Blood Cancer 66:e27426, 2019 [PubMed: 30256525]
30. Roth ME, O'Mara AM, Seibel NL, et al. : Low enrollment of adolescents and young adults onto cancer trials: Insights from the community clinical oncology program. JCO Oncol Pract 12:e388–e395, 2016
31. Albritton KH, Wiggins CH, Nelson HE, et al. : Site of oncologic specialty care for older adolescents in Utah. J Clin Oncol 25:4616–4621, 2007 [PubMed: 17925557]
32. Yeager ND, Hoshaw-Woodard S, Ruymann FB, et al. : Patterns of care among adolescents with malignancy in Ohio. J Pediatr Hematol Oncol 28:17–22, 2006 [PubMed: 16394887]
33. Krebs P, Rogers E, Greenspan A, et al. : Utility of using cancer registry data to identify patients for tobacco treatment trials. J Registry Manag 46:30–36, 2019 [PubMed: 32010425]
34. Collins CL, Malvar J, Hamilton AS, et al. : Case-linked analysis of clinical trial enrollment among adolescents and young adults at a National Cancer Institute-designated comprehensive cancer center. Cancer 121:4398–4406, 2015 [PubMed: 26393950]
35. NCCR*Explorer: An interactive website for NCCR cancer statistics. National Cancer Institute. https://NCCRExplorer.ccdi.cancer.gov/

36. Major A, Cox SM, Volchenboum SL: Using big data in pediatric oncology: Current applications and future directions. Semin Oncol 47:56–64, 2020 [PubMed: 32229032]

37. Plana A, Furner B, Palese M, et al. : Pediatric cancer data commons: Federating and democratizing data for childhood cancer research. JCO Clin Cancer Inform 10.1200/CCI.21.00075

38. Childhood Cancer Data Initiative (CCDI). National Cancer Institute, 2022. https://www.cancer.gov/research/areas/childhood/childhood-cancer-data-initiative

39. The White House. FACT SHEET: On one year anniversary of reignited cancer moonshot, biden-harris administration announces new actions to end cancer as we know it. https://www.whitehouse.gov/briefing-room/statements-releases/2023/02/02/fact-sheet-on-one-year-anniversary-of-reignited-cancer-moonshot-biden-harris-administration-announces-new-actions-to-end-cancer-as-we-know-it/

## CONTEXT

### Key Objective

Most children and adolescents with cancer are not enrolled in clinical trials. Evaluating which patients are not enrolled may help oncologists increase enrollment and improve outcomes. We investigated if data from a state cancer registry can be linked to clinical trial data to describe gaps in clinical trial enrollment.

### Knowledge Generated

By matching state cancer registry data with data from a clinical trial organization, we found that only 47% of patients enrolled in clinical trials and enrollment in clinical trials was most likely for age 0-4 years, for non-Hispanic White patients, and for patients with kidney cancer, neuroblastoma, and leukemia.

### Relevance

This study demonstrated the feasibility of linking central cancer registry data with data from a pediatric cancer clinical trial consortium. This method could be used to assess diagnosis and referral patterns, address barriers to clinical trial enrollment, and improve quality of pediatric cancer care and research.

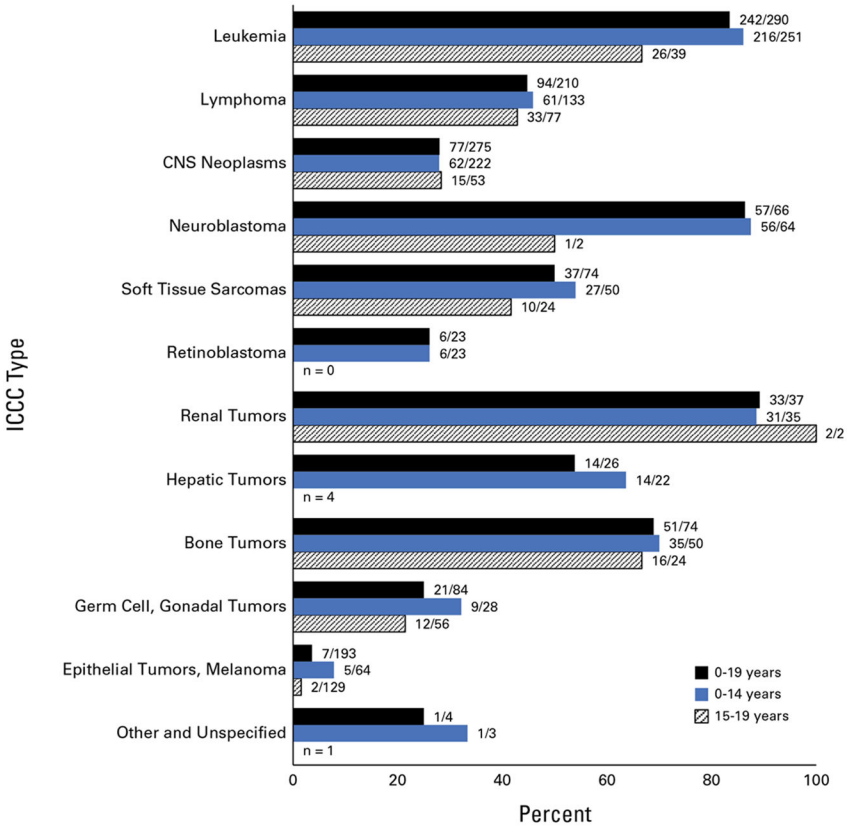


**FIG 1.**

Percent of patients with cancer among Kentucky residents younger than 20 years at diagnosis enrolled in a Children's Oncology Group study by age and cancer type, 2012-2017. ICCC, International Classification of Childhood Cancer.

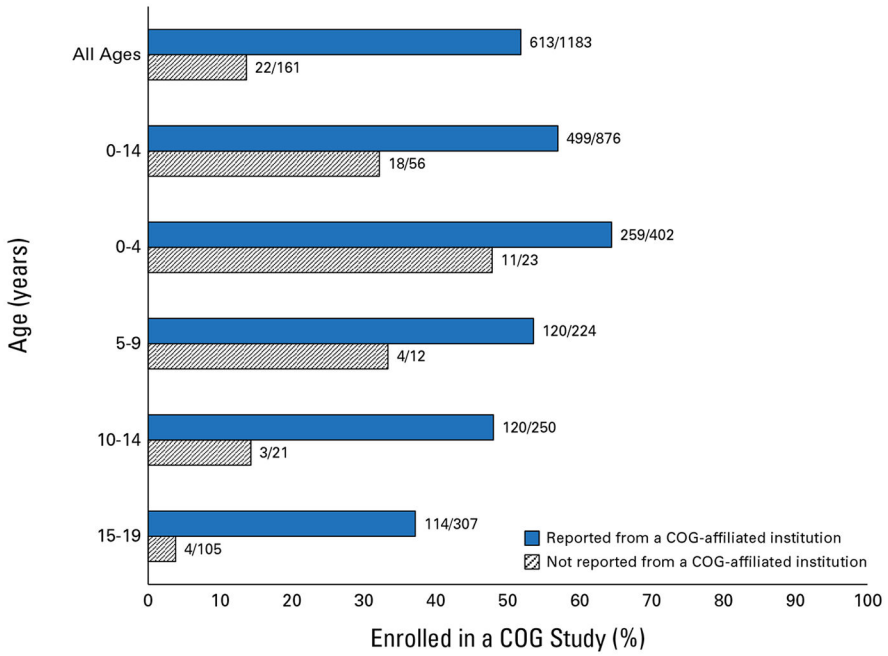


**FIG 2.**

Percent of Kentucky residents younger than 20 years with cancer enrolled in a COG study by age and reporting source, 2012-2017. This figure does not include 13 cases of unknown reporting source. COG, Children's Oncology Group.

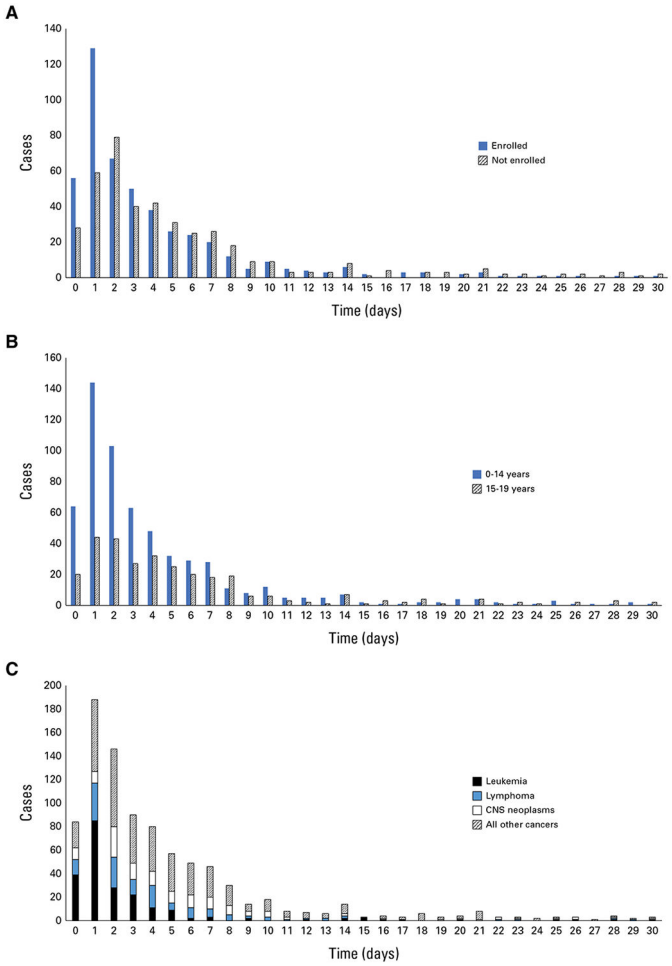


**FIG 3.**

Time between diagnosis and case report to the Kentucky Cancer Registry among patients with cancer younger than 20 years reported in ≤30 days, 2012-2017. (A) By enrollment status with Children's Oncology Group, (B) by age, (C) by ICCC type. ICCC, International Classification of Childhood Cancer.

**TABLE 1.**

Patients Younger Than 20 Years Registered by the COG and KCR, 2012-2017

| Characteristic | COG Database,[a] No. (%) | KCR Database,[b] No. (%) |
|---|---|---|
| Total | 43,298 (100) | 1,357 (100) |
| Sex[c] | | |
| Female | 19,352 (45) | 626 (46) |
| Male | 23,944 (55) | 731 (54) |
| Age at diagnosis, years | | |
| 0-14 | 35,688 (82) | 945 (70) |
| 0-4 | 17,725 (41) | 430 (32) |
| 5-9 | 9,230 (21) | 238 (18) |
| 10-14 | 8,733 (20) | 277 (20) |
| 15-19 | 7,610 (18) | 412 (30) |
| Race and ethnicity | | |
| Black, non-Hispanic | 4,384 (10) | 102 (8) |
| Hispanic | 7,047 (16) | 48 (4) |
| Other, non-Hispanic | 1,959 (5) | 19 (1) |
| Unknown | 5,374 (12) | 24 (2) |
| White, non-Hispanic | 24,534 (57) | 1,164 (86) |
| Cancer type (ICCC)[d] | | |
| Leukemia | 17,412 (40) | 290 (21) |
| Lymphoma | 5,411 (12) | 210 (15) |
| CNS neoplasms | 5,487 (13) | 275 (20) |
| Neuroblastoma and sympathetic nervous system tumors | 3,508 (8) | 66 (5) |
| Soft tissue sarcomas | 2,633 (6) | 74 (5) |
| Retinoblastoma | 481 (1) | 23 (2) |
| Renal tumors | 2,844 (7) | 37 (3) |
| Hepatic tumors | 627 (1) | 26 (2) |
| Bone tumors | 2,444 (6) | 74 (5) |

| Characteristic | COG Database,[a] No. (%) | KCR Database,[b] No. (%) |
|---|---|---|
| Germ cell and gonadal tumors | 1,486 (3) | 84 (6) |
| Epithelial tumors and melanoma | 492 (1) | 193 (14) |
| Other and unspecified | 223 (1) | 4 (0) |
| Unknown/unclassified | 250 (1) | 1 (0) |
| Diagnosis year | | |
| 2012 | 7,977 (18) | 220 (16) |
| 2013 | 7,515 (17) | 209 (15) |
| 2014 | 7,423 (17) | 212 (16) |
| 2015 | 7,207 (17) | 230 (17) |
| 2016 | 6,878 (16) | 234 (17) |
| 2017 | 6,298 (15) | 252 (19) |
| Geographic location | | |
| Appalachian | | 379 (28) |
| Non-Appalachian | | 978 (72) |
| Reported from a COG-affiliated institution (per KCR) | | |
| Yes | 43,298 (100) | 1,183 (87) |
| No | | 161 (12) |
| Unknown | | 13 (1) |

Abbreviations: COG, Children's Oncology Group; ICCC, International Classification of Childhood Cancer; KCR, Kentucky Cancer Registry.

[a]Total number of patients; patients with multiple primaries are counted by their first primary; the denominator for percent calculations is 43,298.

[b]Total number of patients; patients with multiple primaries are counted by their first primary; the denominator for percent calculations is 1,357.

[c]COG database had two patients with sex not specified.

[d]For patients with two or more cancers, only the first cancer diagnosis was described by cancer type.

**TABLE 2.**

Patients Younger Than 20 Years Registered by the KCR and Enrolled in a COG Study, 2012-2017

| | Patients in the KCR | | Patients Reported to the KCR in ≤30 Days | |
|---|---|---|---|---|
| **Characteristic** | **Total (%)[a]** | **Enrolled With COG (%)[b]** | **Total (%)[c]** | **Enrolled With COG (%)[d]** |
| Total | 1,357 (100) | 641 (47) | 892 (66) | 475 (53) |
| Sex | | | | |
| Female | 626 (46) | 263 (42) | 419 (67) | 195 (47) |
| Male | 731 (54) | 378 (52) | 473 (65) | 280 (59) |
| Age at diagnosis, years | | | | |
| 0-14 | 945 (70) | 523 (55) | 593 (63) | 385 (65) |
| 0-4 | 430 (32) | 274 (64) | 262 (61) | 199 (76) |
| 5-9 | 238 (18) | 124 (52) | 150 (63) | 87 (58) |
| 10-14 | 277 (20) | 125 (45) | 181 (65) | 99 (55) |
| 15-19 | 412 (30) | 118 (29) | 299 (73) | 90 (30) |
| Race and ethnicity | | | | |
| Black, non-Hispanic | 102 (8) | 40 (39) | 75 (74) | 33 (44) |
| Hispanic | 48 (4) | 15 (31) | 37 (77) | 11 (30) |
| Other, non-Hispanic | 19 (1) | 7 (37) | 15 (79) | 7 (47) |
| Unknown | 24 (2) | 9 (38) | 7 (29) | 2 (29) |
| White, non-Hispanic | 1,164 (86) | 570 (49) | 758 (65) | 422 (56) |
| Cancer type (ICCC)[e] | | | | |
| Leukemia | 290 (21) | 242 (83) | 214 (74) | 187 (87) |
| Lymphoma | 210 (15) | 94 (45) | 145 (69) | 74 (51) |
| CNS neoplasms | 275 (20) | 77 (28) | 136 (49) | 50 (37) |
| Neuroblastoma and sympathetic nervous system tumors | 66 (5) | 57 (86) | 45 (68) | 39 (87) |
| Soft tissue sarcomas | 74 (5) | 37 (50) | 57 (77) | 31 (54) |
| Retinoblastoma | 23 (2) | 6 (26) | 5 (22) | 2 (40) |
| Renal tumors | 37 (3) | 33 (89) | 25 (68) | 22 (88) |
| Hepatic tumors | 26 (2) | 14 (54) | 18 (69) | 9 (50) |

| Characteristic | Patients in the KCR | | Patients Reported to the KCR in ≤30 Days | |
|---|---|---|---|---|
| | Total (%)[a] | Enrolled With COG (%)[b] | Total (%)[c] | Enrolled With COG (%)[d] |
| Bone tumors | 74 (5) | 51 (69) | 48 (65) | 36 (75) |
| Germ cell and gonadal tumors | 84 (6) | 21 (25) | 55 (65) | 18 (33) |
| Epithelial tumors and melanoma | 193 (14) | 7 (4) | 142 (74) | 6 (4) |
| Other and unspecified | 4 (0) | 1 (25) | 1 (25) | 0 (0) |
| Unknown/unclassified | 1 (0) | 1 (100) | 1 (100) | 1 (100) |
| Diagnosis year | | | | |
| 2012 | 220 (16) | 122 (55) | 133 (60) | 82 (62) |
| 2013 | 209 (15) | 105 (50) | 130 (62) | 69 (53) |
| 2014 | 212 (16) | 106 (50) | 124 (58) | 70 (56) |
| 2015 | 230 (17) | 94 (41) | 160 (70) | 75 (47) |
| 2016 | 234 (17) | 108 (46) | 167 (71) | 95 (57) |
| 2017 | 252 (19) | 106 (42) | 178 (71) | 84 (47) |
| Geographic location | | | | |
| Appalachian | 379 (28) | 195 (51) | 266 (70) | 159 (60) |
| Non-Appalachian | 978 (72) | 446 (46) | 626 (64) | 316 (50) |
| Reported from a COG-affiliated institution | | | | |
| Yes | 1,183 (87) | 613 (52) | 784 (66) | 463 (59) |
| No | 161 (12) | 22 (14) | 99 (61) | 8 (8) |
| Unknown | 13 (1) | 6 (46) | 9 (69) | 4 (44) |

Abbreviations: COG, Children's Oncology Group; ICCC, International Classification of Childhood Cancer; KCR, Kentucky Cancer Registry.

[a]Total number of patients; patients with multiple primaries are counted by their first primary; the denominator for percent calculations is 1,357.

[b]Denominator for percent calculations is row count from the column with footnote a.

[c]Denominator for percent calculations is row count from the column with footnote a.

[d]Denominator for percent calculations is row count from the column with footnote c.

[e]For patients with two or more cancers, only the first cancer diagnosis was described by cancer type.

**TABLE 3.**

Multivariable Regression Analysis Showing Odds of Enrollment in a Children's Oncology Group Study Among Patients Reported by Kentucky Cancer Registry, Age <20 Years, 2012-2017

| Characteristic | Sample Size | Odds Ratio (95% CI) | *P* |
|---|---|---|---|
| Sex | | | |
| Female | 626 | Reference | |
| Male | 731 | 1.15 (0.87 to 1.52) | .32 |
| Age at diagnosis, years | | | |
| 0-4 | 430 | 1.75 (1.17 to 2.62) | .007 |
| 5-9 | 238 | 1.31 (0.86 to 2.00) | .21 |
| 10-14 | 277 | 1.43 (0.95 to 2.14) | .09 |
| 15-19 | 412 | Reference | |
| Race and ethnicity | | | |
| Black, non-Hispanic | 102 | 0.72 (0.43 to 1.22) | .23 |
| Hispanic | 48 | 0.33 (0.15 to 0.70) | .004 |
| Other, non-Hispanic | 19 | 0.44 (0.15 to 1.30) | .14 |
| Unknown | 24 | 0.56 (0.20 to 1.60) | .28 |
| White, non-Hispanic | 1,164 | Reference | |
| Cancer type (ICCC)[a] | | | |
| Leukemia | 290 | 16.12 (10.41 to 24.99) | <.001 |
| Lymphoma | 210 | 2.63 (1.74 to 3.97) | <.001 |
| CNS neoplasms | 275 | Reference | |
| Neuroblastoma and sympathetic nervous system tumors | 66 | 17.80 (8.02 to 39.55) | <.001 |
| Soft tissue sarcomas | 74 | 3.16 (1.79 to 5.60) | <.001 |
| Retinoblastoma | 23 | 0.68 (0.25 to 1.86) | .45 |
| Renal tumors | 37 | 21.27 (7.11 to 63.65) | <.001 |
| Hepatic tumors | 26 | 2.93 (1.25 to 6.85) | .01 |
| Bone tumors | 74 | 6.18 (3.41 to 11.18) | <.001 |
| Germ cell and gonadal tumors | 84 | 1.28 (0.69 to 2.36) | .43 |

| Characteristic | Sample Size | Odds Ratio (95% CI) | *P* |
|---|---|---|---|
| Epithelial tumors and melanoma | 193 | 0.16 (0.07 to 0.36) | <.001 |
| Other and unspecified | 4 | 1.87 (0.16 to 21.81) | .62 |
| Year of diagnosis | | | |
| 2012 | 220 | 1.76 (1.10 to 2.80) | .02 |
| 2013 | 209 | 1.68 (1.04 to 2.71) | .03 |
| 2014 | 212 | 1.34 (0.84 to 2.14) | .22 |
| 2015 | 230 | 0.84 (0.52 to 1.36) | .48 |
| 2016 | 234 | 1.02 (0.64 to 1.61) | .95 |
| 2017 | 252 | Reference | |
| Geographic area | | | |
| Appalachian county | 379 | Reference | |
| Non-Appalachian county | 978 | 0.56 (0.41 to 0.77) | <.001 |
| Reported from a COG-affiliated institution | | | |
| No | 161 | Reference | |
| Yes | 1,183 | 6.54 (3.58 to 11.94) | <.001 |
| Unknown | 13 | 3.34 (0.80 to 13.85) | .10 |

Abbreviation: ICCC, International Classification of Childhood Cancer.

[a]For patients with two or more cancers, only the first cancer diagnosis was described by cancer type.