NIH NATIONAL CANCER INSTITUTE

Childhood Cancer Data Initiative

Poster Abstracts

July 2019

Contents

Creating Meaningful Datasets for Clinical Care and Associated Research Progress

Poster Number	Title	Authors and Affiliations	Abstract
5	Gabriella Miller Kids First Data Resource Center: Collaborative platforms for accelerating cross-disease pediatric research across development and cancer	Allison P. Heath PhD[1], Christina Yung PhD[2], Yuankun Zhu BS[1], Michele Mattioni PhD[3] ,Zachary L. Flamig PhD[4], Yajing Tang MS[4], Bailey Farrow MS[1], Jena Lilly MS[1], Yiran Guo PhD[1], Pichai Raman PhD[1,5], Phillip B. Storm MD[1,9], Luca Graglia, MS[13], Samuel L. Volchenboum MD PhD[8, 13], Javad Nazarian PhD[6], Nicole Vasilevsky PhD[11], Jack DiGiovanna PhD[3], Melissa Haendel PhD[11,12], Robert L. Grossman PhD[4], Brandi Davis-Dusenbery PhD[3], Deanne M. Taylor PhD[5,7], Vincent Ferretti PhD[10], Adam Resnick PhD[1,5,9] [1]Center for Data-Driven Discovery (D3b) Children's Hospital of Philadelphia, Philadelphia, PA, USA; [2]Ontario Institute of Cancer Research, Toronto, Ontario, Canada; [3]Seven Bridges Genomics, Cambridge, MA, USA; [4]Center for Translational Data Science, The University of Chicago, Chicago, Illinois, USA; [5]Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA;; [6]Children's National Medical Center, Washington, DC, USA; [7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [8]Department of Pediatrics, Biological Sciences Division, The University of Chicago, Chicago, Illinois, USA; [9]Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA;	Pediatric cancer and structural birth defects represent diseases of childhood development proposed to share common genetic etiologies. The Gabriella Miller Kids First Data Resource Center (DRC) is a recently launched, multi-year initiative seeking to harness the potential of cloud-based collaborative research platforms supporting FAIR data principles. These platforms are brought together on the behalf of empowering integration, use, and clinical translation of diverse, multi-disease, large-scale genomic and clinical datasets. With a particular focus on family-based genomic studies with associated clinical and phenotypic data, more than 35 cross-disease pediatric cohorts, and associated biospecimens, have been defined, providing for a first-in kind network of leading researchers, sequencing facilities, and patient communities. By the end of 2019, genomic and clinical data from over 15,000 participant samples will be available across a variety of structural birth defects and pediatric cancer cohorts.The newly launched DRC provides for a secure, cloud-based platform that supports the ability of researchers to not only find, access, and reuse data, but also integrate, collaborate, and analyze data quickly and at scale. Recognizing the value in both original source data and the power of harmonized data for cross-disease analyses, the DRC supports the availability data across the entire data life cycle for all cohorts in its platforms. A "best-of-breed" approach has been taken by our multi-institutional team to develop a platform comprising reusable technology stack components and a portal entrypoint that provides comprehensive file browsing and dynamic cohort creation. Users have direct access to analytic workspaces and multi-user collaborative projects, such as Cavatica for bioinformatics workflow deployment and pedCBioPortal for cancer genomic visualizations. Additionally, the DRC has implemented a set of core services, powered by Gen3, that ensure a foundation for interoperability with other large-scale NIH data sources and the NCI Data Commons Framework. The allows the DRC to participate in the emerging ecosystem of analysis and visualization applications that provide a highly diverse pediatric research community secure access to large-scale federated via multi-environment user authentication. On the near term roadmap is the ability to create cohorts utilizing genomic features and improved visualizations for germline and family-based genomic data.

Poster Number	Title	Authors and Affiliations	Abstract
		[10]CHU Sainte-Justine Research Center, Montreal, Quebec, Canada; [11]Translational and Integrative Sciences Lab, Oregon Health & Science University, Portland, OR, USA;; [12]Oregon State University, Corvallis, Oregon, USA; [13]Center for Research Informatics, The University of Chicago, Chicago, Illinois, USA;	
42	Multi-Ethnic Survivors of Childhood Cancer (MESCA) Cohort: Development and Possibilities	Amie E. Hwang[1,2*], Richard Sposto[2], Chunqiao Luo[2], Dennis Deapen[1,2] [1]Los Angeles County Cancer Surveillance Program, 2Norris Comprehensive Cancer Center, University of Southern California	Background: Childhood cancer survivors continue to experience substantial risks of life-threatening late effects and premature mortality. Longitudinal hospital-based survivorship cohorts have shaped current understanding of late effects. But a research framework for population-based approach with racial/ethnic minorities and unbiased outcome assessment is lacking. We developed a population-based cohort of childhood cancer survivors of diverse race/ethnicity, and obtained severe medical outcomes using individual-level linkages with California Department of Public Health administrative datasets. This retrospective cohort, "Multi-ethnic Survivors of Childhood Cancer (MESCA)" will serve as an important data resource for childhood cancer survivorship research. Cohort Description: Childhood cancer survivors diagnosed before age 20 from 1988 to 2010 and who survived 5+ years were identified from the California Cancer Registry. Severe medical outcomes, defined as conditions that required hospitalization or led to premature mortality, from 1993 to 2016 were identified through individual-level linkage to hospital discharge records (Office of Statewide Health Planning and Development) and to vital statistics data (Center for Health Statistics). California's general population of the same age was used as non-cancer comparison group. 34,506 eligible childhood cancer survivors were identified, and hospitalization records of 10,342 cancer survivors, and 1,741 death records were obtained. There were 18,565,955 hospitalization records and 249,550 deaths identified from the non-cancer population. Preliminary Assessment: We used Poisson regression modeling to compare mortality in cancer survivors and non-cancer population with total person-years at risk accumulated by calendar year attained age within one-year intervals, race/ethnicity and sex. Racial/ethnic variation in all-cause mortality rates for childhood cancer survivors was distinctively different than non-cancer population especially for Asians

Poster Number	Title	Authors and Affiliations	Abstract
			and American Indians. Childhood cancer survivors were at substantially higher risk of premature death compared to non-cancer population at ages 5-15. Applicability/Potential: We demonstrate preliminary evidence for successful utilization of this dataset to further investigate the dependence of excess mortality risk in cancer survivors on sex or race/ethnicity and validate the significant excess risks of mortality for childhood cancer survivors especially during early years of survivorship. Model development for assessing severe medical outcomes that require hospitalization is ongoing and expect to offer significant contribution to the current understanding of childhood cancer survivorship.
25	The Utah Population Database: A unique resource for research across the lifespan of childhood, adolescent, and young adult cancer	Anne C. Kirchhoff,[1] Judy Ou,[2] Heydon Kaddas,[2] Joemy Ramsay,[2] Sapna Kaul,[3] Heidi Hanson[4] [1]University of Utah, Department of Pediatrics, Salt Lake City, UT [2]Huntsman Cancer Institute, Salt Lake City, UT [3]University of Texas Medical Branch, Preventive Medicine and Community Health, Galveston, TX [4]University of Utah, Department of Surgery, Salt Lake City, UT	Data Description: Population level data sources that encompass demographic, clinical, and genetic information across time are needed to identify risk factors and related outcomes of childhood, adolescent, and young adult (AYA) cancer throughout the life course. The Utah Population Database (UPDB) is one of the world's richest sources of epidemiologic data containing cancer diagnosis and treatment information from the Utah Cancer Registry linked to an individual's demographic and electronic healthcare records from statewide resources including: healthcare systems serving >85% of Utahans, insurance claims, state-wide vital statistics, hospitalization data, administrative records, and residential history. In addition, UPDB has pedigree information that spans 18 generations in some cases and can be used identify individuals belonging to families with a history of cancer. Applicability: Using UPDB data, we created a cohort of over 24,000 childhood, adolescent, and young adult (AYA) cancer patients diagnosed in Utah since 1986. In this cohort, there are 4,435 patients diagnosed ages 0-14 years and 20,155 patients diagnosed ages 15-39 years. This cohort can be accessed via approval from the Utah Resource for Genetic and Epidemiologic Research and the University of Utah Institutional Review Board. Potential for Broader Use: UPDB data has been linked to other rich data resources, including environmental, geographic, and Census data. The UPDB provides access to information on more than 8 million individuals and currently supports over 200 research projects. The research community could use this resource to answer novel questions regarding family cancer history, environmental risk factors for cancer, gene-environment interaction, long-term health effects of cancer treatment, and other topics pertaining to childhood and AYA cancers. Evidence of success: To date, our team has utilized the UPDB childhood and AYA cancer data to publish 13 manuscripts with an additional four in development on health care costs, utilization, and outcomes, and well as manuscripts on

Poster Number	Title	Authors and Affiliations	Abstract
			environmental exposures and health, among both patients and survivors of childhood and AYA cancer. Our studies often use patients' siblings or same age/sex individuals from the Utah population without a history of cancer as comparison samples. We have obtained both federal and foundation funding to support research with these data.
47	International Collaborations to Advance Etiologic Research on Leukemia and Other Cancers in Children	Catherine Metayer University of California Berkeley, School of Public Health, On behalf of the Childhood Leukemia International Consortium (CLIC), as Chair	Incidence rates of leukemia and other cancers in children have increased in the past 40 years in certain populations in the US and other countries, likely due to increased prevalence of environmental factors, possibly interacting with genetic predisposition. It is therefore critical to invest in childhood cancer etiologic research to identify preventive factors and promote research translation. However, funding is limited putting more children at risk to be diagnosed with cancer and to suffer treatment-related complications throughout their lives. Epidemiologic studies of childhood cancers are challenging due to their rarity and heterogeneity, thus necessitating international collaborations. To overcome the limitations of small individual studies, a group of international researchers established the Childhood Leukemia International Consortium (CLIC) in 2007. This collective of investigators from 21 countries share their expertise, resources, and data for ~40,000 leukemia cases and 350,000 controls from 40 case-control studies. This unprecedented large sample size allows conduct of pooled and meta-analyses for major leukemia subtypes, providing the most robust findings to date on pre- and postnatal exposures including pregnancy and birth characteristics, medical conditions, diet and vitamin supplementation, chemical use at home and at work, and measures of immune function and infections. CLIC analyses incorporate methodological evaluations of systematic errors inherent to case-control design by comparing interview-based vs. registry-based studies, and by conducting quantitative bias analyses. As CLIC studies increasingly generate genome-wide data, we are now entering an exciting new phase for genetic and gene-environment analyses. CLIC has shared governance and transparent procedures for membership, data sharing, scientific review, and authorship. CLIC's success is evidenced by our 15 publications and more than 20 ongoing pooled analyses, and also by the training of young investigators in the field of molecular and environmental epidemiology of childhood cancers. The CLIC consortium recently expanded to other pediatric cancers to form the CLIC+ consortium, which currently focuses on childhood brain tumors (~14,000 children) and non-CNS embryonal tumors (~18,000 children).

Poster Number	Title	Authors and Affiliations	Abstract
48	The Life Cancer Survivorship Research Database: A Resource for Studying Transitional Care Status and Other Health Outcomes Among Young Adult Survivors of Childhood Cancer	David R. Freyer1,2,3,4, Amie E. Hwang2,5, Yi Juin Tan1, Maureen Cairns2, Annie Chen1,3, Joel E. Milam2,5, and Kimberly A. Miller2,5,6 [1]Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA [2]USC Norris Comprehensive Cancer Center, Los Angeles CA Departments of [3]Pediatrics, [4]Medicine, [5]Preventive Medicine, and Dermatology[6], Keck School of Medicine, University of Southern California, Los Angeles, CA	Background Young adult survivors of childhood cancer (YASCC) are at risk for developing health problems caused by cancer treatment, necessitating transition of survivorship-focused care from pediatric to adult-focused facilities to support life-long surveillance and risk reduction strategies. Unfortunately, a large proportion of YASCC fail to complete transition, but little is known about factors influencing this disparity. The objective of this current study is to determine the efficacy of our transitional care model and the biomedical (host, disease and treatment) factors associated with transition failure. Description of Resource The LIFE Cancer Survivorship Research Database was established in 2009 to support survivorship research at Children's Hospital Los Angeles (CHLA). As of June 24, 2019, it comprised 1,655 survivors characterized for host demographics, disease characteristics, detailed cancer treatment exposures, current health problems with severity grade and attribution, and late effects risk classification (higher vs. lower, by program criteria). Higher-risk YASCC ≥ 21 years old are eligible to initiate transition from CHLA to our adult-focused survivorship clinic at USC Norris Comprehensive Cancer Center. Description of Cohort For this analysis, 742 higher-risk YASCC were transition-eligible. Of these, 408 (55.0%) were male; 431 (58.1%) were Latino. 389 (52.4%) had leukemia/lymphoma, 150 (20.2%) CNS tumors, 70 (9.4%) bone/soft tissue sarcoma, 40 (5.4%) neuroblastoma, and 81 (10.9%) other solid tumors. Age at diagnosis was < 16 years old for 656; attained age is ≥ 26 years for 430. 683 (92%) survivors were exposed to chemotherapy and 325 (43.8%) to irradiation. At initial survivorship visit, 400 had 1-5 established health problems, 138 had 6-10, 34 had 11-18, and 170 had none. Of 366 YASCC whose transition status has been determined, 216 (59.0%) initiated transition whereas 148 (40.4%) were lost to follow-up before that point. Preliminary Conclusions and Planned Analyses Pre-transition attrition of survivors is substantial and warrants upstream interventions to improve retention. For those YASCC who do initiate transition, this database study when completed will reveal what proportion actually complete transition in the context of our transitional care model and biomedical risk factors associated with transition failure, including demographics, diagnosis, treatment exposures, and health status at time of transition.

Poster Number	Title	Authors and Affiliations	Abstract
27	Scalable Consortium Approaches Facilitate Cooperative Research into Childhood Cancer Etiology and Outcomes	Jeremy M. Schraw[1], Olga A. Taylor[2,3,] Karen R. Rabin[2,3,4,] Philip J. Lupo[2,3,4,] and Michael E. Scheurer[2,3,5] 1. Department of Medicine, Baylor College of Medicine, Houston, TX USA 2. Department of Pediatrics, Baylor College of Medicine, Houston, TX USA 3. Texas Children's Cancer and Hematology Centers, Texas Children's Hospital, Houston, TX USA 4. Reducing Ethnic Disparities in Acute Leukemia Consortium 5. Adolescent and Childhood Cancer Epidemiology and Susceptibility Service for Texas	Rationale: Cancer is a leading cause of death by disease in children and adolescents/young adults (AYAs). Important gaps that limit research into childhood cancer etiology and outcomes include: 1) lack of population-based resources collecting both biological samples and epidemiologic questionnaire data on newly diagnosed cases; and 2) limited cohort studies of multi-ethnic populations collecting detailed clinical information and biological samples. We highlight two ongoing integrated consortia attempting to address these gaps. Applicability: The Adolescent and Childhood Cancer Epidemiology and Susceptibility Service for Texas (ACCESS-Texas) was established to facilitate rapid and complete ascertainment of childhood and AYA cancer cases, perform biobanking of multiple biospecimens from patients and family members, and systematically centralize and harmonize clinical and epidemiologic data. ACCESS-Texas recruits approximately 800 cases annually from all major pediatric cancer treatment centers in Texas and its infrastructure supports numerous research efforts including the Reducing Ethnic Disparities in Acute Leukemia (REDIAL) Consortium. REDIAL is focused on investigating the clinical and genetic features underlying disparities in outcomes in Latino children with acute leukemia. The REDIAL Consortium is anticipated to enroll a total of 1,000 children with acute leukemia, and will include genome-wide genotyping and collection of serial blood and bone marrow samples during therapy, linked to well-annotated clinical and epidemiologic data collected and managed through ACCESS-Texas. Potential for Broader Use: ACCESS-Texas and REDIAL are efficient and scalable mechanisms for accelerating research into childhood cancer etiology and outcomes. They integrate multiple data sources through a connected data infrastructure. Data and specimens collected by these consortia are a tremendous resource for childhood cancer researchers. This approach is broadly applicable to the study of childhood and adult cancer, as well as other rare diseases. Evidence of Success: Since their inception, ACCESS-Texas and REDIAL have collectively enrolled more than 2,000 patients and 1,500 family members, and provide centralized management of associated data and biospecimens. Importantly, these resources facilitate research into both etiology and outcomes within a single patient population. They support an array of research, including studies of ethnic differences

Poster Number	Title	Authors and Affiliations	Abstract
			in acute leukemia risk and outcomes; tumor and germline genomic sequencing; metabolomic profiling for biomarker discovery, and more.
36	NCI Provides Researchers with De-identified Cancer Imaging Data on The Cancer Imaging Archive (TCIA)	John Freymann,[1] Justin Kirby,[1] Fred Prior,[2] Keyvan Farahani,[3*] Janet Eary[3] [1]Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Frederick, MD [2]Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR [3]Cancer Imaging Program, National Cancer Institute, Bethesda, MD	The Cancer Imaging Archive (TCIA), supported by the Cancer Imaging Program at the National Cancer Institute, is a persistent information repository that increases public availability of high-quality cancer imaging data, supports NIH data sharing requirements, facilitates reproducibility of research, and creates a culture of open data sharing among collaborators. In support of the Childhood Cancer Data Initiative, TCIA welcomes proposals for submission of childhood, adolescent, and young adult cancer imaging datasets. TCIA (www.cancerimagingarchive.net) hosts de-identified publicly accessible cancer imaging data and clinical information related to imaging studies and trials1. TCIA provides services that enable submission, management and retrieval of data based on the FAIR principle2. TCIA data collections are managed as data publications, with each assigned a Digital Object Identifier (DOI) and citation3. Each collection includes a descriptive summary, licensing details, multiple search and download interfaces, citations, version control, and links to supporting data. TCIA data providers are encouraged to submit data descriptor manuscripts to peer-reviewed journals such as Nature Scientific Data. TCIA also supports submission of histology image sets, third-party analysis results such as tumor segmentations and other image annotations, radiomic and pathomic features including advanced analyses such as tumor-infiltrating lymphocyte maps4. Data submission to TCIA begins with a request to NCI which is reviewed by the TCIA governance committee. Once the request is approved a TCIA submissions team works with the submitting site and implements a standards-based, IRB approved process to ensure accurate and timely acquisition. The data curation processes from submitting sites is also reviewed for compliance with privacy, confidentiality and human subjects' regulations. TCIA is visited by over 15,000 users monthly from more than 125 countries that download approximately 75 TB of data per month. Over 700 peer-reviewed publications have utilized TCIA data. A help desk provides email and phone support for both data submitters and researchers who download and use TCIA data. TCIA currently hosts 92 collections derived from major research initiatives including the Cancer Genome Atlas (TCGA), Clinical Proteomic Tumor Analysis Consortium (CPTAC), and the Cancer Moonshot and NCI funded clinical trials as well as independent research initiatives. The pediatric oncology research community is

Poster Number	Title	Authors and Affiliations	Abstract
			encouraged to submit imaging data to TCIA for access by the broader research community.
45	Center for Pediatric Tumor Cell Atlas	Kai Tan and Stephen Hunger Children's Hospital of Philadelphia University of Pennsylvania University of California San Diego	Childhood Cancer is the second most common causes of childhood deaths. Cancers in children are highly distinct from those in adults. Causes, mechanisms and therapeutic approaches cannot be extrapolated from the study of adult malignancies. To impact the burden of pediatric cancers requires the specific study of pediatric cancers. We therefore will conduct in depth characterization of three specific subtypes of pediatric malignancies which, together, account for over 50% of all pediatric cancer deaths: high grade glioma (pHGG), high risk neuroblastoma (NB), and very high risk acute lymphoblastic B-cell precursor leukemia (VHR-ALL). All three tumor types share a characteristic of typically initial response to therapy, followed by the emergence of resistance and refractory disease. We will map molecular and cellular changes in tumor cells, microenvironment and the immune system using comprehensive multi-dimensional single-cell and in situ technologies associated with two critical transitions: initial response, and emergence of resistant disease – both high-priority transitions. Treatment modalities will include standard chemotherapy, molecular targeted therapies and chimeric antigen receptor (CAR-) T cell therapy, capturing the cutting edge of current cancer therapy. The final product will impact global childhood cancer research in the following manner: 1) Tumor Atlases: atlases will provide a highly user friendly, publicly available, searchable database of the most comprehensive multiomic, single cell analysis of the three most lethal childhood cancers. Molecular data will be richly annotated with access to additional pathological evaluation, clinical and outcomes data, and grant access to source data such tumor imaging. 2) Computational methods: in addition to the data, the critical computational tools and pipelines used in this project will be available to the research community. These include methods and pipelines for processing multi-omics and in situ data, inference of cancer phylogeny, inference of cell- and clone-specific pathways, methods for inferring cell-cell crosstalk, as well as database algorithms for the query, exploration and visualization of highly complex data. 3) Access to biospecimens for follow up studies: biospecimens collected in this project will be banked and made available to the biomedical research community. These include tissue sections, viably frozen specimen, and patient derived xenograft (PDX models). In summary, the CPTCA project will seek to address a major public health need in pediatrics, broadly impact the entire research community, and jumpstart target discovery based on a sophisticated understanding of the key molecular circuits that drive pediatric cancer.

Poster Number	Title	Authors and Affiliations	Abstract
44	My Pediatric and Adult Rare Tumor Network (MyPART): a Cancer Moonshot™ funded patient engagement network for pediatric, adolescent, and young adult rare solid tumors	Karlyne M. Reilly, PhD[1]; Abby Sandler, PhD[1]; Taryn Allen, PhD[1]; Donna Bernstein[1]; Jaydira del Rivero, MD[1]; Sarah Fuller[1]; John Glod, MD, PhD[1]; Maran Ilanchezhian[1]; Jason Levine, MD[2]; Staci Martin, PhD[1]; Margarita Raygada, PhD[1]; Nurlanbek Shonkoev, M.Eng[2]; BJ Thomas[1]; Lori Wiener, PhD[1]; and Brigitte C. Widemann, MD[1] [1]Pediatric Oncology Branch and Rare Tumor Initiative, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892 [2]Office of the Director, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892	The Center for Cancer Research (CCR) in the NCI Intramural Research Program has launched a Cancer Moonshot™ funded initiative focused on pediatric and young adult rare solid tumors. The goals of MyPART (My Pediatric and Adult Rare Tumor Network) are to: 1) develop a network of clinical/research sites and a patient and advocacy interface, 2) provide state-of-the-art expertise, health care, and data to patients with rare tumors, and 3) build databases and tools to advance research on new treatments for these tumors. We developed a natural history and biospecimen acquisition study for children and young adults with rare solid tumors (NCT03739827), which collects standardized medical and family history, patient-reported outcomes, and biospecimens. Patients can participate remotely or be evaluated at the NIH Clinical Center (NIHCC). Patients are followed longitudinally including recording of response to treatments such as standard of care, clinical trials, or other interventions received outside of the NIHCC. Patient data are collected in CCR's electronic research information management system, Labmatrix, by three methods: 1) direct data entry by clinical/research staff, 2) electronic transfer from the patient's NIHCC electronic medical record using the Biomedical Translation Research Information System (BTRIS), and 3) electronic transfer of patient-entered data via Scribe. We are developing ways for data to be batched, anonymized, and exported to an appropriate public repository, such as the Cancer Research Data Commons, to facilitate research on rare tumors worldwide. We are also developing ways to harmonize our data with other ongoing data collection efforts for pediatric and other rare solid tumors. MyPART recently held the first pediatric-young adult chordoma clinic at the NIHCC in April that brought together 7 chordoma patients and their families from 3 countries with 9 extramural experts in chordoma care and the Chordoma Foundation. Additionally, 4 patients participated remotely. Data and biospecimens from this clinic are currently being analyzed. This clinic approach for very rare tumors allows patients to 1) get expert clinical advice, 2) meet with other patients with the same rare tumor, and 3) share issues of importance related to their specific rare tumor.
66	Childhood Cancer Data Collection: A Trend Analysis of Timeliness from NPCR-ECC (October	Kevin Zhang[1], Jon Stanger[1], Olga Galin[1], Yuan Ren[1] Toye Williams[2], Reda Wilson[2], Vicki Bernard[2] [1]ICF, Fairfax, VA	Introduction Compared with cancers diagnosed in adults, cancer incidence in children and young adults (aged 19 years and younger) is less common. Therefore, a focused approach is needed to obtain timely report and sufficient data to support scientific research and public health surveillance. CDC's Early Case Capture (ECC) of Pediatric and Young

Poster Number	Title	Authors and Affiliations	Abstract
	2012- April 2019 Submissions)	[2]Centers for Disease Control and Prevention, Atlanta, GA	Adult Cancers (PYAC) program was created to address this issue. Built on the existing National Program of Cancer Registries – Cancer Surveillance System (NPCR-CSS), the ECC project captures state surveillance data on childhood cancers from the latest available year, sometimes within 30 days of diagnosis for specific sites. Purpose This study examines trends in the timeliness of childhood cancer incidence data from the ECC system, which began submitting data in October 2012 through April 2019. Methods Measures of timeliness are an essential aspect of the ECC project and must be calculated in a similar manner across ECC registries. Three timeliness measures have been applied by the ECC registry and the results have been reported to CDC during each ECC data submission. These measures assess improvements in reporting timeliness and data availability for use by a researcher. They address the intent of the Caroline Pryce Walker Conquer Childhood Cancer Act and have been shared with Congress to demonstrate progress. These three measures are: 1) Timeliness of first source record; 2) Timeliness of reporting a case to the central registry; and 3) Timeliness of data availability for use by a researcher. For all the three measures, cases with an unknown day of first contact or day of diagnosis have been excluded from the calculation. Results and Conclusion The changes in timeliness over time may suggest the overall improvements in data collection among the participating states. Areas for improvement will also be revealed.
49	Understanding pediatric osteosarcoma genetic susceptibility through large multi-institutional collaborative studies	Lisa Mirabello[1], Bin Zhu[1], Roelof Koster[1,] Eric Karlins[2,] Michael Dean[1,2,] Meredith Yeager[2], Matthew Gianferante[1,] Logan Spector[4], Lindsay Morton[1], Leslie L. Robison[3,] Gregory T. Armstrong[3,] Smita Bhatia[5], Lei Song[1], Nathan Pankratz[4,] Maisa Pinheiro[1], NCI DCEG Cancer Genomics Research Laboratory [2,6,] Julie M. Gastier-Foster[7,] Richard Gorlick[8,] Silvia Regina Caminada de Toledo[9,] Antonio S. Petrilli[9], Ana Patiño-Garcia[10,] Fernando Lecanda[10,] Massimo Serra11, Claudia Hattinger[11,] Piero Picci[11], Katia Scotlandi[11,] Adrienne M. Flanagan[12,13,] Roberto Tirabosco[13,] Maria Fernanda Amary[13,] Nilgün Kurucu[14,] Inci Ergurhan Ilhan[14,]	Osteosarcoma is the most common malignant bone tumor in children and adolescents, although it is a rare cancer with only approximately 400 new pediatric cases diagnosed per year in the United States. Survival rates and treatments for osteosarcoma, and several other pediatric cancers, have changed little over the past three decades. Studies of its germline genetic profile, in combination with the somatic landscape, inform our understanding of tumorigenesis and new insights may lead to improved clinical outcomes. We perform large, multi-institutional collaborative genomic studies to identify novel germline variants predisposing to osteosarcoma and associated with clinical outcomes. We have assembled approximately 2,000 osteosarcoma cases with germline DNA samples, and conducted genotyping using genome-wide SNP microarrays in 1,990 osteosarcoma cases and whole-exome sequencing in 1,244 cases to investigate the underlying germline genetic architecture of osteosarcoma. Our exome sequencing data were compared

Poster Number	Title	Authors and Affiliations	Abstract
		Neriman Sari[14], Mandy L. Ballinger[15,] David M. Thomas[15,] Donald A. Barkauskas[16,] Neyssa Marina[17], Gerardo Mejia-Baltodano[18,] Patricia Valverde[19,] Danielle Karyadi[1], Belynda D. Hicks[2,] Bin Zhu[2], Mingyi Wang[2], Kathleen Wyatt[2], Amy A. Hutchinson[2], Casey Dagnall[2], Margaret Tucker[1], Joshua Sampson[1,] Maria T. Landi[1,] Neal D. Freedman[1], Susan Gapstur[20,] Robert N. Hoover[1,] Stephen J. Chanock[1,21,] Sharon A. Savage[1,21] [1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [2] Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Frederick, MD 20877, USA; [3] Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN; [4] Department of Pediatrics, University of Minnesota Minneapolis, MN, 55455, USA; [5] Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, AL; [6] The NCI DCEG Cancer Genomics Research Laboratory includes Sara Bass, Joseph F. Boland, Laurie Burdett, Salma Chowdhury, Michael Cullen, Herbert Higson, Kristine Jones, , Hyo Jung Lee, Wen Luo, Michael Malasky, Michelle Manning, Adri O'Neil, David Roberson, Shalabh Suman, and Aurelie Vogt; [7] Nationwide Children's Hospital, and The Ohio State University Department of Pathology and Pediatrics, 700 Children's Dr., C0988A, Columbus, OH 43205, USA;	with both 1,062 cancer-free NCI-DCEG controls whole-exome sequenced simultaneously and 27,173 ExAC non-Finnish European (NFE) cancer-free controls to investigate the frequency of germline pathogenic mutations in high interest cancer-susceptibility genes (CSG). We observed a higher predicted pathogenic CSG variant burden in the cases compared to both the DCEG controls (P=1.3x10^{-18}) and ExAC NFE controls (P=2.3x10^{-53}) in an analysis restricted to European ancestry (732 cases). Notably, a pathogenic CSG variant was identified in 28% of osteosarcoma cases of which 19.6% mapped to an autosomal dominant gene or a known osteosarcoma-associated cancer predisposition syndrome gene, whereas the frequencies in the control sets were each approximately 5%. We also observed a higher than expected frequency of pathogenic variants in genes not previously linked to osteosarcoma. Our data establishes that approximately one fourth of individuals with osteosarcoma, unselected for family history, may harbor a highly penetrant germline mutation necessitating follow-up studies and possible genetic counseling with cascade testing. Our data underscore the need for collaborative efforts and data sharing to enable large studies of rare pediatric cancers.

Poster Number	Title	Authors and Affiliations	Abstract
		[8] University of Texas MD Anderson Cancer Center, Houston, TX, USA; [9] Pediatric Oncology Institute, GRAACC/UNIFESP, Rua Botucatu, 743 8º andar Laboratorio de Genética, CEP 04023-061 São Paulo SP Brazil; [10] Department of Pediatrics, University Clinic of Navarra, Universidad de Navarra, Pío XII 36, 31080 Pamplona, Spain; [11] Laboratory of Experimental Oncology, Orthopaedic Rizzoli Institute, Bologna, Italy; [12] UCL Cancer Institute, Huntley Street, London WC1E 6BT, UK; [13] Royal National Orthopaedic Hospital NHS Trust, Stanmore, Middlesex HA7 4LP, UK; [14] A.Y. Ankara Oncology Training and Research Hospital, Department of Pediatric Oncology, Yenimahalle- Ankara, Turkey; [15] The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Darlinghurst, NSW, Australia; [16] Department of Preventive Medicine, Keck School of Medicine of the University of Southern California, Los Angeles, CA, 90032; [17] Stanford University & Lucile Packard Children's Hospital, 1000 Welch Rd, Suite 300 Palo Alto, CA 94304, USA; [18] Hospital Infantil Manuel De Jesus Rivera, Managua, Nicaragua; [19] Unidad Nacional de Oncología Pediátrica, Guatemala City, Guatemala; [20] Epidemiology Research Program, American Cancer Society, Atlanta, GA; [21] contributed equally.	
28	Up-front collection of critical epidemiologic	Logan G. Spector, Todd Alonzo, Negar Fallahazad, Philip Lupo, Peter Adamson	Description of Need and Relevant Methodology: Though a thorough understanding of etiology is necessary to inform prevention efforts, the causes of childhood cancer remain mostly obscure. However, the nearly forty years of research into childhood

Poster Number	Title	Authors and Affiliations	Abstract
	variables in Project:EveryChild		cancer causation have identified a number of critical epidemiologic variables of interest across tumor types. Thus we established an up-front collection of critical variables in Project:EveryChild (PEC). Applicability to Childhood/AYA Cancers: As the combined research registry and biobank of the Children's Oncology Group, PEC serves as the major platform for non-therapeutic correlative or biologic investigations of pediatric cancer in North America. Using information collected up-front from patients and parents investigators may now directly identify potential subjects for further in-depth study. Potential for Broader Use: Variables collected by PEC include demographics (parental age, detailed and specific race/ethnicity, place of birth), parental address, conception by in vitro fertilization, family history of cancer in first degree relatives, and presence of genetic syndromes, structural birth defects, and autoimmune disease; twinship and presumed zygosity is also collected to enable twin studies. PEC registrants may indicate permission for future contact at the time of enrollment, thus researchers have the ability to validate and expand upon the preliminary data available up-front. Evidence of Success: The current phase of PEC was activated on 8/21/2017 with groupwide adoption of the protocol required by 12/8/2017. From 8/21/2017 through 6/10/2019, PEC has enrolled 11,304 patients with consent to recontact; 9,806 (86.7%) were newly diagnosed. While PEC has not yet achieved population-wide enrollment, preliminary analyses suggest the prevalence of these critical epidemiologic variables are consistent with previous reports. For instance, the male excess of pediatric cancer is apparent, with 6,399 (56.6%) of patients being male. Additionally, 518 (4.6%) patients have a first-degree relative with history of cancer, 381 (3.4%) have a known genetic syndrome, 362 (3.2%) have a structural birth defect, 253 (2.2%) were twins or triplets, 148 (1.3%) were conceived by IVF, and 135 (1.2%) were diagnosed with autoimmune disease. These data will be valuable for future epidemiologic studies of childhood cancer seeking to better characterize the role of these variables on susceptibility.
43	Integrating genomic and clinical data for pediatric cancer across California	Michelle Turski, Carlos Espinoza, Courtney Onodera, Jessica Van Ziffle, Marcus Breese, Atul Butte, Eugenia Rutenberg, Gundolf Shenck, Atul Butte, Adam Olshen, Boris Bastian, E. Alejandro Sweet-Cordero.	Background: Many pediatric cancer subtypes are rare and thus hard to study without robust data sharing. Methods: The University of California San Francisco has developed an active and robust cancer sequencing program revolving around an in-house assay of almost 500

Poster Number	Title	Authors and Affiliations	Abstract
			genes (UCSF500). This assay analyzes single-nucleotide variants, copy number alterations and fusions in pediatric and adult cancers. Importantly, the assay also analyzes germline DNA to identify genetic predispositions. To date over 600 pediatric/AYA patients have been sequenced. The UCSF500 pediatric cohort is unusual in that almost half represent central nervous system (CNS) tumors with the remaining half consisting of both non-CNS solid tumors and hematological malignancies. Approximately 90% of cases were clinically informative, as defined by detection of either pathogenic or likely pathogenic somatic or germline alterations. Approximately 11% were found to have a genetic predisposition. We will present additional details on this cohort including germline findings and therapeutically-actionable variants. We will also describe internal efforts to establish a data-sharing platform for all UCSF500 data using the cBioPortal interface. A customized cBioPortal interface is already available and actively utilized to share data on this cohort. Importantly, while most of the cases sequenced using the UCSF500 are from UCSF patients, we also sequence a significant number of patients from outside referring institutions. cBioPortal provides an interface for data sharing with these external referral centers. To annotate this genomic data with clinical information, we are using the Information Commons (IC), a fast and easily searchable repository of all UCSF clinical data extracted from the electronic health record (EHR). IC was developed by the UCSF Bakar Institute for Computational Health Sciences led by Dr. Atul Butte. IC has the ability to extract data from clinical records including clinic notes and radiological images in a structured and de-identified manner, store these data in a queryable database, and provide different tools to query the database. IC represents a powerful resource in providing the clinical data elements needs to contextualize genomic data to provide new clinical and scientific insights. We will describe our ongoing efforts to integrate genomic data already shared in ciboportal with an enriched clinical annotation provided by the IC. Finally, we will briefly share ongoing efforts to build a genomic database that will include both pediatric and young adult patients from five different University of California institutions to be made available for research purposes. Conclusions: We have sequenced over 600 pediatric cancer cases across a wide variety of cancers and are developing new approaches to integration with the

Poster Number	Title	Authors and Affiliations	Abstract
			medical record. This dataset is unique as it is enriched for patients with the wide ethnic, racial and national origin diversity seen in the state of California.
41	Optimizing Pediatric Cancer Sequencing Data Generated in Clinical Laboratories Through a Collaborative Annotation Initiative of the AACR Genomic Evidence Neoplasia Information Exchange (GENIE) Project	Neerav Shukla[1], Eva Lepisto[2], Abigail Ward[2], Stephanie Suser[1], Nancy Bouvier[1], Seth Sheffler-Collins[3], Marilyn M Li[4], Sarah K Tasian[4], Alejandro Sweet-Cordero[5], Deborah Schrag[2], Andrew Kung[1], Katherine A. Janeway[2], and the AACR Project GENIE Consortium. [1]Memorial Sloan Kettering Cancer Center [2]Dana-Farber Cancer Institute [3]American Academy of Cancer Research [4]Children's Hospital of Philadelphia [5]University of California, San Francisco	Background: All pediatric malignancies are considered rare or ultra-rare diseases which necessitates data sharing to identify recurrent genomic alterations, key clinical prognostic factors, and patterns of treatment response and resistance. As sequencing becomes more common in clinical practice, maximizing the utility of data generated in individual clinical laboratories through data sharing and established annotation practices is essential. Methods: Project GENIE is an international data registry linking genomic and clinical data collected during the routine care of cancer patients treated at 19 participating centers. GENIE data are shared publicly after 6 months of embargoed use by consortium members. We report contribution of sequencing data from pediatric patients in the GENIE repository and associated characteristics. The age variable currently available for analysis (and used in the results section below) is age at the time of sequencing. Results: 1,487 pediatric cancer samples from 1,410 individuals age ≤18 years are currently included in the publicly-available GENIE dataset. The most common pediatric cancer types are (n=samples/individuals): neuroblastoma (n=192/174), osteosarcoma (n=68/63), B-ALL (n=63/63), medulloblastoma (n=53/53), and Ewing sarcoma (n=49/48). For these 5 most common pediatric cancer types, an additional 225 samples from 224 individuals age >18 years are present in the dataset. The most commonly altered genes identified in cancer specimens sequenced at age ≤25 were *TP53, KMT2D, NF1, BRAF, CDKN2A/B, MYCN, CDKN1B,* and *CCND3*. Conclusions: We report a publicly-available dataset which will include 2,200 pediatric cancer specimens from 2,100 individuals age ≤18 years with the next data release planned for July, 2019. As pediatric cancer types are also sequenced when individuals are >18 years, the total number of pediatric cancer cases reported here is an underestimate. More complex case selection criteria are being applied to ascertain pediatric cancer cases. A pilot project to annotate genomic data in GENIE with treatment outcome data is underway. By enlisting data scientists and disease-specific pediatric oncology experts to establish shared annotation methodologies, we aim to

Poster Number	Title	Authors and Affiliations	Abstract
			provide a cohesive genomic database linked to relevant clinical factors across the spectrum of childhood cancer, leveraging the infrastructure of Project GENIE as an emerging standard for aggregation of cancer genomic data.
18	Deep Proteogenomic Survey across Seven Childhood Brain Tumors: A CBTTC/CPTAC Pilot	Pei Wang, Nicole Tignor, Steven Gygi, Richard G. Ivey, Amanda Paulovich, Alexey Nesciżhskii, Sanjukta Guha Thakurta, Jeffery R. Whiteaker, Jacob J. Kennedy, Uliana J. Voytovich, Li Ding, Liang Bo Wang, Pichai Raman, Yuankun Zhu, Tara Hiltke, Henry Rodriguez, Brian R. Rood, and Adam Resnick on behalf of the NCI's Clinical Proteomic Tumor Analysis Consortium (CPTAC) and the Children's Brain Tumor Tissue Consortium (CBTTC) The Children's Hospital of Philadelphia, Children's National Medical Center, Icahn School of Medicine at Mount Sinai, Fred Hutchinson Cancer Research Center, Harvard Medical School, Washington University School of Medicine	Genomic characterization has allowed for the differentiation of different tumor types based upon the abundance of gene transcripts. However, owing to the many layers of regulation between transcript and the post-translationally modified protein, it has been challenging to extrapolate biology from transcriptional differences alone. Working with the NCI's Clinical Proteomic Tumor Analysis Consortium (CPTAC) and a consortium of more 18 partnered institutions comprising the Children's Brain Tumor Tissue Consortium (CBTTC), we hypothesized that a comparative analysis of the proteome and phosphoproteome across a pilot initiative of 7 childhood brain tumors would yield a deeper understanding of differences in functional biology in ways that could be harnessed for clinical translation. We performed tandem mass tag labeling and triple mass spectrometry of 226 fresh frozen tumor samples representing histologic diagnoses of: high grade astrocytoma (27), low grade astrocytoma (97), ganglioglioma (20), ependymoma (32), medulloblastoma (22), atypical teratoid rhabdoid tumor (12), and craniopharyngioma (16). Among these samples were 22 pairs from pre/post recurrence. Across this sample set, we quantified 9155 proteins and 13632 phospho sites. After data preprocessing, consensus clustering showed that protein profiles effectively distinguish major histology types. Additional regression-based analyses revealed groups of proteins and pathways showing distinct activities across histologies and clinical features/outcomes. Further leveraging tumor/normal WGS, RNAseq data and longitudinal clinical data collected under the CBTTC, the functional impact of mutations and fusions was assessed, focusing initially on low grade glioma and the proteomic ramifications of BRAF V600E, BRAF-fusion and wild type BRAF. Moreover, from the primary and recurrent tumor pairs, the upregulation of proteins associated with immune evasion was also identified in more advanced LGG tumors. Building on this study, the consortia are further expanding studies across the age continuum, beginning with an integrated prospective and retrospective cohort of pediatric, adolescent and young adult gliomas, and adult glioblastoma multiforme (GBM). These efforts, supported by the emerging network of interoperable data-sharing and collaborative cloud-based NIH/NCI platforms provide for a first-in-kind multiomic analysis that includes the proteomic dimension alongside large scale genomic data, providing new functional insights that are poised to accelerate clinical translation.

Poster Number	Title	Authors and Affiliations	Abstract
17	Policies and procedures for data use in the Pediatric Proton and Photon Consortium Registry (PPCR)	Sara L. Gallotto[1], Benjamin V. Bajaj[1], Miranda P. Lawell[1], Daniel Indelicato[2], Arnold Paulino[3], William Hartsell[4], Nadia Laack[5], Ralph Ermoian[6], John Perentesis[7], Stephanie Perkins[8], Victor Mangona[9], Christine Hill-Kayser[10], Suzanne Wolden[11], Young Kwok[12], Michael Confer[13], J. Ben Wilkinson[14], Torunn I. Yock[1] [1]Department of Radiation Oncology, Harvard Medical School, [2]Department of Radiation Oncology, University of Florida, [3]Department of Radiation Oncology, MD Anderson Cancer Center, [4]Department of Radiation Oncology, Northwestern Medicine Chicago Proton Center, [5]Department of Radiation Oncology, Mayo Clinic, [6]Department of Radiation Oncology, University of Washington, [7]Division of Oncology, Cincinnati Children's Hospital Medical Center, [8]Department of Radiation Oncology, Washington University in St. Louis, [9]Department of Radiation Oncology, Texas Center for Proton Therapy, [10]Department of Radiation Oncology, University of Pennsylvania, [11]Department of Radiation Oncology, ProCure Proton Therapy Center, New Jersey, [12]Department of Radiation Oncology, University of Maryland, [13]Department of Radiation Oncology, ProCure Proton Therapy Center, Oklahoma City, [14]Department of Radiation Oncology, Provision Healthcare	Background/Objectives: The Pediatric Proton and Photon Consortium Registry (PPCR) opened to enrollment in 2012 with the primary goal of expediting outcomes research in the population of pediatric patients requiring radiotherapy and provide a platform and community for collaborative research. Design/Methods: The PPCR allows each Consortium member full access to their own institution's data, with users capable of selecting, building, and exporting their own reports and data for internal use. Data exports from the aggregate data of all the participating institutions are available to consortium members and outside investigators through the PPCR-specific "Request for Data" (RFD) form which is processed by the data management team at MGH after site PI's consent to allow their institutional data to be used for this purpose. Data is de-identified and distributed to the investigator within 15-30 days of the request. The MGH PPCR team assists in the creation of routine electronic data imports from other hospital-wide databases, such as the Radiation Oncology Charts (MOSAIQ and ARIA) and other electronic medical record (EMR) systems at the consortium institutions. Data collection field updating occurs regularly as the field of cancer evolves, a process made very simple by use of the non-proprietary REDCap platform. Results: The PPCR has published over 40 abstracts, papers, and manuscripts. With the creation of the REDCap RFD form, projects are tracked from hypothesis to publication. Electronic connections to other databases reduce hands-on time for routine tasks such as patient record creation and data validation. The PPCR can accept patient information as entered within REDCap or in bulk through imports coordinated by the MGH-based data management team. Such techniques have assisted in entering data backlog and database restructuring and recoding with little disruption or downtime to users. Conclusion: The PPCR is an extensive consortium of pediatric radiation oncologists throughout the United States. By using a non-proprietary web-based data storage platform and policies of free and transparent data access for all members to view their institution's study records, the PPCR has positioned itself as a standard for pediatric cancer registries capable of more expansion as it continues to develop and evolve.
50	The Long Tail of Diagnoses in a Cohort of Pediatric Patients with Solid	Suzanne J. Forrest[1], Abigail Ward[1], Sanda Alexandrescu[2], Keith L. Ligon[2], Pratiti Bandopadhayay[1], Alyaa Al-Ibraheemi[2], Alanna Church[2], Katherine A. Janeway[1]	<u>Background:</u> Research discovery sequencing studies in pediatric malignancies utilize a particular cancer diagnosis as their starting point. In contrast, basket trials and clinical sequencing studies allow patients with multiple diagnoses to enroll. One important

Poster Number	Title	Authors and Affiliations	Abstract
	Tumors in a Clinical Sequencing Study	[1]Pediatric Oncology, Dana-Farber / Boston Children's Cancer and Blood Disorders Center, Boston, MA [2]Department of Pathology, Boston Children's Hospital, Boston, MA	objective of such studies is to describe the spectrum of mutations seen in the cancers of the study participants. To facilitate such analyses, patient diagnoses need to be systematically classified. Methods: All patients seen at Dana-Farber/Boston Children's with a suspected neoplasm were offered participation in the PROFILE Cancer Research Study since 2013. Tumor samples were submitted for sequencing with OncoPanel, a next-generation targeted DNA sequencing panel, in which exons of 300-447 cancer genes are sequenced for variants and copy number alterations, and introns of 35-60 genes are sequenced for fusions. Patient diagnosis was determined by committee review of the pathology report from the specimen that underwent sequencing using the International Classification of Diseases for Oncology (ICD-O) coding system, Version 3.1. Results: At the data cut-off date for this analysis (8/1/17), 953 patients with solid tumors had enrolled in the PROFILE Study. Within the cohort, 75% (n=719) had adequate leftover material sent for sequencing. For the 719 patients with tumor sample submitted, 680 (94%) had a successful sequencing result. In the patients with successful sequencing, 50% (338/679) had extracranial solid tumors, 45% (n=308/679) had intracranial solid tumors, and 5% (33/679) were excluded based on having benign lesions or vascular malformations. 57% (370/646) had one of ten common pediatric cancer diagnoses (neuroblastoma, low-grade glioma, medulloblastoma, Wilms tumor, pilocytic astrocytoma, ependymoma, osteosarcoma, glioblastoma, rhabdomyosarcoma, and Ewing sarcoma). The remaining 43% (n=276) had a long tail of rare diagnoses with less than 20 cases for each diagnosis. Conclusion: The pediatric solid tumor patient population is characterized by a collection of rare diagnoses. Understanding the mutational spectrum in these patients requires international sharing of sequencing data. To facilitate data sharing in rare pediatric cancers, a standard ontology system should be adopted for classification. For our patient population, we used the ICD-O coding system. Importantly, the ICD-O code can be mapped to other ontology systems including the NCI Thesaurus and the International Classification of Childhood Cancer.

Poster Number	Title	Authors and Affiliations	Abstract
29	None specified	Timothy Triche	While major advances in the understanding and treatment of adult cancer have eventuated from large scale cancer genomic profiling initiatives like TCGA, there is no such broad genomic database for cancer in the young. At Children's Hospital Los Angeles, part of the NCI supported USC Norris Comprehensive Cancer, we seek to redress that issue across all types of cancer in the young. We created a childhood cancer genomic profiling panel, OncoKids (aka Oncomine Childhood Cancer Research Assay, OCCRA), modeled after the NCI MATCH program Oncomine Comprehensive Assay (OCA). Both panels use FFPE or fresh tumor tissue, interrogate both DNA and RNA, and were developed in collaboration with Thermo Fisher. In contrast to OCA, OncoKids incorporates all known DNA mutations, amplified genes, gene fusions, and over-expressed genes of clinical, diagnostic, prognostic, or therapeutic importance. Over 400 cases of liquid, solid, and brain tumors from multiple institutions were used to validate the CAP/CLIA certified LDT. In use for over a year at CHLA, essentially every patient (~600 per year) is profiled prior to treatment. Results are discussed bi-weekly in a Molecular Pathology conference attended by oncologists and pathologists. All accrued data, as well as all publicly available genomic data, are archived in a growing genomic database of childhood cancer, Childhood Cancer Knowledge Base (CCKB), hosted at CHLA and available throughout the world. A comprehensive suite of analytic tools for data mining suitable for both clinical and research analysis has been developed to analyze the data which are available to anyone who elects to join the international collaborative network ICON (International Childhood Oncology Network), sponsored by Thermo Fisher. Over 30 institutions worldwide have joined ICON and will have access to the data, analytic tools, and expertise hosted at CHLA. Over 65% of patients have clinically relevant genomic defects, 15% show evidence of germline inheritance, and in several instance, critically important treatment decisions have been based on the results, available in emergent circumstances within 48 hours. Ultimately, the goal is to create an international collaborative consortium that will ensure that children with cancer will receive optimal treatment and outcomes, regardless of their geographic location.
7	Pediatric Proton/Photon Consortium Registry is a Multicenter Data Repository for Pediatric Cancer	Torunn I. Yock[1], Benjamin V. Bajaj[1], Miranda P. Lawell[1], Daniel Indelicato[2], Arnold Paulino[3], Susan McGovern[3], William Hartsell[4], Nadia Laack[5], Anita Mahajan[5], Ralph Ermoian[6], John Perentisis[7], Ralph Vatner[8], Luke Pater[8], Stephanie Perkins[9], Victor Mangona[10], Christine Hill-Kayser[11], Suzanne Wolden[12], Young Kwok[13], Michael Confer[14], J. Ben	Background/objectives: The Pediatric Proton and Photon Consortium Registry (PPCR) was established in 2010 to expedite proton outcomes research in the pediatric population requiring radiotherapy. We are introducing the PPCR to the Childhood Cancer Data Initiative (CCDI) and providing an overview of the data available for further study and collaboration. Design/methods: A multi-institutional registry of clinical, dosimetric, radiographic, and patient-reported data for pediatric patients undergoing radiation therapy was

Poster Number	Title	Authors and Affiliations	Abstract
	Patients Requiring Radiotherapy	Wilkinson[15], Matthew Ladra[16], Bree R. Eaton[17], Sara L. Gallotto[1] [1]Department of Radiation Oncology, Harvard Medical School, [2]Department of Radiation Oncology, University of Florida, [3]Department of Radiation Oncology, MD Anderson Cancer Center, [4]Department of Radiation Oncology, Northwestern Medicine Chicago Proton Center, [5]Department of Radiation Oncology, Mayo Clinic, [6]Department of Radiation Oncology, University of Washington, [7]Division of Oncology, Cincinnati Children's Hospital Medical Center, [8]Department of Radiation Oncology, Cincinnati Children's Hospital Medical Center, [9]Department of Radiation Oncology, Washington University in St. Louis, [10]Department of Radiation Oncology, Texas Center for Proton Therapy, [11]Department of Radiation Oncology, University of Pennsylvania, [12]Department of Radiation Oncology, ProCure Proton Therapy Center, New Jersey, [13]Department of Radiation Oncology, University of Maryland, [14]Department of Radiation Oncology, ProCure Proton Therapy Center, Oklahoma City, [15]Department of Radiation Oncology, Provision Healthcare, [16]Department of Radiation Oncology, Johns Hopkins University, [17]Department of Radiation Oncology, Emory University	established in May 2010 with enrollment beginning in July 2012. Currently there are 14 institutional members with 6 in the process of opening to enrollment in 2019. Patients may be enrolled prospectively during treatment or during follow up. An optional patient-reported quality of life (PedsQoL) survey is implemented at eight institutions. Baseline and annual follow up health status, symptoms, medications, neurocognitive status, audiogram findings, and neuroendocrine testing are collected. Treatment details of surgery, chemotherapy, and radiation therapy are documented, and DICOM radiation CTs and dosimetric plans are archived with diagnostic imaging (e.g. MRIs). Data is stored in a REDCap database. REDCap is a free, web-based NIH-supported platform housed at and maintained by the core team at MGH. Results: A total of 2,698 patients have consented and enrolled in the PPCR from October 2012 until June 12, 2019. The median age of the cohort is 9.6 years (range 0.5-21.9), 56% male, 32% non-white, and comprised of 70% United States residents. Central nervous system (CNS) tumors comprise 62% of the cohort. The most common CNS histologies are as follows: medulloblastoma (n=418), ependymoma (n=305), glioma/astrocytoma (n=195), craniopharyngioma (n=153), and germ cell tumors (n=252). The most common non-CNS tumors diagnoses are: rhabdomyosarcoma (n=276), Ewing sarcoma (n=179), Hodgkin lymphoma (n=115), and neuroblastoma (n=127). The median follow-up is 1.5 years (range, 0.14-16.2) with 72% disease controlled; 8% with progression, and 7% deceased. Conclusion: This registry of pediatric radiation treated patients receiving radiotherapy has reached a critical milestone of patient numbers and can now facilitate robust clinical outcomes research. It is an important resource for consortium investigators and has the potential for broader use in the academic research community through linking to other datasets.
40	Integrated Pediatric Omics Web Resource to Explore Genomic Data and to Enable Genome Guided Precision Therapeutics	Wei JS, Hsien-Chao C, Tyagi M, Wen X, Khan J. Oncogenomics Section, Genetics Branch, Center for Cancer Research, National Cancer Institute. Bethesda, MD 20892	Cancer is a complex, multifactorial disease that is associated with aberrations in chromosomal ploidy, DNA copy number, epigenetic, mRNA and protein expression profiles, and altered signaling networks. In recent years, the application of high-throughput genomic and proteomic technologies, such as microarrays, next-generation sequencing, and mass spectroscopy, to the study of cancer has greatly improved our understanding of the molecular machinery underlying this disease. These studies, broadly defined as genomics and proteomics, have uncovered several biological themes. Firstly, that

Poster Number	Title	Authors and Affiliations	Abstract
			cancers of different histological types have diagnostic specific gene expression profiles and somatic mutational spectra. Furthermore, cancers can be more accurately classified into clinical, prognostic, and molecular sub-groups by these methods than routine histology. Finally, functional genomics including siRNA screening and pathway analyses, through the enumeration of mutated genes, has implicated modules within cellular circuitry that seem to be essential, across all cancers, for disease initiation and progression. We have developed the Oncogenomics Web Resource to support the use of genomics as a tool to explore the biological underpinnings of cancer and drive genome guided precision therapy. The Oncogenomics web resource provides users with a single-point access to a catalog of data sets generated from microarrays, next-generation sequencing (NGS)-based RNAseq, and mass-spectrometry, for a wide range of cancers and normal tissues and includes data from cell lines, multiple species, and model organisms. Registered users can query each data-set to access data on their genes or functional groups of interest; query results include heat-map and bar chart representations of experimental data as well as link-outs to more detailed annotations. As all data-sets are pre-normalized, users can perform cross-experimental comparisons on their gene sets of interest. All query results can be downloaded as text files. For gene expression data-sets the site enables users to run Gene set enrichment analyses (GSEA) against lists of both curated and custom gene sets. For datasets with patient survival data, users can view Kaplan-Meier plots and the results of pre-run log-rank tests. Finally, we have developed a patient centric database that integrates germline, somatic, RNAseq, gene expression, copy number analysis, gene expression, gene mutational signatures, HLA typing, neoantigen prediction and generates CIRCOS plots for individual patients from Whole Genome, Exome, Panel and RNAseq data. All variants are linked in real time to dbSNP, 1000 gnomes, TCGA, ExAC, Cosmic, Clinvar, and HGMD. Bioinformatic functional predictions for variants are performed in real time using PolyPhen-2, SIFT, FATHMM, MA, VEST, and CADD. The database predicts and presents germline and somatic clinical actionability scores that enable clinical management. Thus, the Oncogenomics web resource provides bench-scientists, bioinformatics scientists, pathologists, geneticists, and clinicians a simple, intuitive interface to explore, visualize, filter, compute and download a large collection of pre-normalized data-sets. Probe, annotation and ontology-based navigation allows users to easily distill data and meta-data on their genes, drugs and pathways of interest as a starting

Poster Number	Title	Authors and Affiliations	Abstract
			point for data-mining, functional genomic studies and generating testable hypotheses. This is also a patient-centric tool for geneticists and molecular pathologists that is currently being utilized at genetic and molecular tumor boards for clinical sign out. We propose that is could be one of the platforms for data visualization and information dissemination for Pediatric Cancers for the Childhood Cancer Data Initiative.
26	Opportunities for epidemiologic and genetic studies of childhood and AYA cancers in an ethnically-diverse population in California	Xiaomei Ma[1], Joseph Wiemels[2], Libby Morimoto[3], Adam de Smith[2], Andrew DeWan[1], Catherine Metayer[3] [1]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT [2]Center for Genetic Epidemiology, Keck School of USC, Los Angeles, CA [3]Division of Epidemiology/Biostatistics, UC Berkeley School of Public Health, Berkeley, CA	The California Childhood Cancer Record Linkage Project (CCRLP) and the California Linkage Study of Early-onset Cancers (CALSEC) are population-based linkages of all cancer cases diagnosed in children, adolescents and young adults up to age 37 years in California from 1988 to 2015, birth records, and non-cancer controls selected from the statewide birth records (N~59,000 cases, N~2,930,000 controls). To date, CCRLP and CALSEC have enabled 14 independently funded projects on pediatric leukemia, CNS tumors, rhabdomyosarcoma, osteosarcoma, and testicular cancer. These data have been utilized for epidemiological studies of childhood cancer risk, including parental age, cesarean section, birth weight, neonatal steroid sex hormones, GIS, epigenetic, and exposome studies. We have optimized a protocol to link subjects to the California Biobank Program at the California Department of Public Health, which has been archiving dried blood specimens (DBS) since 1982. DBS were initially collected by a heel prick of newborns, usually within 24 hours after birth, for the purpose of screening for specific serious health problems that may be successfully treated early. The availability of DBS has enabled us to extract sufficient genomic DNA for both genotyping and sequencing. This led to the largest genome-wide association study (GWAS) of childhood acute lymphoblastic leukemia, to date, and identified three new risk loci. Given the high proportion of individuals of Hispanic/Latino ancestry within California, we were able to conduct a GWAS in which the majority of the subjects were minorities, allowing us to leverage this ancestral diversity to map novel genetic loci. The blood spots are also used for protein and small molecule extractions other than DNA.

Poster Number	Title	Authors and Affiliations	Abstract
			All requests for data use within the approved institutions (Yale University, University of California at Berkeley and San Francisco, and the University of Southern California) are reviewed by the Principal Investigators of CCRLP/CALSEC to ensure that: (1) it does not conflict with planned analyses; (2) it has scientific merit; and (3) there will be no impact on the rights and privacy of human research subjects. Because of California state regulations, data sharing outside the approved institutions is subject to additional review and approval to safeguard the data confidentiality.

Infrastructure to Enable Federation Among Disparate Pediatric Data

Poster Number	Title	Authors and Affiliations	Abstract
31	Towards a Data Ecosystem for Childhood Cancers Using the Data Commons Framework Services	Abby George, Fantix King, Gina Kuffel, Dmitry Grigoryev, Phillis Tang, Zac Flamig and Robert L. Grossman Center for Translational Data Science University of Chicago	There has been progress developing data commons that hold data for childhood cancers, including the NCI Genomic Data Commons, which contains genomic data from the TARGET Project (gdc.cancer.gov), and the Kids First Data Resource, which contains data generated by the Gabriella Miller Kids First Pediatric Research Program (kidsfirstdrc.org). There are also platforms that bring data together from international collaborations, such as the International Neuroblastoma Risk Group (inrgdb.org) An important challenge is to create what is being called a data ecosystem. By a *cancer data ecosystem* we mean a collection of data commons, cloud-based computing services (sometimes known as data clouds), data resources, notebooks and applications over them. One approach to developing a cancer data ecosystem is to interoperate the data commons, data clouds, data resources, notebooks and applications using a set of core software services. These core software services are sometimes called *framework services*. In this poster, we describe how a set of hosted framework services called the Data Commons Framework Services (DCFS) and how they are being used to create a preliminary data ecosystem for childhood cancer data. We show how the DCFS are being used so that Jupyter Notebook-based applications can be built that access data from the NCI GDCI and the Kids First Data Resource within a secure compliant infrastructure, enabling a bioinformatician to analyze the data easily. We also show applications with this architecture can be enriched, in certain cases, using data from resources such as NHLBI's DataSTAGE. The DCFS are built over the open source Gen3 software stack (gen3.org) and include services for authentication, authorization, assigning digital IDs to data objects, such as BAM, CRAM or image files, and services for working with structured data, such as clinical data or biospecimen data. Finally, we identify some challenges, including when data resources limit the ability to analyze data using other resources in the data ecosystem (imposed by the data

Poster Number	Title	Authors and Affiliations	Abstract
			resource) and when secondary use restrictions on data limit how it can be used (imposed by the data contributor)
19	A Synthesized Common Data Model and Data Standards for the University of Chicago's Pediatric Cancer Data Commons	Alejandro Plana, BA; Brian Furner, MS; Monica Palese, MPH; Samuel Volchenboum, MD PhD	Pediatric cancer research is hindered by limited cases available for properly-powered analyses. To increase available sample sizes, the University of Chicago Pediatric Cancer Data Commons (PCDC) has established a novel system for harmonizing international clinical data for storage and sharing in scalable and extensible data commons infrastructure. Here, we present our efforts to build a set of data standards facilitating interoperability across disparate data repositories. The National Cancer Institute's Enterprise Vocabulary Service (NCI-EVS) provides the PCDC with a comprehensive terminology system including term mappings and metadata. Because each cancer research community has developed its own unique set of clinical data elements, the PCDC works with international leaders in several pediatric cancer types to build a master list of common data fields with permissible values and definitions. This allows for harmonization between completed studies, easier data sharing, and sustainable prospective data collection across institutions and consortia. To facilitate the integration of these data standards with a data model, the Biomedical Research Integrated Domain Group (BRIDG) project brings a structured protocol and study representation to the PCDC. It augments the CDISC ODM (Clinical Data Interchange Standards Consortium Operational Data Model) with detailed semantics. The PCDC has identified this data model as the most robust and internally-consistent method for building a domain-specific data model. With recent developments of BRIDG mappings to several Fast Healthcare Interoperability Resource (FHIR) resources, this model also leaves room for maturation as more data commons develop and as data standards evolve. By employing data standards that leverage the breadth of the NCI-EVS and the use-case independent BRIDG data model, the PCDC has rooted its system in a stable data ecosystem that fosters sustainable and collaborative growth. This system serves as the foundation upon which an instance of the Bionimbus Gen3 data commons platform has been instantiated to store these data, as well as for future mapping to the FHIR model. Ultimately, this rubric of engaging the international pediatric oncology community and leveraging well-established data standards resources provides a method for building data commons capable of integration into the NCI's proposed framework of federated tools and commons.
20	Rapid consensus building and development of the International Soft	Alejandro Plana, Brian Furner, Monica Palese, Suzi Birz, Douglas S. Hawkins, Samuel L. Volchenboum	Background: Large-scale pediatric cancer research requires consortium-driven data commons. The International Soft Tissue Sarcoma Consortium (INSTRuCT) formed in 2018 for harmonization of data from trials into the UChicago Pediatric Cancer Data Commons (PCDC). In 18 months, we built and deployed an international data

Poster Number	Title	Authors and Affiliations	Abstract
	Tissue Sarcoma Consortium (INSTRuCT) data commons		commons through consistent project and team management, collaborative consensus-driven development, and an aggressive stakeholder engagement timeline enabled by a strong foundation of trust developed by the cooperative group leads. Methods: Since 2018, INSTRuCT convened five in-person meetings with representation from the Children's Oncology Group and the European CWS and EpSSG, with MMT and STSC contributing as legacy groups to drive the process. Through these meetings and routine calls and guided by a consensus-derived list of scientific questions, case report forms from over 10 completed studies were used to build a consensus data dictionary. Statisticians from each of the cooperative groups harmonized the trials data to the new standard. A UChicago expert in HIPAA and GDPR worked closely with lawyers and consortium investigators to execute data contributor agreements, and the data were securely stored within UChicago's HIPAA-compliant infrastructure. Data governance discussions culminated in a set of operating policies and procedures for data contribution and use. Results: The INSTRuCT commons contains data on over 4,000 children with from multiple continents. The publicly-available INSTRuCT cohort discovery tool allows researchers to query over 30 clinical features to discover patients matching select criteria. Query results can be used to request data from the Executive Committee. Additional rhabdomyosarcoma data are being harmonized and will soon be deposited into the commons, and outcomes data will be updated periodically. Conclusions: Development of data commons can be slow. Persistent meeting cadence and momentum, an aggressive development timeline, and rigorous stakeholder engagement enabled INSTRuCT and the UChicago PCDC team to successfully ballot a consensus pediatric rhabdomyosarcoma data dictionary and stand up a data commons cohort discovery tool within an 18 month timeline. Data harmonization for non-rhabdomyosarcoma soft tissue sarcomas, enhanced graphic analysis tools, automated data ingestion with quality control, connection to external data sources, and alignment with NCI's federated cloud-based infrastructure ecosystem are critical next steps.
32	Enabling Data Discovery and Analysis in the Placental Atlas Tool (PAT) Through Semantically-Driven Data Federation	Alexandra Shlionskaya, M.S., Christopher H. Ferguson, Ph.D., Bianca Patel, M.S., Andrijana Dabic, B.S., Brett Pickett, Ph.D., Frederick Dong, Ph.D., Michael Keller, Ph.D. Booz Allen Hamilton, Inc. McLean, Virginia, USA	Description: In October 2018, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) launched the Placental Atlas Tool (PAT) system to provide a comprehensive placental knowledgebase, analytic tools, and relevant publications linked to curated molecular and image data. This supports visualization of placental function and development down to the molecular level, melding of diverse datasets into an integrated resource, and scientific data analysis across studies.

Poster Number	Title	Authors and Affiliations	Abstract
			Applicability: PAT is built on the principles of data accessibility and reuse. PAT facilitates harmonization of placental research terminology that is used as metadata for data discovery, annotation, association, and analysis. Using the terminologies and native APIs, data is extracted from multiple disparate public and proprietary resources including NCBI Gene Expression Omnibus (GEO), EMBL-EBI ArrayExpress, NLM Open-i. Use of standardized terminologies enhances data findability, links disparate data entities, and facilitates knowledge integration. Analytic pipelines enable users to extract knowledge from data through bioinformatic analysis of molecular datasets accessible to all researchers regardless of bioinformatic experience. Potential for broader use: The PAT system employs a microservice architecture and terminology-based, data-driven design principles to facilitate a flexible, scalable, and user-friendly experience that is adaptable to any research area of interest. The use of standardized terminologies allows for PAT to be easily extended to other research areas, such as pediatric neuro-oncology, by extending or replacing of the system terminologies. Evidence of success & lessons learned: Since release, the PAT system has garnered enthusiasm and positive feedback from users across the placental research community. Currently, PAT contains over 500 datasets and over 800 images and has been explored by more than 3000 users. Additionally, PAT is being incorporated into college curriculum as a learning tool to introduce undergraduates to bioinformatics and molecular data analysis. Analysis of system use has demonstrated that molecular data findability and annotation accuracy are the most valued aspects of PAT, while publication-derived images and duplication of PubMed functionality have limited value.
5	Gabriella Miller Kids First Data Resource Center: Collaborative platforms for accelerating cross-disease pediatric	Allison P. Heath PhD[1], Christina Yung PhD[2], Yuankun Zhu BS[1], Michele Mattioni PhD[3] ,Zachary L. Flamig PhD[4], Yajing Tang MS[4], Bailey Farrow MS[1], Jena Lilly MS[1], Yiran Guo PhD[1], Pichai Raman PhD[1,5], Phillip B. Storm MD[1,9], Luca Graglia, MS[13], Samuel L. Volchenboum MD PhD[8, 13], Javad Nazarian PhD[6], Nicole Vasilevsky PhD[11], Jack	Pediatric cancer and structural birth defects represent diseases of childhood development proposed to share common genetic etiologies. The Gabriella Miller Kids First Data Resource Center (DRC) is a recently launched, multi-year initiative seeking to harness the potential of cloud-based collaborative research platforms supporting FAIR data principles. These platforms are brought together on the behalf of empowering integration, use, and clinical translation of diverse, multi-disease, large-scale genomic and clinical datasets. With a particular focus on family-based genomic

Poster Number	Title	Authors and Affiliations	Abstract
	research across development and cancer	DiGiovanna PhD[3], Melissa Haendel PhD[11,12], Robert L. Grossman PhD[4], Brandi Davis-Dusenbery PhD[3], Deanne M. Taylor PhD[5,7], Vincent Ferretti PhD[10], Adam Resnick PhD[1,5,9] [1]Center for Data-Driven Discovery (D3b) Children's Hospital of Philadelphia, Philadelphia, PA, USA; [2]Ontario Institute of Cancer Research, Toronto, Ontario, Canada; [3]Seven Bridges Genomics, Cambridge, MA, USA; [4]Center for Translational Data Science, The University of Chicago, Chicago, Illinois, USA; [5]Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA;; [6]Children's National Medical Center, Washington, DC, USA; [7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [8]Department of Pediatrics, Biological Sciences Division, The University of Chicago, Chicago, Illinois, USA; [9]Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [10]CHU Sainte-Justine Research Center, Montreal, Quebec, Canada; [11]Translational and Integrative Sciences Lab, Oregon Health & Science University, Portland, OR, USA;; [12]Oregon State University, Corvallis, Oregon, USA; [13]Center for Research Informatics, The University of Chicago, Chicago, Illinois, USA;	studies with associated clinical and phenotypic data, more than 35 cross-disease pediatric cohorts, and associated biospecimens, have been defined, providing for a first-in kind network of leading researchers, sequencing facilities, and patient communities. By the end of 2019, genomic and clinical data from over 15,000 participant samples will be available across a variety of structural birth defects and pediatric cancer cohorts.The newly launched DRC provides for a secure, cloud-based platform that supports the ability of researchers to not only find, access, and reuse data, but also integrate, collaborate, and analyze data quickly and at scale. Recognizing the value in both original source data and the power of harmonized data for cross-disease analyses, the DRC supports the availability data across the entire data life cycle for all cohorts in its platforms. A "best-of-breed" approach has been taken by our multi-institutional team to develop a platform comprising reusable technology stack components and a portal entrypoint that provides comprehensive file browsing and dynamic cohort creation. Users have direct access to analytic workspaces and multi-user collaborative projects, such as Cavatica for bioinformatics workflow deployment and pedCBioPortal for cancer genomic visualizations. Additionally, the DRC has implemented a set of core services, powered by Gen3, that ensure a foundation for interoperability with other large-scale NIH data sources and the NCI Data Commons Framework. The allows the DRC to participate in the emerging ecosystem of analysis and visualization applications that provide a highly diverse pediatric research community secure access to large-scale federated via multi-environment user authentication. On the near term roadmap is the ability to create cohorts utilizing genomic features and improved visualizations for germline and family-based genomic data.

Poster Number	Title	Authors and Affiliations	Abstract
10	Building a pediatric oncology research database using point-of-care electronic data capture and automated data extraction software	Angela Holmes[1], Christopher Ahern[1], Christopher Berlind[1], Benjamin Smith[2], Tim Edwards[2], William D. Lindsay[1], Susan L. McGovern[2] [1]Oncora Medical, Inc. [2]Department of Radiation Oncology, University of Texas MD Anderson Cancer Center	Background: Pediatric oncology research is limited by a lack of structured data on patient characteristics, treatments, and outcomes. Such information is distributed across multiple software systems at disparate clinical sites in unstructured form. In pediatric oncology, where the volume of patients treated is small, it is difficult to build meaningful datasets for research use, limiting the utility of this valuable patient experience for addressing outstanding questions in treatment and outcome. Methods: To address this problem, the radiation oncology service at MD Anderson implemented a point-of-care electronic data capture solution to prospectively capture structured data on every patient treated and automated data extraction software to combine these data with information obtained from the electronic medical record, the radiation record and verify system, and the institutional tumor registry, creating holistic patient data from diagnosis through treatment and follow up. Source data are automatically processed into clinically-relevant formats by the software, making more than 1,000 individual data elements available for each patient including demographics, medical history, pathology, tumor markers, staging, systemic therapy, surgical procedures, radiation treatments, vital signs, laboratory values, adverse events, hospitalizations, and vital status. Results: Data have been captured and unified for more than 700 pediatric cancer patients treated at MD Anderson since January 2016. This dataset grows daily with each new patient treated. Physicians have real-time access to this information and custom patient cohorts can be created based on user-specified inclusion and exclusion criteria. Kaplan-Meier estimates of survival can be generated instantly. With appropriate IRB approval, investigators can download data to support trials and registry submissions. Conclusions: Structured data capture combined with automated data extraction can be implemented at a high-volume academic radiation oncology department. Future effort will be directed at improving the scale and quality of data collected from pediatric patients on treatment and in follow up. Multi-institutional data collection and data sharing will be essential for maximizing the potential of real world data, predictive analytics, and research platform technologies to improve outcomes for children with cancer.

Poster Number	Title	Authors and Affiliations	Abstract
25	The Utah Population Database: A unique resource for research across the lifespan of childhood, adolescent, and young adult cancer	Anne C. Kirchhoff,[1] Judy Ou,[2] Heydon Kaddas,[2] Joemy Ramsay,[2] Sapna Kaul,[3] Heidi Hanson[4] [1]University of Utah, Department of Pediatrics, Salt Lake City, UT [2]Huntsman Cancer Institute, Salt Lake City, UT [3]University of Texas Medical Branch, Preventive Medicine and Community Health, Galveston, TX [4]University of Utah, Department of Surgery, Salt Lake City, UT	Data Description: Population level data sources that encompass demographic, clinical, and genetic information across time are needed to identify risk factors and related outcomes of childhood, adolescent, and young adult (AYA) cancer throughout the life course. The Utah Population Database (UPDB) is one of the world's richest sources of epidemiologic data containing cancer diagnosis and treatment information from the Utah Cancer Registry linked to an individual's demographic and electronic healthcare records from statewide resources including: healthcare systems serving >85% of Utahans, insurance claims, state-wide vital statistics, hospitalization data, administrative records, and residential history. In addition, UPDB has pedigree information that spans 18 generations in some cases and can be used identify individuals belonging to families with a history of cancer. Applicability: Using UPDB data, we created a cohort of over 24,000 childhood, adolescent, and young adult (AYA) cancer patients diagnosed in Utah since 1986. In this cohort, there are 4,435 patients diagnosed ages 0-14 years and 20,155 patients diagnosed ages 15-39 years. This cohort can be accessed via approval from the Utah Resource for Genetic and Epidemiologic Research and the University of Utah Institutional Review Board. Potential for Broader Use: UPDB data has been linked to other rich data resources, including environmental, geographic, and Census data. The UPDB provides access to information on more than 8 million individuals and currently supports over 200 research projects. The research community could use this resource to answer novel questions regarding family cancer history, environmental risk factors for cancer, gene-environment interaction, long-term health effects of cancer treatment, and other topics pertaining to childhood and AYA cancers. Evidence of success: To date, our team has utilized the UPDB childhood and AYA cancer data to publish 13 manuscripts with an additional four in development on health care costs, utilization, and outcomes, and well as manuscripts on environmental exposures and health, among both patients and survivors of childhood and AYA cancer. Our studies often use patients' siblings or same age/sex individuals from the Utah population without a history of cancer as comparison samples. We have obtained both federal and foundation funding to support research with these data.
39	Bridging the Gap: A Two Center Pediatric and Adult	Anthony Audino, MDNCH Nicholas Yeager, MD NCH Tammi Young-Saleme, PhD, NCH Samantha Hulett, SW James Marie Michalik SW NCH Annie	Introduction; In Ohio, the cancer incidence rate of adolescents and young adults (AYAs) is 337 cases/100,000 people. Prior investigation of the 15-19 year old group suggests that age of the patient, specific diagnosis, and distance from an academic

Poster Number	Title	Authors and Affiliations	Abstract
	Hospital Adolescent and Young Adult Program	Trance, SW James Mary Connelly SW, James Dori Klemanski DNP, James Maryam Lustberg, MD James Bhavana Bhatnagar, DO, James	center have significant impact on the location of treatment for this group. Our collaborative program aims provides a means to bridge the gap and address the unmet needs of this age group. Methods; The James Cancer Hospital and Nationwide Children's Hospital (NCH) are two of the largest cancer programs in the United States that provide care to this age group. Both hospitals are centerpieces of the NCI-designated Comprehensive Cancer Center at the OSUCCC-James. Development of this joint AYA program at OSUCCC-James/NCH aims to unify efforts and bolster activities in this age group, as there is currently a gap in AYA services (programming and research) at both institutions. Recently awarded a grant from Teen Cancer America, this Collaborative program aims to towards the following: 1.Navigate through the obstacles of rooted cultures, physical space constraints, and provider expertise in the attempt to increase access to appropriate care and achieve a desired change in the outcomes of AYA patient with cancer. 2. Develop an AYA single registry database. The joint AYA program includes oncofertility services, symptom management and palliative care, psychosocial oncology care, social worker and survivorship care. Results: Between 2016 and 2019, total of 4730 AYA patients have been seen at the OSUCCCJames/NCH. Predominant Cancer diagnoses amongst this population are Hematologic 20.8% (n=944), Thyroid 10.8% (n=490), Brain 9.9% (n=451) and Breast cancer 9.1% (n=414). Cancer diagnosis for AYA patients seen at NCH included Hematologic malignancy 36% (n=72), Brain tumors 25% (n=51) Sarcoma 27% (n=54), Germ Cell tumors 5% (n=11) and Thyroid 4% (n=9). Conclusion: Development of a two center AYA program will nurture a close collaboration between the pediatric and adult oncology teams in clinical care, programming and research. Ultimately, an improvement in the care of this vulnerable population is projected.
1	Describing the database infrastructure for a multi-institutional pediatric radiation registry	Benjamin V. Bajaj[1], Sara L. Gallotto[1], Miranda P. Lawell[1], Daniel Indelicato[2], Arnold Paulino[3], William Hartsell[4], Nadia Laack[5], Ralph Ermoian[6], John Perentesis[7], Stephanie Perkins[8], Victor Mangona[9], Christine Hill-Kayser[10], Suzanne Wolden[11], Young Kwok[12], Michael Confer[13], J. Ben Wilkinson[14], Torunn I. Yock[1] [1]Department of Radiation Oncology, Harvard Medical School, [2]Department of Radiation Oncology, University of Florida, [3]Department of Radiation Oncology, MD Anderson Cancer Center,	Background/objectives: The Pediatric Proton and Photon Consortium Registry (PPCR) was established in 2010 at Massachusetts General Hospital (MGH) and currently consists of 14 institutions across the United States. Here we describe the Registry's data storage infrastructure and database management processes. Design/methods: The PPCR uses REDCap, a free, NIH supported web-based database platform for storage of all treatment, diagnostic, and clinical information for patients enrolled on the study. REDCap allows for manual data entry or data imports from other sources. The PPCR core team at MGH has established electronic data transfers from hospital-wide electronic medical records (EMR) directly into the database using Application Programming Interfaces (APIs) provided by REDCap. Data may be extracted by manual download or electronically using APIs and other built-in

Poster Number	Title	Authors and Affiliations	Abstract
		[4]Department of Radiation Oncology, Northwestern Medicine Chicago Proton Center, [5]Department of Radiation Oncology, Mayo Clinic, [6]Department of Radiation Oncology, University of Washington, [7]Division of Oncology, Cincinnati Children's Hospital Medical Center, [8]Department of Radiation Oncology, Washington University in St. Louis, [9]Department of Radiation Oncology, Texas Center for Proton Therapy, [10]Department of Radiation Oncology, University of Pennsylvania, [11]Department of Radiation Oncology, ProCure Proton Therapy Center, New Jersey, [12]Department of Radiation Oncology, University of Maryland, [13]Department of Radiation Oncology, ProCure Proton Therapy Center, Oklahoma City, [14]Department of Radiation Oncology, Provision Healthcare	reporting features within REDCap. APIs are also used to build connections between the PPCR REDCap database and other institutional databases. Results: The use of REDCap as a data storage platform for the PPCR allows for easy management of all data management tasks. APIs allow seamless communication between the PPCR database and other clinical trials' REDCap databases for real-time data comparison and verification. With connections to EMRs, patient records in REDCap can be created centrally by the PPCR core team for all registered patients, eliminating errors that can be associated with manual data entry. Weekly and monthly reports confirm information is accurate, identify corruptions in the data that would cause failure of future data exports, and generate Missing Fields Reports and Data Quality Reports for all institutions within the PPCR to ensure the data is accurate and remains up to date at all participating sites. Conclusion: The PPCR's REDCap database stores detailed records for over 2700 patients enrolled on the study. Fast, reliable management and sharing of data allows our collaborators easy access to view, analyze, and report on their own patients. Utilizing REDCap's APIs and data management tools maintains data integrity and insures data completeness, which are key priorities for the Registry. A framework that capitalizes on user-friendly data storage coupled with modern EMR data extraction tools improves data quality and overall scalability of the PPCR.
67	Diverse noncoding mutations contribute to deregulation of cis-regulatory landscape in pediatric cancers	Bing He[1], Peng Gao[1], Yang-Yang Ding[1,2], Chia-Hui Chen[1], Gregory Chen[3], Hannah Kim [1,] Sarah K. Tasian[1,2], Stephen P. Hunger [1,2], Kai Tan [1,2,4,5,6] [1] Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA [2] Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA 3 Medical Scientist Training Program, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA 4 Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania	Various types of mutations can disrupt the function of cis-regulatory sequences, including point mutations, small insertions and deletions (indels), and large structural variants such as copy number variations and translocations. Very few studies have conducted comprehensive analysis of the full spectrum of noncoding mutations in large cohorts of cancer genomes. Here we develop PANGEA (Predictive Analysis of Noncoding Genomic Enhancer/Promoter Alterations), an analytical framework to identify simultaneously all classes of risk noncoding mutations by joint analysis of mutation and gene expression data. By applying PANGEA to matched whole genome sequencing and RNA-Seq data of 501 pediatric cancer patient samples of five histotypes, we identified 1,405 single nucleotide variants (SNVs)/small indels and 1,137 structural variants (SVs) in noncoding regions as candidate risk mutations. They affect gene expression by either disrupting enhancer/promoter sequences or perturbing enhancerpromoter communication. We experimentally validated the oncogenic role of CHD4 translocation in B-ALL that results in enhancer hijacking and

Poster Number	Title	Authors and Affiliations	Abstract
		19104, USA 5 Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA 6 Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA	CHD4 overexpression. We observed a general exclusivity of coding and noncoding mutations affecting the same genes and pathways. We found that the level of disruption of a transcription factor regulon could be used to stratify patients in terms of their relapse-free survival. Finally, we showed that integrated mutation signatures can help define novel patient subtypes with favourable and unfavourable clinical outcomes. In summary, our study introduces a general strategy to identify systematically and characterize the full spectrum of noncoding mutations in a variety of childhood cancers.
2	A Novel Cloud Infrastructure for DICOM Images and Radiation Treatment Data Analysis (BrlCS) for the Pediatric Proton/Photon Consortium Registry and Beyond: Facilitating Outcome Analysis and Multi-institutional Clinical Trials to Improve Outcomes for Children with Brain Tumors	Bree Eaton[1], Torunn I. Yock[2], Alexander Giuffrida[1], Karthik Ramesh[1], Sara L Gallotto[2], Miranda P Lawell[2], Brent Weinberg[3], Hyunsuk Shim [1,3] [1]Department of Radiation Oncology, Emory University, [2]Department of Radiation Oncology, Harvard Medical School, [3]Department of Radiology & Imaging Sciences, Emory University	Multi-institutional collaboration and data sharing is essential to advancing care for rare and fatal diagnoses, such as pediatric brain tumors. The Pediatric Proton/Photon Consortium Registry (PPCR) is a multi-institutional prospective childhood cancer registry established to expedite outcome research for children receiving radiation therapy. The PPCR is a unique in that high-quality treatment delivery and outcome data such as DICOM radiation dose files and baseline and follow-up imaging data are collected, along with toxicity, disease status and health outcomes data. The successful collection of DICOM imaging and radiation data as part of multi-institutional pediatric registry is unprecedented, yet there remains a critical need for infrastructure to facilitate the analyses and real-time sharing of the data. The Emory Brain Imaging Group's Brain Imaging Collaboration Suite (BrICS) is a novel cloud infrastructure enabling real-time collaboration among multisite investigators for imaging and treatment of brain tumors. A browser interface allows users to upload, access, and process de-identified data from multisites, and to collaborate with ease and includes tools to interface with HIPAA-compliant databases, such as REDCap. BrICS was developed to promote multisite implementation of spectroscopic MRI (sMRI), a quantitative imaging technique that measures endogenous metabolic differences without contrast agents, and has successfully been used in a multi-site clinical trial of sMRI-guided radiation therapy in adult patients (NCT03137888). BrICS includes modules for image registration, quality control, spectra visualization, clinical data retrieval, automatic segmentation, and radiation target volume generation. A longitudinal imaging tracking (LIT) module incorporates the assistive tools for implementing the Brain Tumor Reporting and Data System (BT-RADS), a management-based structured reporting system for brain tumor imaging developed at Emory. LIT automatically co-registers all of a patient's scans and includes automatic segmentation tools and a radiation dose map overlay to allow for automated

Poster Number	Title	Authors and Affiliations	Abstract
			volumetric assessment of tumor response, disease progression or radiation associated toxicity. Combining BrICS with the PPCR will facilitate the automated analysis of imaging, radiation, and clinical data and presents unique opportunity to directly impact care of pediatric brain tumor patients. Furthermore, BrlCS can be utilized to facilitate multi-institutional clinical trials involving advanced imaging to further advance care for pediatric brain tumor patients.
20	Linking clinical trials data with images via the Pediatric Cancer Data Commons and the National Biomedical Imaging Archive (NBIA)	Brian Furner[1], Tomasz Oliwa[1], Luca Graglia[1], Bron Kisler[2], Fran Laurie[3], Julie Evans[2], Jian Tian[1], Julie Johnson[1], Rakesh Veetekat[3], Richard Hanusik[3], Allen Dearry[2], Susan Cohn[1], Samuel Volchenboum [1]University of Chicago [2]NIH/NCI [3]Imaging and Radiation Oncology Core (IROC) Rhode Island Center	Background: For pediatric cancer research, data commons are becoming essential for storing and democratizing harmonized clinical trials data from international consortia. In collaboration with the NCI, the University of Chicago has implemented the National Biomedical Imaging Archive (NBIA) platform providing researchers with an easy-to-use interface for querying harmonized clinical data alongside information about multiple types of DICOM-standardized radiology images. Methods: The International Neuroblastoma Risk Group (INRG) data commons at UChicago contains clinical data on over 21,000 children treated on clinical trials sourced from international consortia and harmonized to a common data model. New data are added every six months, and rigorous quality control and governance processes have been implemented for access. Images for clinical trials sponsored by the Children's Oncology Group (COG) are stored in the Imaging and Radiation Oncology Core (IROC) Rhode Island Center. By linking on the USI, a publicly-available identifier from COG, these images are associated with the clinical data in the INRG commons. This unique NCI-UChicago-IROC collaboration links and makes available MIBG, CT, and MRI images alongside the clinical data in the INRG commons. Results: NBIA software provides an API-enabled platform in which to store DICOM images for retrieval and visualization. An instance of NBIA was deployed to UChicago's HIPAA-compliant infrastructure. Working closely with the IROC team, images were de-identified, associated with COG USI's, and transferred to UChicago for storage and integration. Reports are processed, text extracted, parsed, and added to the commons. Through the cohort discovery tool, clinical data can be queried and image availability can be ascertained. Some DICOM metadata and report elements will be made available for searching in the INRG commons interface. Conclusions: Data commons are most useful when they contain linked multiple disparate data types. Because the COG provides a universal identifier for all patient data and samples, it is possible to link clinical data with other sources. The INRG commons, with its access to rich clinical and imaging data mapped to the BRIDG

Poster Number	Title	Authors and Affiliations	Abstract
			model, is a robust platform for research, allowing investigators to query images, clinical elements, genomic data, and biospecimen availability.
19	The Pediatric Acute Leukemia (PedAL) initiative - an innovative platform for real-time matching of children with relapsed AML to early-phase clinical trials	Brian Furner[1], Alejandro Plana[1], Monica Palese[1], Gwen Nichols[2], E. Anders Kolb[3], Samuel Volchenboum[1] [1]University of Chicago [2]Leukemia & Lymphoma Society [3]Nemours Alfred I. duPont Hospital for Children	Background: Pediatric acute myelogenous leukemia (AML) remains deadly for 30% of children, many after relapse. The Leukemia & Lymphoma Society (LLS) Pediatric Acute Leukemia (PedAL) Master Trial Initiative will provide target-specific treatment within 48 hours for all children with relapsed or refractory AML. The University of Chicago Pediatric Cancer Data Commons (PCDC) team is building informatics infrastructure to support PedAL and to create the world's largest AML data commons. Methods: The PCDC is implementing three innovative approaches for PedAL: (1) Working with Children's Oncology Group (COG) and the Berlin-Frankfurt-Münster (I-BFM) group to ballot an AML data dictionary, enabling harmonization of completed clinical trials data and future trial and registry forms. (2) Creating interfaces for automatic data transfer between stakeholders and the COG Medidata Rave platform, alleviating the burden of manual data extraction and entry. (3) Creating an platform that provides clinicians with real-time patient-trial matching and enrollment instructions. Results: Starting with case report forms from 6 US / European trials, a set of common data elements was built and reviewed at the May SIOP-E meeting. A final data dictionary will be presented at the fall 2019 SIOP meeting. Inclusion of the NCI's Enterprise Vocabulary Service team ensures that elements are available for future data collection. The PedAL team and COG are updating collection forms for the Project EveryChild registry and upcoming Phase 3 AML trial. The PedAL team is working with Foundation Medicine and Hematologics to ensure that panel sequencing and flow cytometry data will be automatically moved into the Rave platform, eliminating costly and slow manual data extraction and entry. Finally, PedAL is developing the Genomic Eligibility Algorithm at Relapse for Better Outcomes (GEARBOx) platform, for clinicians to query real-time trial matching using a patient's clinical and genomic data. Conclusions: The PedAL initiative is a collaboration between UChicago, COG, and I-BFM, with support from LLS and backing from the NCI to develop sustainable informatics infrastructure that will revolutionize the treatment and outcomes of children with AML. The collection of data through this initiative will result in the largest data commons for AML and will enable research and innovation for better cures.
12	Streamlined sharing of clinical patient	Brigitte E. Raumann[1] and Ian T. Foster[2]	Advances in genomics and data analytics create new opportunities to advance cancer research via large-scale sharing of genomic, clinical, imaging and other data types

Poster Number	Title	Authors and Affiliations	Abstract
	data for cancer research networks	[1]Globus, University of Chicago [2]Department of Computer Science, University of Chicago	from patients across institutions around the world. Yet these opportunities are often stymied by a lack of tools for the reliable, secure, rapid, and easy transfer and sharing of large collections of human research data. In the absence of such tools, security and performance concerns often force researchers to resort to slow and error prone shipping of physical media, or worse still, prevent sharing altogether. If data are received, timely analysis is further impeded by the difficulties inherent in verifying data integrity and managing who can access data and for what purpose. The Globus research data management platform (www.globus.org) addresses these obstacles to discovery in data-intensive cancer research with secure, high-quality cloud-hosted Software-as-a-Service (SaaS). It provides intuitive, web-based interfaces that people without specialized IT knowledge can use to move, replicate, synchronize, and share data sets with high reliability and speed, thanks to integrated monitoring, failure recovery, and optimization. Globus capabilities are also accessible via simple REST application programming interfaces (APIs), allowing developers to provide robust file transfer and sharing capabilities within their own research data applications and services, while leveraging advanced identity management, single sign-on, and authorization capabilities. Encryption in transit, auditing, access control, and other features are provided to meet the security requirements of human research data and HIPAA. Globus is widely used in the research community, with over 20,000 users in the past year, representing most leading US universities, national laboratories, and many sites overseas. Globus capabilities are ideal for managing the voluminous datasets produced by next generation sequencing and the many biomedical imaging modalities, data types especially relevant in cancer research. Our technology has been applied in a variety of biomedical research contexts where science can be accelerated with rapid, reliable, and secure data transfer and sharing, such as collaborative networks, sequencing and imaging facilities, data portals, campus computing clusters, supercomputers, and public and private clouds.
15	Defining Childhood/AYA Melanoma through a Multi-Institutional Database	Brittani K. Seynnaeve, MD, MS[1,2], John M. Kirkwood, MD[2] [1]University of Pittsburgh Medical Center Children's Hospital of Pittsburgh, [2]University of Pittsburgh	Cutaneous melanoma is the deadliest of all skin cancer and is linked to nearly 75% of skin cancer-related deaths. Because of the lower frequency and often atypical presentation of childhood and AYA melanoma, there is often diagnostic and management uncertainty for this potentially deadly disease, leading to suboptimal outcomes. There has been a trend toward less aggressive management for this entity, although specific features predicting those cases that are more or less likely to have

Poster Number	Title	Authors and Affiliations	Abstract
		Medical Center Hillman Cancer Center, Pittsburgh, Pennsylvania, United States	progression and/or mortality are not clearly defined. Standard recommendations for the management of childhood and AYA melanocytic neoplasms are desperately needed, but large cohorts of organized patient data to serve as the groundwork for such recommendations are lacking. Our goal was to assemble a current multi-institutional database to better define this challenging cohort of melanoma patients, to improve prognostication and tailor therapy accordingly. Additionally, difficulties pervade childhood/AYA melanoma literature as pathologic case validation has often been omitted. Thus, our retrospective database has included a linked digitized pathology archive of specimens for multi-observer pathology validation. We have created a REDCap database with extensive diagnostic, clinical course, and long-term outcome data. We have established collaboration with eight large melanoma treatment sites that have contributed more than 250 fully annotated cases to date, with ongoing acquisition. We aim to thoroughly describe this population and to identify prognostic factors affecting survival in patients ≤25 years of age. Through this work, we have established our ability to lead a multi-institutional database effort and have learned the critical importance of inter-institutional collaboration in the assessment of rare diseases. Within this solid framework, there is a vast potential for developing a broader scope. Specifically, we hope to expand this effort to (1) include the challenging full gamut of atypical melanocytic and Spitzoid neoplasms and (2) prospectively collect data for all childhood and AYA melanocytic neoplasms for ongoing clinical care, research and data sharing. Significant impact upon the approach to these young patients relies on such efforts, to finally formulate care guidelines for optimal management and outcomes.
13	Unified access to cancer proteogenomics data	Caleb M. Lindgren, David W. Adams, Sadie Taylor, Sean J.I. Beecroft, Isaac D. Taylor, Samuel H. Payne Biology Department, Brigham Young University, Provo UT	Cancer data has many audiences. Although a small number of scientists may be involved in generating the clinical and molecular characterization of cancer cohorts, the data is a national resource that should be broadly distributed to facilitate analysis by a wider community of researchers, data scientists and even the lay public. Therefore, expanding access to data is an important goal for publicly funded research. To seamlessly enable analysis, data must be distributed in a convenient and accessible manner. Although storing data in supplemental tables or cloud-based archives is fine for historical records, it is not the optimal dissemination method for active collaborations or ongoing analyses; researchers need to be able to access data

Poster Number	Title	Authors and Affiliations	Abstract
			in their analytical software via APIs. No convenient dissemination method currently exists for the quantitative molecular data tables that are the primary input of data interpretation and analysis codes. We present an example of a unified API for accessing proteogenomic data from CPTAC cancer cohorts, a model that could be adopted by NCI for childhood cancer data. Each CPTAC dataset contains comprehensive genomics, transcriptomics, proteomics, and clinical data for a tumor-specific cohort - e.g. ovarian, endometrial, colon, etc. This data is contained within a Python package, cptac, that is freely distributed through the Python Package Index (PyPI). Our package removes many common barriers to analysis by automating all data loading and formatting to make the data ready for statistical and visual analytics. Additionally, the package handles complex merging between data tables and includes common algorithms for analyses. The module contains extensive tutorials and documentation to assist users in understanding the data and analysis methodology.
14	Gabriella Miller Kids First Pediatric Research Program (Kids First)	D. Winchester, Ph.D.[2], V. Cotton[2], The Gabriella Miller Kids First Pediatric Research Program Working Group [1]Office of the Director, Office of Strategic Coordination -The Common Fund, National Institutes of Health, [2]Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health	Epidemiological studies have reported that children with congenital anomalies are at a higher risk of developing certain childhood cancers. Little is known about the genetic cause-and -effect relationship between these disorders. Furthermore, a large dataset is needed to examine the role of genetics in childhood cancer and birth defects. The National Institutes of Health Common Fund's Gabriella Miller Kids First Pediatric Research Program (Kids First) is developing a large-scale data resource of clinical and genetic data from patients with childhood cancers and structural birth defects and their families. These data are now available through the Gabriella Miller Kids First Data Resource Portal which was launched by the program's Data Resource Center in 2018. To date, Kids First has selected 27 patient cohorts for whole genome sequencing through a peer-review process. Clinical and genetic sequence data from over 5,000 patient samples are accessible through the portal, including data from patients with Ewing Sarcoma and Congenital Diaphragmatic Hernia. Data from more than 30,000 DNA and RNA samples are expected to be added to the Kids First Data Resource Portal over the next few years. The Kids First program is focused on data sharing to develop tools and resources to foster collaborative analyses. The program aims to help researchers uncover new insights into the biology of childhood cancer and structural birth defects and identify shared genetic pathways among these various pediatric conditions. A complete list of conditions represented in the Kids

Poster Number	Title	Authors and Affiliations	Abstract
			First data resource and additional information about the Kids First program will be described in this presentation.
24	Rapid Case Ascertainment in Childhood Cancer Surveillance Data Collection	David A. Siegel, Taylor Ellington, Toye Williams, Reda Wilson Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA	Program: National cancer registry incidence data are typically not available for public analysis until 24-36 months post-diagnosis. As directed by the Caroline Pryce Walker Act (2008), the Centers for Disease Control and Prevention established the Early Case Capture of Pediatric and Young Adult Cancers (ECC) program. Nine US states during 2011-2019 were funded to capture new cases of cancer diagnosed in individuals aged 0-19 years within 30 days of diagnosis, as opposed to six months. ECC variables included patient demographics, tumor characteristics, and follow-back contact information for the patient and provider. Applicability: ECC data could be used by cancer registries or linked to clinical trial databases to ascertain diagnosis patterns of patients that could be targeted for clinical trials, and allow for changes in clinical trial recruitment efforts in a timely manner. Rapid case ascertainment could potentially help state registries identify patient needs for treatment, supportive care, or long-term follow-up. Potential for broader use: The Childhood Cancer STAR Act (2018) directs CDC to “enhance and expand infrastructure to track the epidemiology of cancer in children, adolescents, and young adults.” Methods of the ECC project will inform implementation of the Childhood Cancer STAR Act. Specific goals of implementation will be to improve timeliness and completeness of rapid case ascertainment data, expand the age range of rapid case ascertainment to include young adults, and scale the data collection beyond the current registries involved. Evidence of success and critical lessons learned: Seven states were funded under ECC in 2018. These states recorded 84%, 77%, 72%, 63%, 60%, 56%, and 44% of their cases by 30 days after diagnosis, and the majority of their cases were submitted to state registries 60 days after diagnosis. Improvement in timeliness was seen during the ECC program years. Experience from the ECC program will inform implementation of the Childhood Cancer STAR Act. Electronic reporting was found to be the most timely method of data transmission. However, sole reliance on electronic pathology reporting can decrease completeness because it misses clinically or radiologically diagnosed cases, and might not capture race/ethnicity data. Specific plans could be initiated to address these limitations.

Poster Number	Title	Authors and Affiliations	Abstract
65	Childhood Cancer Long-Term Follow-Up Program (LTFC)	Dennis Deapen[1], Amie Hwang[1], Mark Krailo[1], Wendy Landier[2], Smita Bhatia[2] [1]Department of Preventive Medicine, University of Southern California/Norris Comprehensive Cancer Center, [2]Department of Pediatrics, University of Alabama at Birmingham	It is well recognized that the survival rates for many childhood cancers have improved at a remarkable pace, such that nearly 85% survive 5 or more years, and the number of survivors in the US is approaching 500,000. However, long-term survivors are at risk for developing adverse outcomes including early death, second neoplasms, organ dysfunction, impaired growth and development, decreased fertility, impaired intellectual function, and overall diminished quality of life. Successfully contacting patients many decades following completion of their Children's Oncology Group (COG) treatment protocols is essential to adequately assessing the health and quality of life of childhood cancer survivors, especially since late effects may not be evident in these patients for many years. Long term follow-up of children poses unique challenges including special protections for minors, and name, family structure and lifestyle changes that come with young adulthood. Most COG institutions are unable to support long-term follow-up of their patients. The LTFC has developed a systematic, cost-effective mechanism for tracking and maintaining contact with childhood cancer patients enrolled in the program. New patients are enrolled monthly and identification and contact information updated annually to maintain currency of contact information and to re-establish contact with patients (or parents) when contact has been lost. The LTFC provides updated data to the COG Statistics and Data Center on a regular basis. This information will assist COG institutions in conducting follow-up examinations of their patients and collecting protocol-specific data. The LTFC will be a national resource, creating the capacity for clearer understanding of childhood cancer survivorship issues and providing many research opportunities. COG established the Umbrella Long-Term Follow-up Protocol (ALTE05N1). There are now 142 member institutions with IRB approval to enroll patients. One hundred twenty six of those institutions have enrolled a combined 3,360 patients currently with up to 8 years of active follow up. The LTFC is located at the Keck School of Medicine at the University of Southern California and is directed in collaboration with the University of Alabama at Birmingham Department of Pediatrics. This resource is available to any qualified researcher interested in conducting survivorship and outcomes research through COG.
16	Experience of Building a Pediatric	Donghan M. Yang[1], Laura Klesse[2], Bo Yao[1], Bo Ci[1], Shin-Yi Lin[1], Danni Luo[1], James F. Amatruda[2],	The Pediatric Cancer Data Commons (PCDC) is a platform of data collection, harmonization, and sharing for the pediatric cancer research community1. The

Poster Number	Title	Authors and Affiliations	Abstract
	Cancer Data Commons in Texas	Richard Gorlick[3], A. Christopher Menzies[4], Guanghua Xiao[1], Lin Xu[1], Lindsay A. Frazier[5], Patrick Leavey[2], Stephen X. Skapek[2], and Yang Xie[1] 1. Quantitative Biomedical Research Center, Department of Population & Data Sciences, UT Southwestern Medical Center, Dallas, TX 2. Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX 3. Division of Pediatrics, The University of Texas MD Anderson Cancer Center, Houston, TX 4. Children's Medical Center, Dallas, TX 5. Department of Pediatrics, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA	development started in 2018 with $5.3M support from the Cancer Prevention and Research Institute of Texas and is empowered by a team of experts with over 10 years' experience in developing cancer-related data analytics tools and platforms. Various types of pediatric cancer data are collected from multiple institutes and research programs in Texas. At the initiation stage, Children's Health System of Texas serves as the main Electronic Health Record (EHR) contributor, which operates the largest childhood cancer treatment center in North Texas. As part of the Malignant Germ Cell International Consortium (MaGIC), we have developed a clinical data standard to share clinical trial data for germ cell tumors across different countries2. We are collaborating with experts in ontology and biodata standards to expand it to encompass EHR data for other pediatric cancer types. Advanced artificial-intelligence-based Natural Language Processing tools (eg, BERT and CLAMP) were investigated to convert free-text medical notes into structured and standardized clinical variables. An Aperio VERSA 200 pathology scanner with brightfield and fluorescence modalities has been dedicated to digitizing Children's pathological tissue slides into whole slide images, which will be analyzed using the deep-learning-based analysis pipelines developed by our team3-14 and incorporated onto PCDC. Large volume and diverse types of genomics data generated by multiple research teams across Texas will be integrated and processed using our bioinformatics pipelines that have been developed over the past decade15-24. A biospecimen management system has been developed to connect patient-level clinical information with pathology imaging, genomics data and data from patient-derived xenograft (PDX); this system will be deployed in different institutes across Texas. To integrate and harmonize these data of various types and sources, i2b2-like data and information models are being optimized for childhood cancer, which will feature our long-established user-friendly online analytics and visualization tools for data query and analysis. The experience and tools generated in the PCDC project can be used for nation-wide data management and sharing demands for childhood cancer.
34	Supporting the Childhood Cancer Data Initiative through NCI Cancer Research Data Commons by enhancing large-	Erika Kim[1], Tanja Davidsen[1], Juli Klemm[1], Allen Dearry[1], Jaime Guidry Auvil[2], Zhining Wang[3], Chris Kinsinger[4], David Patton[1], Keyvan Farahani[1], Anthony Kerlavage[1] [1]Center for Biomedical Informatics and Information Technology, [2]Office of Data Sharing, [3]Center for Cancer Genomics, and [4]Center for Strategic	Basic and clinical research is increasingly focused on the generation of rich datasets to accelerate our understanding of childhood cancer and develop better treatment strategies. NIH has supported numerous programs including Therapeutically Applicable Research to Generate Effective Treatments (TARGET), Gabriella Miller Kids First Pediatric Research Program (GMKF), and Pediatric-MATCH to generate a wealth of data to be leveraged by the pediatric research and clinical communities and to test the effectiveness of therapies targeted to specific genetic changes. However, we are

Poster Number	Title	Authors and Affiliations	Abstract
	scale data analysis and data sharing	Scientific Initiative, National Cancer Institute, National Institutes of Health, Bethesda, MD	still limited in our ability to draw meaningful interpretations by challenges in integration across disparate datasets. To progress towards this goal, the pediatric research and clinical communities will need to access, aggregate, analyze, and interpret multi-modal data, including genomics, proteomics, imaging, and clinical data. NCI is investing in the informatics infrastructure to democratize access to and analysis of these diverse data types by developing the Cancer Research Data Commons (CRDC). The vision for the CRDC is a virtual, expandable infrastructure supporting data discovery, visualization, analysis, and sharing. CRDC allows researchers to securely access petabyte-scale data and to analyze, share, and store results, leveraging the storage and elastic compute of the cloud. Currently, operational components include Genomics Data Commons (GDC) and Proteomics Data Commons (PDC). Additional components, including Imaging Data Commons (IDC), Clinical Trials Data Commons (CTDC) and Integrated Canine Data Commons (ICDC) are under active development. Additionally, CRDC includes three Cloud Resources (CRs) that provide innovative computational and analytical platforms to enable large-scale, in-depth integrated data analysis. GDC and PDC currently host data from large genomic and proteomic programs including TARGET and Clinical Proteomic Tumor Atlas Consortium (CPTAC). TARGET provides rich molecular datasets from five major childhood cancer types and is available through GDC for the greater research community. PDC houses proteomic datasets from CPTAC as well as from a cohort of 199 tumors from Children's Brain Tumor Tissue. CRDC is currently available to the scientific community and CRDC will, in the future, interoperate with available childhood cancer data commons such as GMKF's Data Resource Center to further empower collaborative research to promote new discoveries to ultimately cure childhood cancer.
3	Unique features of the Ewing sarcoma transcriptome	Garrett T. Graham, Jeffrey A. Toretsky	Pediatric tumors are often driven by a limited set of mutations, and in some cases, as few as one major coding mutation has been identified. This is in contrast with adult tumors, which frequently carry orders of magnitude more coding mutations. These 'quiet' genomes require mining at higher resolution to detect targetable mutations for both small molecule pharmacology and immunotherapy approaches. Ewing sarcoma is characterized by a t(11;22) translocation, leading to the EWS-FLI1 fusion protein, but there are few other genomic events that would produce neoantigens. Because EWS-FLI1 is known to alter mRNA splicing patterns, we looked at mRNA splicing as an avenue for neoantigen production and patient stratification. Both

Poster Number	Title	Authors and Affiliations	Abstract
			endpoints require rich data collection and analysis within a framework that permits analysis of sequencing data without compromising data security. We have developed a strategy to identify potential discriminating coding features by analyzing mRNA splicing patterns in Ewing sarcoma patient tumor samples. Over the course of this project, we have used single-molecule real time (SMRT) long read sequencing of Ewing sarcoma to produce a detailed cancer-specific reference transcriptome. Our first objective is to establish a meaningful dataset of mRNA splicing variants for research into both splicing mechanisms and immunotherapy applications in Ewing sarcoma through analysis of existing tumor sequencing data. Thus far we have collected and analyzed splicing patterns at over 300k exons for altered variants in 109 Ewing sarcoma tumor samples with a reference comparison to 350 other solid tumors of childhood and AYA. Through this work we have identified a panel of novel exon skipping events that enable identification of Ewing sarcoma from all normal tissue tested thus far. We also expect that this information will be useful to researchers interested in exploring circulating cell-free RNA detection and quantification. One of the major benefits of information transfer at this analysis level is that, once processed, downstream steps would not require the same encryption on-disk and data protection required by raw sequence and mutation data. Additional samples and detailed clinical information will allow us to further refine this model, and potentially extend into other pediatric tumor types.
22	A collaborative pediatric cancer precision medicine model leveraging shared genomic, clinical, and research data from high-risk solid tumor patients across 45 clinical sites.	Giselle L. Saulnier Sholler[1,2,] Jacqueline M. Kraveka[3], William Ferguson[4,] Genevieve Bergendahl[1], William P. D. Hendricks[5], Abhinav Nagulapally[1], Sara A. Byron[5], James Lowey[5], William Burleson5, Elizabeth VanSickle[1], Muhammed Murtaza[5], Matija Snuderl [6], Timothy J Triche Jr [7], and Jeffrey M. Trent[5] on behalf of the Beat Childhood Cancer Consortium. [1]Helen DeVos Children's Hospital at Spectrum Health, [2]Michigan State University College of Human Medicine, [3]Medical University of South Carolina, [4]Saint Louis University School of Medicine, [5]Translational Genomics Research Institute (TGen), [6]New York University, [7]Van Andel Research Institute.	The Beat Childhood Cancer (BCC) Consortium is an international collaboration of academic hospitals/research institutions (40 in the United States, 4 in Canada and 1 in Lebanon). In collaboration with the Translational Genomics Research Institute, the BCC has developed a robust data-sharing infrastructure that integrates genomic, clinical, and discovery research data to empower advanced clinical trials, drive molecular discovery and hypothesis generation, and enable efficient distribution of data. BCC has enrolled 694 pediatric patients with relapsed, refractory, or high-risk solid tumors onto six clinical trials that have leveraged comprehensive genomic profiling (tumor-normal whole exome and/or RNA sequencing) to inform therapeutic decisions. The consortium also conducts orthogonal molecular studies such as epigenomic, proteomic, and circulating tumor DNA analyses. We have developed an open-source Precision Medicine Portal that integrates genomic, clinical and research data for facile analyses in molecular tumor boards composed of participants from geographically disparate sites. This portal enables integrated data analysis and timely development of treatment plans. Additional resources include established cell lines for high-throughput drug screening (>400 cell lines) as well as xenograft model

Poster Number	Title	Authors and Affiliations	Abstract
			development (>100 xenografts). Jointly, these resources comprise a suite of data and models to power pediatric cancer precision medicine studies that are shared within the BCC and the broader pediatric oncology research community. Key themes that have emerged from this work include: (1) The importance of re-biopsy at relapse due to clonal evolution and the shifting therapeutic landscape of tumors; (2) The identification of shared oncogenic events across diverse tumor pathologies, supporting the potential for tumor-type-agnostic targeting strategies; and (3) The value of multi-omic profiling in generation of a comprehensive molecular profile of pediatric tumors in order both to inform patient treatment and also to accelerate data-driven hypotheses for follow-on preclinical studies in matched patient-derived models. As precision medicine evolves, clinicians must consider many aspects of a patient's tumor biology at the time of treatment to make informed clinical decisions. The infrastructure for sharing data, samples, and model systems across hospitals and research institutions is needed and possible with collaborative efforts.
80	refine.bio: A resource for extracting knowledge from pediatric cancer data	Jaclyn N. Taroni, Kurt G. Wheeler, Richard W. W. Jones, Deepashree Venkatesh Prasad, Ariel Rodriguez Romero, Candace L. Savonen, Casey S. Greene Childhood Cancer Data Lab, Alex's Lemonade Stand Foundation, Philadelphia, PA USA	Millions of genome-wide gene expression assays are publicly available in repositories such as Gene Expression Omnibus and Sequence Read Archive. These petabyte-scale data are a powerful tool for pediatric cancer research, but obstacles for use remain. Samples were assayed on multiple technologies and platforms and it is often unclear how they were processed, leading individual researchers to invest significant time and computing resources in data reprocessing efforts. To address this challenge, we developed *refine.bio*: a multi-organism collection of uniformly processed and normalized transcriptomic data obtained from publicly available repositories. refine.bio allows researchers to select from hundreds of thousands of samples and to build datasets that are tailored to their question of interest. This permits biologists and clinicians to quickly validate results in different cohorts or from model systems. Relatively few publicly available samples directly assay pediatric cancers. Other biological systems, such as cell lines or model organisms, hold rich information for discovery in these conditions. We have previously trained unsupervised models on random collections of human RNA-seq data from diverse biological contexts; these models capture more biological processes and better describe rare disease datasets, even in the absence of the disease in the training set (Taroni et al. *Cell Systems*. 2019.). When the backlog of samples in government-run repositories are processed and made available in refine.bio, it will be an order of magnitude larger than prior resources and, with the inclusion of multiple assay modalities, particularly well-suited

Poster Number	Title	Authors and Affiliations	Abstract
			to the study of rare diseases. The breadth of refine.bio is highly complementary to pediatric cancer-specific curation efforts and will be valuable in the investigation of new treatments for pediatric cancer in cases where successful match to currently available therapies is not possible. Collections of all samples from a species processed by refine.bio, termed species compendia, are well-suited for training unsupervised models that may best describe developmental processes important for understanding the biology underlying pediatric cancers. refine.bio runs on Amazon Web Services. We have successfully processed over half a million samples, but the remote nature of cloud architectures has posed challenges. We will share what we learned processing sequencing data at this scale.
30	Building a Massive Oncology Health Data and Genomics Resource across Adolescent and Young Adult Patients Through Collaborative Partnership: Leveraging the ORIEN Consortium	Jeffrey Trent[1,2], Timothy McDaniel[1], Jill Kolesar[3], Tom C Badgett[3], Cornelia M Ulrich[4], Howard Colman[4], Mikaela Larson[4], Luke Maese[4], Joshua D Schiffman[4], Damon Reed[5], Michael Caligiuri[2], William Dalton[5,6] [1]Translational Genomics Research Institute, Phoenix, AZ, [2]City of Hope Comprehensive Cancer Center, Duarte, CA, [3]University of Kentucky-Markey Cancer Center, Lexington, KY, [4]University of Utah, Huntsman Cancer Center, Salt Lake City, UT, [5]Moffitt Cancer Center, Tampa, FL, [6]M2Gen, Inc., Tampa, FL.	Adolescent and Young Adult (AYA) patients with cancer have a large unmet clinical need due to their rarity and lack of participation in clinical and translational research, leading to unacceptably poor levels of relapse and survival. Genomics has a growing but unrealized capacity to transform cancer research, drug and diagnostics development, and patient care. A prerequisite to fully reaching this potential, including in the AYA community, is to develop a massive and accessible data resource containing genomic information, medical and family histories, and outcomes for very large numbers of consenting cancer patients. Building this resource at sufficient scale has been hindered by a variety of operational and administrative barriers, which we have overcome through ORIEN, the Oncology Research Information Exchange Network. ORIEN is an alliance of 19 U.S.-based cancer centers founded on the principles of Collaboration, Inclusiveness, Data accessibility, and Partnerships. Under a single protocol, known as Total Cancer Care (TCC), patients provide lifetime consent for longitudinal, prospective study, contributing clinical information, tissue specimens, and molecular data, and allow proactive matching and recontact to facilitate clinical trials based on molecular profile. To date over 215,000 cancer patients have consented, including 14,561 AYA patients (ages 18-39) across 16 ORIEN sites. In addition, the Markey Cancer Center began enrolling pediatric cancer patients (ages 0-18) to TCC in April 2019, and currently has seven pediatric patients on-study. Molecular profiling across the network is performed using next generation sequencing to generate deep tumor/normal exome and tumor whole transcriptome data. As of 2019, over 9,000 patients have been profiled by NGS. The de-identified patient data are made available to each ORIEN member for research, and to participating pharmaceutical companies for collaborative research studies. The TCC

Poster Number	Title	Authors and Affiliations	Abstract
			protocol can create *in silico* patient communities to identify those with similar molecular and clinical profiles to facilitate study enrollment. Thus, ORIEN constitutes a vast infrastructural and data resource for genomic oncology to the benefit of all stakeholders, including cancer centers, pharmaceutical companies, and patients. This model could be readily expanded to more systematically serve the pediatric and AYA communities at the national level.
27	Scalable Consortium Approaches Facilitate Cooperative Research into Childhood Cancer Etiology and Outcomes	Jeremy M. Schraw[1], Olga A. Taylor[2,3], Karen R. Rabin[2,3,4,] Philip J. Lupo[2,3,4], and Michael E. Scheurer[2,3,5] 1. Department of Medicine, Baylor College of Medicine, Houston, TX USA 2. Department of Pediatrics, Baylor College of Medicine, Houston, TX USA 3. Texas Children's Cancer and Hematology Centers, Texas Children's Hospital, Houston, TX USA 4. Reducing Ethnic Disparities in Acute Leukemia Consortium 5. Adolescent and Childhood Cancer Epidemiology and Susceptibility Service for Texas	Rationale: Cancer is a leading cause of death by disease in children and adolescents/young adults (AYAs). Important gaps that limit research into childhood cancer etiology and outcomes include: 1) lack of population-based resources collecting both biological samples and epidemiologic questionnaire data on newly diagnosed cases; and 2) limited cohort studies of multi-ethnic populations collecting detailed clinical information and biological samples. We highlight two ongoing integrated consortia attempting to address these gaps. Applicability: The Adolescent and Childhood Cancer Epidemiology and Susceptibility Service for Texas (ACCESS-Texas) was established to facilitate rapid and complete ascertainment of childhood and AYA cancer cases, perform biobanking of multiple biospecimens from patients and family members, and systematically centralize and harmonize clinical and epidemiologic data. ACCESS-Texas recruits approximately 800 cases annually from all major pediatric cancer treatment centers in Texas and its infrastructure supports numerous research efforts including the Reducing Ethnic Disparities in Acute Leukemia (REDIAL) Consortium. REDIAL is focused on investigating the clinical and genetic features underlying disparities in outcomes in Latino children with acute leukemia. The REDIAL Consortium is anticipated to enroll a total of 1,000 children with acute leukemia, and will include genome-wide genotyping and collection of serial blood and bone marrow samples during therapy, linked to well-annotated clinical and epidemiologic data collected and managed through ACCESS-Texas. Potential for Broader Use: ACCESS-Texas and REDIAL are efficient and scalable mechanisms for accelerating research into childhood cancer etiology and outcomes. They integrate multiple data sources through a connected data infrastructure. Data and specimens collected by these consortia are a tremendous resource for childhood cancer researchers. This approach is broadly applicable to the study of childhood and adult cancer, as well as other rare diseases.

Poster Number	Title	Authors and Affiliations	Abstract
			Evidence of Success: Since their inception, ACCESS-Texas and REDIAL have collectively enrolled more than 2,000 patients and 1,500 family members, and provide centralized management of associated data and biospecimens. Importantly, these resources facilitate research into both etiology and outcomes within a single patient population. They support an array of research, including studies of ethnic differences in acute leukemia risk and outcomes; tumor and germline genomic sequencing; metabolomic profiling for biomarker discovery, and more.
8	Feasibility of Linking Childhood Cancer Registry Data and Civil Registries in São Paulo, Brazil: Developing Tools to Find the Clues	Karina B Ribeiro, Carolina T M Luizaga, Rosa M V Freitas, Lilian C C Morais, Valmir J Aranha, Monica L Teixeira, Célia B G Antoneli, Bernadette C Waldvogel, José Eluf Neto	Background: Perinatal exposures have been associated to childhood cancer, but few studies with large samples have been conducted in Latin America. We aimed to investigate this association through the linkage between cancer and civil registries databases in São Paulo, Brazil. Design/Methods: The Seade Foundation is responsible for the Civil Registry Statistics System, based on a monthly survey in the Civil Registry Office, collecting information on marriages, live births, stillbirths, and deaths, covering all municipalities of São Paulo. The Central Hospital Cancer Registry from the state of São Paulo (RHC-SP) was established in 2000, maintained by Oncocentro Foundation (FOSP), and currently comprising 74 cancer hospitals affiliated with the Brazilian Universal Health Care System (SUS). As of March 2019, RHC-SP includes 849,359 analytical cancer cases, holding excellent quality indicators. For the linkage, we have selected all cases of acute leukemia and non-CNS embryonal tumors, diagnosed between 2001 and 2019 among children with age $\leq$ 5 years (n=3,199). These cases were linked to live births from 2001 to 2016 (9,931,861 records, average number/year=620,741). Linkage was deterministic, based on standardized variables and 20 algorithms derived from different combinations of percentages of similarity of the child's name, mother's name, and date of birth. Visual checking of the possible true pairs was performed by two investigators, considering additional variables (home address, date of the first medical visit, date of cancer diagnosis, date of the last follow-up and vital status). Results: A total of 2,830 pairs were successfully linked (88.5% of the initial database), with 2,233 (78.9%) based on the exact matching of the 3 identifiers. The final matched database includes demographics, variables relative to cancer diagnosis and treatment, as well as maternal education, maternal age, race/skin color, gestational age, plurality, delivery method, number of live births, stillbirths, number of prenatal consultations, birthweight, congenital malformations, and APGAR scores 1 and 5 minutes.

Poster Number	Title	Authors and Affiliations	Abstract
			Conclusions: Linkage was feasible, achieving robust rates. Matched database will now serve to build a case-control study; further access to the biospecimens of the neonatal screening program is also an alternative, in order to investigate etiological and prognostic factors for childhood cancer in large scale.
21	Generation of a real-time sample tracking and data sharing system for a multi-institution project to develop patient derived xenografts in pediatric cancer.	Klesse, L.J.[1], Yang, D.[2], Butler, E.[1], Lin, S.[2], Xu, L.[2], Yao, B.[2], Kurmasheva, R.[3], Rogojina, A.[3], Bandyopadhyay, A.[3,] Houghton, P.[3], Xie, Y.[2] and Skapek, S.[1] [1]Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX [2]Quantitative Biomedical Research Center, UT Southwestern Medical Center, Dallas, TX [3]Greehey Children's Cancer Research Institute, UT Health Science Center, San Antonio, TX	In spite of clear advances in cure and survival over the past three decades, cancer remains the leading cause of disease related death in children and teens in the United States. For several tumor types, particularly those with limited response to initial therapy or metastatic disease, survival rates continue to be poor. For these children, novel therapies are clearly needed. Given the overall rarity of childhood cancer, a reproducible experimental system is paramount for testing of novel therapies. A consortium of clinical and laboratory sites in Texas was established to generate, characterize, verify and catalogue patient derived xenografts (PDXs) of childhood cancer. Since its inception in 2017, the consortium has enrolled over 200 patients and generated robust PDXs from over 1/3 of the enrolled patient tumors. The consortium has collected primary tumor samples, PDX samples and germ line samples from each patient, compared DNA, RNA signatures, histology and methylation patterns. A web-based cataloguing and real-time tracking system has been developed to help collate samples from and between the different institutions. The system also serves as a tool for patient de-identified tumor queries for future experimental projects as it records patient-level clinical variables, such as basic demographic, diagnosis and outcome measures. The sample processing and transaction history is automatically tracked by the linkage between each original biospecimen and all subsequently derived DNA, RNA, and xenograft samples. The availability, quantity, and location of a sample can be updated across institutions in real time. Attachable sample barcodes can be directly printed from the system and later scanned to retrieve specific samples from the catalogue. Variables for medical terminologies and sample preparation procedures were standardized based on ICD-O-3 and TCGA codes, ensuring compatibility in data sharing between institutions. A user-friendly, interactive cohort discovery tool is provided in the web portal, which allows patient and sample selection based on interested variables.
4	Development of Cancer Database Platform for Clinical and Research Utilization from	Lilibeth Torno, Ivan Kirov, Elyssa Rubin, Van Huynh, Josephine HaDuong, Carol Lin, Rishikesh Chavan, Ashley Plant, Chenue Abongwa, Jamie Frediani, Richard Chamberlain, Kevin Bostwick, Colette Bruce, Christine Yun	In a pilot project at CHOC Children's, we designed a comprehensive database with datapoints geared to pull pertinent informational and clinical data for pediatric/adolescent young adult (AYA) cancer survivor's treatment summaries. Relevant information included demographics, ethnicity, diagnosis and genetic predisposition syndromes. Treatment modalities such as surgery, radiation,

Poster Number	Title	Authors and Affiliations	Abstract
	Diagnosis to Survivorship	CHOC Children's Hospital Division of Oncology 1201 W. La Veta Orange, CA 92868	chemotherapy/immunotherapy agents and doses were included. Formulas were designed to calculate cumulative anthracycline and alkylator doses. Furthermore, treatment complications and late effects were included. Data entry was initiated at diagnosis, followed by multiple key time points throughout the course of treatment to document recurrences, change in treatment regimens or adverse events. From the database, clinical tools for tracking treatment toxicities, use of new and innovative treatments, assessing mental health metrics, tracking central line use and its relationship to risk of bacteremia and several other Quality Improvement projects have been developed. The database has also been utilized by case management and coordinators to bridge the complexity of navigating the healthcare system. The database is highly utilized in research such as in a current multi-institutional study tracking the use of pre-medications for Asparaginase in leukemia patients as well as in cancer survivorship research. The development of this database has demonstrated utility in improving research efforts, clinical care, communication, ability to bridge the gap between healthcare providers as well as the dyad of provider and survivor. Treatment summaries, letters to providers, and data reports have been generated from the database. The design will be further developed to model its utility in a multi-institutional collaborative setting.
33	Pediatric Normal Tissue Effects in the Clinic (PENTEC): an international collaboration to analyze normal tissue radiation dose-volume-response relationships for pediatric cancer patients	Louis S. Constinea, Cécile M. Ronckersb,c, Chia-Ho Huad, Arthur Olche, Leontien C. M. Kremerb,c, Torunn I. Yockf, Andrew Jacksong, Soren M. Bentzenh [a] University of Rochester Medical Center, Rochester, NY, USA [b] Emma Children's Hospital/Academic Medical Center, Amsterdam, the Netherlands [c] Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands [d] St. Jude Children's Research Hospital, Memphis, TN, USA	Purpose: PENTEC (Pediatric Normal Tissue Effects in the Clinic) is a volunteer research collaboration of more than 150 physicians, medical physicists, biomathematical modelers, and epidemiologists organized into 18 organ-specific task forces conducting critical reviews and synthesis of published toxicity data. PENTEC aims to (1) establish quantitative, evidence-based dose/volume/risk guidelines to inform radiation treatment decision-making to improve efficacy-toxicity outcomes; (2) explore the most relevant patient-level risk factors for toxicity; (3) delineate dose-volume constraints relevant to pediatric radiotherapy plan optimization; and (4) propose dose-volume-outcome reporting standards for publications on childhood cancer outcomes. Methods: Comprehensive PubMed searches were performed in each site-specific task force and data was abstracted from pertinent studies and incorporated into mathematical models. This usually consisted of a dichotomous specific endpoint

Poster Number	Title	Authors and Affiliations	Abstract
		[e] University of Southern California Keck School of Medicine and Children's Hospital of Los Angeles, Los Angeles, CA, USA [f] Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA [g] Memorial Sloan-Kettering Cancer Center, New York, NY, USA [h] University of Maryland School of Medicine, Baltimore, MD, USA	evaluated as a function of dose, volume of the specific organ, age, and other possible contributing variables specific to that site. For example, hearing loss (yes/no) was evaluated according to mean radiation dose to the cochlea and by age, stratified by platinum-based chemotherapy exposure. Results: Eighteen task forces are in varying stages of the methodology, from data acquisition, modeling, and reporting. Educational sessions, status report updates, and task force-specific reports have been and will be presented at American Society of Therapeutic Radiation Oncology (ASTRO) and Pediatric Radiation Oncology Society meetings. Each PENTEC task force is finding significant variability in the completeness and manner that data are presented and analyzed, thus posing challenges when synthesizing data from multiple studies. Each site-specific task force will be recommending a standard approach to the collection and reporting of data to improve outcome research necessary to inform radiotherapy decision making and dose-volume planning constraints. Conclusion: PENTEC has revealed a multiplicity of challenges in exploring and defining normal tissue tolerances in developing children as a function of radiation organ dose/volume, chemotherapy exposure, and age. This work provides a valuable starting point for defining a uniform ontology for recording and grading of treatment adverse effects, as well as critical recommendations for standardized reporting of dose/volume metrics that will allow future refinement of radiation dose guidelines in pediatric cancer patients through analysis of large data registries.
11	Standardizing Data Sharing in Pediatric Oncology Research	Meghan McCormick, MD[l] and Louis Rapkin, MD[l] [l]University of Pittsburgh Medical Center Children's Hospital of Pittsburgh	Over half of pediatric patients diagnosed with cancer in the United States are enrolled and treated on clinical trials. The comprehensive data collection and rapid enrollment of large groups of patients made possible through cooperative group sponsored clinic trials has contributed to rapid advancements in clinical care and associated improvements in survival. However, for the large number of pediatric patients who are not enrolled on clinical trials, the valuable information obtained from their experiences is often not shared with the scientific community. This is particularly troubling in the case of rare oncologic diagnoses. Often clinical trials are not available for these rare diseases and only a small number of patients carrying that diagnosis may be seen at each institution. This can make determination of the best course of action based on personal or collective experience difficult. Administrative databases offer the capability to complete multi-center research, although limitations include undercoding and often inadequate clinical information is available.

Poster Number	Title	Authors and Affiliations	Abstract
			We propose to address this issue through the development of institutionally maintained patient registries that may be linked to facilitate research endeavors. We will complete a comprehensive literature review to identify existing clinical databases aimed at pediatric oncology, in order to determine the strengths and weaknesses of existing infrastructure. We will then contact member institutions of the children's oncology group (COG), the largest cooperative children's cancer research group, through the primary investigator of the institution's COG research program. Through standardized phone interviews we will collect information on the systems currently in place within each institution for clinical data collection, the data elements collected and each institution's sharing policies for clinical data. Lastly, we will utilize an electronic survey system to query COG members on the data elements they would find most critical to conducting research. Our ultimate goal is to utilize this information to develop an electronic registry which will collect de-identified standardized data elements. In our proposed model, each institution would independently maintain the registry within their institution, though the standardized format would allow for easy merging of datasets from multiple institutions for approved research questions.
18	Deep Proteogenomic Survey across Seven Childhood Brain Tumors: A applying CBTTC/CPTAC Pilot	Pei Wang, Nicole Tignor, Steven Gygi, Richard G. Ivey, Amanda Paulovich, Alexey Nescizhskii, Sanjukta Guha Thakurta, Jeffery R. Whiteaker, Jacob J. Kennedy, Uliana J. Voytovich, Li Ding, Liang Bo Wang, Pichai Raman, Yuankun Zhu, Tara Hiltke, Henry Rodriguez, Brian R. Rood, and Adam Resnick On behalf of the NCI's Clinical Proteomic Tumor Analysis Consortium (CPTAC) and the Children's Brain Tumor Tissue Consortium (CBTTC). The Children's Hospital of Philadelphia, Children's National Medical Center, Icahn School of Medicine at Mount Sinai, Fred Hutchinson Cancer Research Center, Harvard Medical School, Washington University School of Medicine	Genomic characterization has allowed for the differentiation of different tumor types based upon the abundance of gene transcripts. However, owing to the many layers of regulation between transcript and the post-translationally modified protein, it has been challenging to extrapolate biology from transcriptional differences alone. Working with the NCI's Clinical Proteomic Tumor Analysis Consortium (CPTAC) and a consortium of more 18 partnered institutions comprising the Children's Brain Tumor Tissue Consortium (CBTTC), we hypothesized that a comparative analysis of the proteome and phosphoproteome across a pilot initiative of 7 childhood brain tumors would yield a deeper understanding of differences in functional biology in ways that could be harnessed for clinical translation. We performed tandem mass tag labeling and triple mass spectrometry of 226 fresh frozen tumor samples representing histologic diagnoses of: high grade astrocytoma (27), low grade astrocytoma (97), ganglioglioma (20), ependymoma (32), medulloblastoma (22), atypical teratoid rhabdoid tumor (12), and craniopharyngioma (16). Among these samples were 22 pairs from pre/post recurrence. Across this sample set, we quantified 9155 proteins and 13632 phospho sites. After data preprocessing, consensus clustering showed that protein profiles effectively distinguish major histology types. Additional regression-based analyses revealed groups of proteins and pathways showing distinct activities across histologies and clinical features/outcomes. Further leveraging tumor/normal

Poster Number	Title	Authors and Affiliations	Abstract
			WGS, RNAseq data and longitudinal clinical data collected under the CBTTC, the functional impact of mutations and fusions was assessed, focusing initially on low grade glioma and the proteomic ramifications of BRAF V600E, BRAF-fusion and wild type BRAF. Moreover, from the primary and recurrent tumor pairs, the upregulation of proteins associated with immune evasion was also identified in more advanced LGG tumors. Building on this study, the consortia are further expanding studies across the age continuum, beginning with an integrated prospective and retrospective cohort of pediatric, adolescent and young adult gliomas, and adult glioblastoma multiforme (GBM). These efforts, supported by the emerging network of interoperable data-sharing and collaborative cloud-based NIH/NCI platforms provide for a first-in-kind multiomic analysis that includes the proteomic dimension alongside large scale genomic data, providing new functional insights that are poised to accelerate clinical translation.
38	Critical Need for Novel Data Sources to Evaluate End-of-Life Care Quality for Children with Cancer	Prasanna Ananth, MD, MPH[1,2;] Sophia Mun, MPH[2]; Tannaz Sedghi, MPH[2]; Randall Li, BA[3]; Cary Gross, MD[1,2;] Xiaomei Ma, PhD[2],[4;] Joanne Wolfe, MD, MPH[5]. [1]Yale School of Medicine, New Haven, CT. [2]Yale Cancer Outcomes, Public Policy and Effectiveness Research Center, New Haven, CT. [3]The Warren Alpert Medical School of Brown University, Providence, RI. [4]Yale School of Public Health, New Haven, CT. [5]Department of Psychosocial Oncology and Palliative Care, Dana-Farber Cancer Institute; Department of Pediatrics, Boston Children's Hospital, Boston, MA	Background: Without standardized approaches to providing high quality end-of-life (EOL) care for children, adolescents, and young adults (AYAs) with cancer, many receive intensive services near the EOL that may heighten patient and family suffering, without offering benefit. Objectives & Applicability to Childhood/AYA Cancers: We sought (1) to elicit EOL care priorities from key stakeholders; and (2) to develop a framework for future data collection that enables systematic measurement of high quality EOL care in the context of childhood cancer. Methods: In a multi-center qualitative study at Yale Cancer Center and Dana-Farber Cancer Institute, we interviewed AYAs with advanced cancer, parents of children with advanced cancer, and bereaved parents (n=23). We also conducted focus groups and in-depth interviews with interdisciplinary providers (n=19). We explored participant priorities for a child with cancer near the EOL and whether they perceived existing quality measures for adults with cancer to be relevant to pediatrics. Results: Participants prioritized optimal symptom control near the EOL, direct communication with the affected child, and involvement of interdisciplinary providers (e.g. child life, psychosocial clinicians). Participants reflected on the importance of access to the hospital for symptom management or supportive care, diverging from adult quality measures wherein minimizing hospital use is preferred. Several identified the use of chemotherapy near the EOL to be appropriate, particularly if chemotherapy might relieve symptoms. Most participants preferred to avoid cardiopulmonary resuscitation or mechanical ventilation near the EOL. Hospice was

Poster Number	Title	Authors and Affiliations	Abstract
			valued by some participants, but providers commented on limited availability of pediatric hospice expertise. Potential for Broader Use in Research & Lessons Learned: Key stakeholders identified several priorities that define high quality EOL care in childhood cancer. Administrative datasets may capture some of these quality measures. However, more nuanced elements of high quality EOL care, such as symptom control, necessitate novel approaches to data acquisition. This may entail employment of patient- and parent-reported outcome tools to evaluate symptoms, combined with unique strategies, such as natural language processing, that harness the content of electronic health records. Integrating data sources would facilitate more person-centered evaluation of quality benchmarks in EOL care and development of interventions to enhance care.
9	Universal deployment of a multi-functional information system for patient care and research improves childhood cancer care and outcomes	Scott C. Howard, MD, MSc	Background Children with cancer are cured more than 80% of the time in high-income countries (HIC), but less than 30% of the time in low- and middle-income countries (LMIC). Improving survival requires massive reductions in preventable treatment failure in LMIC by addressing non-diagnosis, misdiagnosis, lack of access to treatment, abandonment, toxic death, and excess relapse. It also requires acceleration of research in all countries by increasing access to clinical trials and the speed of testing new treatment strategies. High-quality information systems are critical for progress in both LMIC and HIC. Program description The Resonance Patient Center (www.ResonanceOncology.org, RPC) provides a fit-for-purpose pediatric cancer registry; pathology module to address information needs for cancer diagnosis; patient tracking module to prevent missed appointments and abandonment; toxicity module to address toxic death; quality improvement (QI) module to allow clinicians to implement their own hospital-specific QI programs for specific issues (e.g. febrile neutropenia management); a chemotherapy roadmap that calculates doses and can be used to promote adherence and facilitate forecasting of drugs quantities needed; a clinical trials management system to enable research; and a palliative care module. Data is stored in the Amazon HIPAA-compliant cloud using end-to-end encryption with password protected access and transaction logging. Data sharing is controlled by the system administrator at each center and uses an anonymization engine that automatically removes protected health information (PHI). The system is provided at no cost to LMIC centers.

Poster Number	Title	Authors and Affiliations	Abstract
			Applicability to pediatric cancer Currently, RPC stores information about 24144 children with cancer from 46 countries; uptake of the system is now growing exponentially and will soon constitute the global standard for pediatric oncology information systems. A global quality improvement project to reduce toxicity from high-dose methotrexate launched recently. The system supports individual centers, specific projects, daily healthcare operations, and clinical trials networks. Conclusions RPC provides the clinical information system on which tissue banks and specific projects can be built. For example, the consult system for second opinions by clinical, pathology, and radiology experts launches soon.
89	Improving pediatric cancer patient follow up data collection through smartphone-based short health status check-ins using REDCap	Susan L. McGovern[1], Sara L. Gallotto[2], Benjamin V. Bajaj[2], Miranda P. Lawell[2], Daniel Indelicato[3], Arnold Paulino[1], William Hartsell[4], Nadia Laack[5], Ralph Ermoian[6], Ralph Vatner[8], Stephanie Perkins[9], Victor Mangona[10], Christine Hill-Kayser[11], Suzanne Wolden[12], Young Kwok[13], Michael Confer[14], J. Ben Wilkinson[15], Torunn I. Yock[2] [1]Department of Radiation Oncology, MD Anderson Cancer Center, [2]Department of Radiation Oncology, Harvard Medical School, [3]Department of Radiation Oncology, University of Florida, [4]Department of Radiation Oncology, Northwestern Medicine Chicago Proton Center, [5]Department of Radiation Oncology, Mayo Clinic, [6]Department of Radiation Oncology, University of Washington, [7]Division of Oncology, Cincinnati Children's Hospital Medical Center, [8]Department of Radiation Oncology, Cincinnati Children's Hospital Medical Center, [9]Department of Radiation Oncology, Washington University in St. Louis, [10]Department of Radiation Oncology, Texas Center for Proton Therapy, [11]Department of Radiation Oncology, University of	Background: Radiation therapy (RT) is an integral component of curative treatment of many childhood solid tumors but also a major contributor to late morbidity for long-term survivors. Proton radiotherapy (PRT) is gaining prominence to diminish late effects of treatment in children and there are now 31 centers operating in the US. However, PRT treatment centers are usually quaternary (referral) centers, and patients often return to their home/referring institution for follow-up care. The difficulty in obtaining comprehensive clinical follow-up has manifested into a resource-intensive challenge for our multicenter Pediatric Proton/Photon Consortium Registry (PPCR). For that reason, we have piloted a Patient/Parent-proxy reported outcomes (PROs) health status survey at one of our PPCR sites to help alleviate the burden of follow-up data collection. Design: A PROs health status check-in survey was developed in REDCap which allows participants to securely complete online surveys which deposit directly into the REDCap database. The survey questions include current contact information, disease and health status, recent medical care, and social information including education or employment status. The survey is designed to be completed in under 5 minutes on a smartphone. A survey non-response or an affirmative response to any health question indicates a change in health status (i.e. new diagnosis, or new medication) which triggers a review of the electronic medical record (EMR). The PRO screens for patients with a change in health status and reduces the number of medical records requiring review, thereby saving person-hours or expense for data collection and input. Results: To date, 170 registry participants have received a mailed letter containing a QR Code survey link which can be opened with a smartphone camera and 38 have

Poster Number	Title	Authors and Affiliations	Abstract
		Pennsylvania, [12]Department of Radiation Oncology, ProCure Proton Therapy Center, New Jersey, [13]Department of Radiation Oncology, University of Maryland, [14]Department of Radiation Oncology, ProCure Proton Therapy Center, Oklahoma City, [15]Department of Radiation Oncology, Provision Healthcare, [16]Department of Radiation Oncology, Johns Hopkins University, [17]Department of Radiation Oncology, Emory University	responded (22.4%). To improve the response rate, the next pilot will use text messaging (avoiding the mail) for easy mobile access. Conclusion: PROs hold great promise in improving efficiency of data collection in our large referral proton pediatric population in the context of a limited resources for medical record review and input. Further evolution and testing of this approach are warranted to maximize collection of accurate patient follow-up data.
23	Clinical and Translational Data Processes and Workflows for a Pediatric Genomic Profiling Protocol	Tara Lichtenberg, Catherine E. Cottrell, Vincent Magrini, Kathleen Schieffer, Elizabeth Varga, Susan Vear, Kristen Leraas, Katherine Miller, Stephanie LaHaye, Amy Wetzel, Daniel Koboldt, Ben Kelly, James Fitch, Patrick Brennan, Gregory Wheeler, Peter White, Ruthann Pfau, Selene Koo, Julie Gastier-Foster, Richard K. Wilson, Elaine R. Mardis.	The Institute for Genomic Medicine (IGM) at Nationwide Children's Hospital has developed a translational protocol to evaluate the genomic landscape of tumors and blood disorders among patients with rare or treatment refractory disease. Patients enrolled in this study receive disease-involved/comparator-normal whole exome sequencing and whole transcriptome analysis of the disease involved specimen. To date over 100 patients have been nominated for this protocol, 77 have been consented and 68 have completed sequencing results. Central to the protocol is a multifunctional REDCap database designed to streamline workflows including enrollment, consenting, specimen selection, assay methodology and the documentation of results. Data collection begins when a clinical provider nominates a patient using a REDCap survey. This initial survey kicks off a series of forms completed by the IGM team as well as other groups within the hospital. Each form in the series uses information from previous forms to track the overall status of the patient as they move through the IGM workflow from consent to the discovery of medically meaningful findings. The implementation of the REDCap database has improved the flow of information among IGM team members as well as communication with patient care providers. In addition to tracking the patient through the workflow of the protocol, we are abstracting a dataset from the patient's medical record. This dataset is being recorded in REDCap where it is harmonized and standardized using classifications such as the World Health Organization's classification of diseases. These data are used in parallel with the genomic data during analysis.

Poster Number	Title	Authors and Affiliations	Abstract
			The organization and management of all data generated as a result of this protocol is important to IGM, as we are broadly committed to sharing data. We are actively submitting to the Treehouse Childhood Cancer Initiative where we have sent data from 68 NCH patients, including 35 patients from this protocol. We are also in the process of submitting sequencing and phenotype data to the Database of Genotypes and Phenotypes (dbGaP). We continue to improve data organization and are creating new ways to visualize the data using the Qlik Sense application. In the future, as additional data standards become more widely recognized, we plan to further harmonize our data to be compatible with developing models.
37	Patient Advocacy Call to Action for Childhood Cancer Data and Biospecimens	The members of the Alliance for Childhood Cancer (Alliance) and the Coalition Against Childhood Cancer (CAC2) recognize the possibilities to enhance the collection, organization, and use of biological samples and corresponding clinical and demographic childhood cancer data afforded by coordinating provisions of the Childhood Cancer Survivorship, Treatment, Access and Research (STAR) Act and the emerging Childhood Cancer Data Initiative.	Issues with Current System The National Cancer Institute (NCI) supports a variety of networks, through different funding mechanisms, related to collaborative pediatric oncology clinical research that involve both biorepository and data repository activities. Because specimen and management and quality control procedures can vary between and across networks, investigators trying to answer research questions can find it difficult to progress their work quickly and efficiently. Improve Quality Control of Biospecimens and Clinical Data Advances in childhood cancer research depend on specimen and data quality control and accuracy. Accordingly, available resources should enhance staff capacity to ensure that data management and specimen handling procedures follow the highest possible quality guidance and standards. Funding should be used to guarantee that samples are appropriately coded, annotated, and checked for quality before they enter a biorepository. Develop Uniform Coding To improve researchers' access to quality data, a coding system or adoption of common application programing interfaces (APIs) that would standardize molecular and clinical data characterization should be developed and implemented. Ideally, the

Poster Number	Title	Authors and Affiliations	Abstract
			solution would be uniform both within and across network biorepositories to ease researchers' data queries of different sources. Providing Funding for Long-Term Follow-Up Patients' privacy must be maintained through individually unique identifiers and made secure over the course of data collection, transmission to, and storage in a repository. With these identifiers in place, funding should be provided so that patients could be tracked throughout their treatment course and followed up annually for ten years to expand the usefulness of each patient's biospecimen and clinical data. Improve Researchers' Access to Data A clear and easy online application system should be developed so that researchers' requests of data are responded to expeditiously, consistently and accurately.
29	None specified	Timothy Triche	While major advances in the understanding and treatment of adult cancer have eventuated from large scale cancer genomic profiling initiatives like TCGA, there is no such broad genomic database for cancer in the young. At Children's Hospital Los Angeles, part of the NCI supported USC Norris Comprehensive Cancer, we seek to redress that issue across all types of cancer in the young. We created a childhood cancer genomic profiling panel, OncoKids (aka Oncomine Childhood Cancer Research Assay, OCCRA), modeled after the NCI MATCH program Oncomine Comprehensive Assay (OCA). Both panels use FFPE or fresh tumor tissue, interrogate both DNA and RNA, and were developed in collaboration with Thermo Fisher. In contrast to OCA, OncoKids incorporates all known DNA mutations, amplified genes, gene fusions, and over-expressed genes of clinical, diagnostic, prognostic, or therapeutic importance. Over 400 cases of liquid, solid, and brain tumors from multiple institutions were used to validate the CAP/CLIA certified LDT. In use for over a year at CHLA, essentially every patient (~600 per year) is profiled prior to treatment. Results are discussed bi-weekly in a Molecular Pathology conference attended by oncologists and pathologists. All accrued data, as well as all publicly available genomic data, are archived in a growing genomic database of childhood cancer, Childhood Cancer Knowledge Base (CCKB), hosted at CHLA and available throughout the world. A comprehensive suite of analytic tools for data mining suitable for both clinical and

Poster Number	Title	Authors and Affiliations	Abstract
			research analysis has been developed to analyze the data which are available to anyone who elects to join the international collaborative network ICON (International Childhood Oncology Network), sponsored by Thermo Fisher. Over 30 institutions worldwide have joined ICON and will have access to the data, analytic tools, and expertise hosted at CHLA. Over 65% of patients have clinically relevant genomic defects, 15% show evidence of germline inheritance, and in several instance, critically important treatment decisions have been based on the results, available in emergent circumstances within 48 hours. Ultimately, the goal is to create an international collaborative consortium that will ensure that children with cancer will receive optimal treatment and outcomes, regardless of their geographic location.
40	Integrated Pediatric Omics Web Resource to Explore Genomic Data and to Enable Genome Guided Precision Therapeutics	Wei JS, Hsien-Chao C, Tyagi M, Wen X, Khan J. Oncogenomics Section, Genetics Branch, Center for Cancer Research, National Cancer Institute. Bethesda, MD 20892	Cancer is a complex, multifactorial disease that is associated with aberrations in chromosomal ploidy, DNA copy number, epigenetic, mRNA and protein expression profiles, and altered signaling networks. In recent years, the application of high-throughput genomic and proteomic technologies, such as microarrays, next-generation sequencing, and mass spectroscopy, to the study of cancer has greatly improved our understanding of the molecular machinery underlying this disease. These studies, broadly defined as genomics and proteomics, have uncovered several biological themes. Firstly, that cancers of different histological types have diagnostic specific gene expression profiles and somatic mutational spectra. Furthermore, cancers can be more accurately classified into clinical, prognostic, and molecular sub-groups by these methods than routine histology. Finally, functional genomics including siRNA screening and pathway analyses, through the enumeration of mutated genes, has implicated modules within cellular circuitry that seem to be essential, across all cancers, for disease initiation and progression. We have developed the Oncogenomics Web Resource to support the use of genomics as a tool to explore the biological underpinnings of cancer and drive genome guided precision therapy. The Oncogenomics web resource provides users with a single-point access to a catalog of data sets generated from microarrays, next-generation sequencing (NGS)-based RNAseq, and mass-spectrometry, for a wide range of cancers and normal tissues and includes data from cell lines, multiple species, and model organisms. Registered users can query each data-set to access data on their genes or functional groups of interest; query results include heat-map and bar chart representations of experimental data as well as link-outs to more detailed annotations. As all data-sets

Poster Number	Title	Authors and Affiliations	Abstract
			are pre-normalized, users can perform cross-experimental comparisons on their gene sets of interest. All query results can be downloaded as text files. For gene expression data-sets the site enables users to run Gene set enrichment analyses (GSEA) against lists of both curated and custom gene sets. For datasets with patient survival data, users can view Kaplan-Meier plots and the results of pre-run log-rank tests. Finally, we have developed a patient centric database that integrates germline, somatic, RNAseq, gene expression, copy number analysis, gene expression, gene mutational signatures, HLA typing, neoantigen prediction and generates CIRCOS plots for individual patients from Whole Genome, Exome, Panel and RNAseq data. All variants are linked in real time to dbSNP, 1000 gnomes, TCGA, ExAC, Cosmic, Clinvar, and HGMD. Bioinformatic functional predictions for variants are performed in real time using PolyPhen-2, SIFT, FATHMM, MA, VEST, and CADD. The database predicts and presents germline and somatic clinical actionability scores that enable clinical management. Thus, the Oncogenomics web resource provides bench-scientists, bioinformatics scientists, pathologists, geneticists, and clinicians a simple, intuitive interface to explore, visualize, filter, compute and download a large collection of pre-normalized data-sets. Probe, annotation and ontology-based navigation allows users to easily distill data and meta-data on their genes, drugs and pathways of interest as a starting point for data-mining, functional genomic studies and generating testable hypotheses. This is also a patient-centric tool for geneticists and molecular pathologists that is currently being utilized at genetic and molecular tumor boards for clinical sign out. We propose that is could be one of the platforms for data visualization and information dissemination for Pediatric Cancers for the Childhood Cancer Data Initiative.

Development of Tools and Methods to Extract Knowledge from Data

Poster Number	Title	Author and Affiliations	Abstract
35	Synergy Between the Childhood Cancer Data Initiative (CCDI) and the Integrated Canine Data Commons (ICDC) to Uniquely Propel Advances in Childhood Cancer	AK LeBlanc1, DW Knapp2, WA Kibbe3, JM Trent4, M Breen5, MR Chambers6, A Dearry1, DL Duval7, AP Heath8, W Hendricks4, PM Jacobs1, AR Kerlavage1, EM Kim1, CA London9, C Mazcko1, EA Ostrander10, CL Sommers1, GJ Tawa11, RGW Verhaak12, S Zhao13, ML Beyers14, P Musk14, J Otridge14, RE Parchment14, TT Hecht1 1National Cancer Institute, National Institutes of Health including: Comparative Oncology Program, Division of Cancer Treatment and Diagnosis, Center for Biomedical Informatics and Information Technology 2Department of Veterinary Clinical Sciences, Purdue University College of Veterinary Medicine 3Duke Cancer Institute, Duke University 4Translational Genomics Research Institute, Phoenix, AZ 5Dept. of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University 6 Department of Neurosurgery, University of Alabama 7Department of Clinical Sciences, Colorado State University 8Children's Hospital of Philadelphia 9Cummings School of Veterinary Medicine, Tufts University 10National Genome Research Institute, National Institutes of Health 11National Center for Advancing Translational Sciences, National Institutes of Health	Novel targeted therapies and immunotherapies are emerging that hold considerable promise to improve the outlook for children with cancer. The number of new drugs and drug combinations, however, far exceeds the capacity of pediatric clinical trial networks. Highly relevant pre-clinical animal models are essential to refine and inform the most promising approaches to advance into human clinical trials. Our group is focusing our collective efforts on studies of spontaneously-arising cancer in pet dogs in which the cancer closely mimics the human condition in pathology, molecular features, biological behavior, host immunocompetence, and treatment response. Canine cancers include many that affect children (osteosarcoma, brain tumors, lymphomas, leukemias), and the canine model can greatly extend and complement existing experimental models. In recognizing the importance of a comparative oncology research approach, the NCI's Division of Cancer Treatment and Diagnosis (DCTD) has launched three initiatives to expand the value, use, and impact of spontaneously-arising cancer in pet dogs. These initiatives have included: (1) Cancer Center Support Grant (P30) supplements to sequence canine cancers including some that affect children (glioblastoma, lymphoma, osteosarcoma), (2) UO1 awards to study immunotherapies in canine cancer clinical trials, and (3) an Integrated Canine Data Commons (ICDC). Data from the P30 supplements and UO1 initiatives are being used to build a prototype for the ICDC. The ICDC leverages the architecture of the NCI Cancer Research Data Commons and will become a publicly accessible data repository of canine genomic, pathological, clinical, imaging, and other data. With strategic development, the ICDC will offer an extraordinary opportunity for cross-species analyses, taking advantage of highly relevant spontaneously-arising animal models to advance the management and outcome of childhood cancer. Currently, steps are being taken to assure synergy between the Childhood Cancer Data Initiative (CCDI) and ICDC efforts including, but not limited to: (1) providing support to sequence canine cancers and provide a mechanism for sharing data and analyzing relevance to human childhood cancer to determine the molecular drivers, and (2) establishing a strategic alignment of the CCDI and ICDC format and structure to allow easy integration of data for hypotheses-driven cross-species analyses.

Poster Number	Title	Author and Affiliations	Abstract
		12The Jackson Laboratory 13Department of Biochemistry and Molecular Biology, University of Georgia 14Frederick National Laboratory for Cancer Research, National Institutes of Health	
32	Enabling Data Discovery and Analysis in the Placental Atlas Tool (PAT) Through Semantically-Driven Data Federation	Alexandra Shlionskaya, M.S., Christopher H. Ferguson, Ph.D., Bianca Patel, M.S., Andrijana Dabic, B.S., Brett Pickett, Ph.D., Frederick Dong, Ph.D., Michael Keller, Ph.D. Booz Allen Hamilton, Inc. McLean, Virginia, USA	Description: In October 2018, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) launched the Placental Atlas Tool (PAT) system to provide a comprehensive placental knowledgebase, analytic tools, and relevant publications linked to curated molecular and image data. This supports visualization of placental function and development down to the molecular level, melding of diverse datasets into an integrated resource, and scientific data analysis across studies. Applicability: PAT is built on the principles of data accessibility and reuse. PAT facilitates harmonization of placental research terminology that is used as metadata for data discovery, annotation, association, and analysis. Using the terminologies and native APIs, data is extracted from multiple disparate public and proprietary resources including NCBI Gene Expression Omnibus (GEO), EMBL-EBI ArrayExpress, NLM Open-i. Use of standardized terminologies enhances data findability, links disparate data entities, and facilitates knowledge integration. Analytic pipelines enable users to extract knowledge from data through bioinformatic analysis of molecular datasets accessible to all researchers regardless of bioinformatic experience. Potential for broader use: The PAT system employs a microservice architecture and terminology-based, data-driven design principles to facilitate a flexible, scalable, and user-friendly experience that is adaptable to any research area of interest. The use of standardized terminologies allows for PAT to be easily extended to other research areas, such as pediatric neuro-oncology, by extending or replacing of the system terminologies. Evidence of success & lessons learned: Since release, the PAT system has garnered enthusiasm and positive feedback from users across the placental research community. Currently, PAT contains over 500 datasets and over 800 images and has been explored by more than 3000 users. Additionally, PAT is being incorporated into college curriculum as a learning tool to introduce undergraduates to bioinformatics

Poster Number	Title	Author and Affiliations	Abstract
			and molecular data analysis. Analysis of system use has demonstrated that molecular data findability and annotation accuracy are the most valued aspects of PAT, while publication-derived images and duplication of PubMed functionality have limited value.
57	Exploring Patterns of Care among American Indian Children with Cancer	Amanda E. Janitz1, Janis E. Campbell1, Julie A. Stoner1, René McNall-Knapp1, Hanumantha R. Pokala1, Jessica Blanchard2 1University of Oklahoma Health Sciences Center, Oklahoma City, OK 2University of Oklahoma – Norman Campus, Norman, OK	Introduction: Childhood cancer is the leading cause of disease-related death among children aged 5-19 years in the US. While there have been great successes in the treatment of cancer, little information is available on disparities in access to treatment and survival among underrepresented populations, especially American Indian/Alaska Native (AI/AN). No recent studies have conducted in-depth analysis of disparities among AI/AN children with cancer. AI/AN children who use Indian Health Service (IHS) or tribal health clinics may have increased challenges due to the need to obtain referrals through the IHS system, which is unique to tribal healthcare and subject to funding limitations. Our goal is to develop a population-based dataset that will aid in identifying and eliminating cancer health disparities, particularly among AI/AN children. Methods: Our aim is to evaluate the relation between race and event-free survival and overall survival among AI/AN and white children diagnosed with acute lymphoblastic leukemia (ALL) prior to age 20 from 1997 to 2018. We have partnered with the two children's hospitals in Oklahoma that treat the majority of children with cancer in the state. We are currently abstracting data from electronic medical records to obtain data on cancer diagnosis, treatment, and outcomes and will link with state cancer registry data to provide additional information on mortality and/or secondary cancers. Linking with state cancer registries, which include IHS-validated race information, will also allow us to better identify AI/ANs as these patients are frequently misclassified as another race in medical records. We will adjust for demographic and clinical factors as potential confounders, while investigating delay in diagnosis and geographic distance from treatment facilities as mediators. Future Directions: This project will generate important preliminary data for future studies, which will evaluate strategies to improve care coordination and incorporate cultural factors important to AI/AN families into pediatric oncology care by tribal health and oncology providers. We aim to partner with pediatric institutions across the US that treat cancer to expand our study and evaluate geographical differences in

Poster Number	Title	Author and Affiliations	Abstract
			patterns of care. Once we understand disparity patterns and mediating factors, we will develop and test interventions to address the identified disparities.
58	Enriching a population-based sample of childhood cancer survivors with multi-level data to examine receipt of recommended follow-up care: The Project Forward study	Ann S. Hamilton1, Kimberly A. Miller1,2, Jessica Tobin1, Anamara Ritt-Olson1, Cynthia N. Ramirez1, Denise Modjeski1, Katherine Y. Wojcik1, David R. Freyer3,4,5, Maryann Pentz1, Lourdes Baezconde-Garbanati1, Joel E. Milam1 1Department of Preventive Medicine, Keck School of Medicine of the University of Southern California, Los Angeles, CA; 2Department of Dermatology, Keck School of Medicine of the University of Southern California, Los Angeles, CA; 3Department of Pediatrics, Keck School of Medicine of the University of Southern California, Los Angeles, CA; 4Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA; and 5USC Norris Comprehensive Cancer Center, Los Angeles, CA	Background: Childhood cancer survivors (CCS) are at high risk for late effects from their treatment and lifelong monitoring is recommended. However, CCS often disengage from healthcare as they transition to adulthood. The Project Forward study examined cancer-related follow-up care and psychosocial outcomes among diverse young adult CCS using population-based sampling methods, medical chart abstraction, and qualitative interviews. Methods: CCS were identified through the Surveillance, Epidemiology, and End Results (SEER) Cancer Registry covering Los Angeles County. Eligible CCS were diagnosed with any cancer (stage II or greater/all stages for brain) between ages 0-19 during 1996-2010 and were at least 5 years from diagnosis when mailed a survey (with Spanish translation as needed). Survey data were linked to registry data and assessed follow-up care, psychosocial factors, and health behaviors. The Dillman Survey Method was used to enhance response (including tracing and multiple contacts by phone, mail, email, and text) and paper and online survey options were offered. HIPAA consent was requested to abstract medical records, and qualitative interviews were conducted to understand patient motivation regarding follow-up care. Results: Of 2,608 eligible cases, 1,169 responded (44.8% recruitment rate and 64.3% participation rate among those who were located (1,817)). Clinical factors were not associated with response; however, demographic factors were, including race/ethnicity (higher response by non-Hispanic whites), socioeconomic status (SES) (higher response in areas with higher SES), and gender (higher response by females.) 39% completed surveys online; 63% signed HIPAA. On average, 11.4 follow-up phone calls were made, and 7.6 mailings sent. Initial treatment and site specific follow-up care were obtained from medical charts. Forty respondents were selected for qualitative interviews stratified on those in follow-up vs. no follow-up care, Hispanic vs. non-Hispanic, and male vs. female. Conclusions: We used multiple methods to enrich data on patterns of follow-up care among diverse young adult CCS. Use of an intensive recruitment strategy and population-based sample resulted in a representative sample of CCS while additional

Poster Number	Title	Author and Affiliations	Abstract
			study components demonstrate how the linkage of these methods enhances the value of this cohort. Follow-up of this cohort will be important to monitor long term survivorship in this vulnerable population.
70	Predicting drug response using integrated genome-wide molecular profiles of pediatric tumors	Aparna Gorthi, Yu-Chiao Chiu and Yidong Chen Greehey Children's Cancer Research Institute, University of Texas Health at San Antonio	Childhood cancers represent a rare and distinct class of genetically driven tumors forming less than 1% of new cancer diagnoses. Given the lower incidence, there is little impetus to drive drug development for this genetically unique cohort. Further, the focus of pediatric oncology unlike adult cancers is "effective" cure that minimizes the long-term adverse effects of treatment. Therefore, there is an urgent need to develop models that provide early insight into what drugs and which potential combinations would be beneficial to test in a preclinical setting. Pharmacogenomics exploration relies on the underlying assumption that drug response is correlated with the inherent molecular makeup of the cells being treated. This is especially relevant for the pediatric population that is known to have distinctly different tumor behavior and chemotherapeutic response from adult cancers. We have developed a deep neural network-based model to predict drug response based on mutation and expression profiles of a given cancer cell line or tumor. The framework includes pre-trained autoencoders for mutation data and expression data each as well as a drug response predictor network that integrates the abstraction of the high-dimensional datasets of the autoencoders along with the drug response data (represented as the IC50 value of the drug). The model was trained on the expression and mutation profiles along with IC50 values to 265 drugs obtained from the Broad Cancer Cell Line Encyclopedia and The Cancer Genome Atlas databases. The model was validated on 9059 tumor datasets from The Cancer Genome Atlas and achieved superior performance in IC50 prediction compared to other methods. Preliminary analysis on pediatric cancer datasets was done with expression and mutation profiles obtained from the Pediatric Pre-clinical Testing Program and the Pediatric Pre-clinical Testing Consortium and the specificity of predictor as well as other metrics are reported. The significance of the proposed study lies in the unbiased exploration of complex interaction maps hidden in large-scale data using advanced machine-learning algorithms.
46	Navigating Access to Care: Understanding the Primary Care and	Becky N. Lowry, Kyla Alsman, Hope Krebill, Gary Doolittle, Jennifer R. Klemp	Background: With continued improvements in survival rates we are seeing a growing numbers of childhood cancer survivors (CCS). Studies have compared the comfort-levels and knowledge between primary care and oncology physicians; however, little focus has been placed on analyzing the primary and subspecialty care needs of these

Poster Number	Title	Author and Affiliations	Abstract
	Subspecialty Care Needs of Adolescent and Young Adult Survivors of Childhood Cancer		patients after active treatment. Studies describe oncology provider shortages and the decline in oncology appointments 5 years after treatment highlighting the importance of primary care in delivery of survivorship care. Literature supports that primary care providers (PCP) prefer to collaborate with a specialized survivorship clinic, but unfortunately, access to dedicated survivorship clinics is limited. Understanding these care gaps by analyzing current populations may help institutions prioritize process development for their regional patient populations. Aim: Understand primary care gaps and subspecialty referral trends to assess impact to care needs for rural and urban young adult survivors of childhood cancer. Method: Retrospective chart review was completed on 182 patients from the University of Kansas Cancer Center Survivorship Transition clinic (STC) from 2014-6/2019. The electronic medical record was utilized to identify the geographic area that patients from the STC reside and to track the primary and subspecialty care needs. Results / Conclusion: Despite complex medical history and the value of maintaining medical care, over half of CCS arrived to STC without an established PCP. As the only dedicated survivorship clinic for adult survivors of childhood cancer in the region, STC provides specialized comprehensive survivorship care for patients from 37 counties, 5 states and an international location. Subspecialty referral trends reflected common late effect manifestations of chemotherapy, surgery, and therapeutic radiation. The most frequent referrals included: dermatologic screenings, fertility / reproduction, cardiac health, secondary malignancy and post-transplant monitoring. The need for mental health referrals echoed the well described impact cancer history has on overall health. Nurse navigation is an essential component in supporting the access and coordination of care for these complex patients.
84	Implementation of the PROMIS(10) Global Health Questionnaire in the Management of Adolescents and Young Adults with	Bhavana Bhatnagar1, Kristine Diener2, Samantha Hulett2,3, Dori Klemanski3, Tarah Amato1, Cassidy Day1 and Maryam Lustberg3,4 1Division of Hematology, The Ohio State University James Cancer Hospital and Solove Research Institute, Columbus OH	Introduction: The Adolescent and Young Adult (AYA) patient population represents a unique and ever-growing demographic with distinct psychosocial needs that are either not discussed or are difficult to capture at the time of clinical assessments. As such, we integrated the use of the PROMIS10 Global Health Questionnaire in an AYA-dedicated hematology clinic in order to better gauge the physical and emotional concerns of these patients. Methods: The James Cancer Hospital Hematology and Transplant Clinic opened a pilot AYA clinic for patients between 18-39 years with

Poster Number	Title	Author and Affiliations	Abstract
	Hematologic Malignancies	2Department of Social Work, The Ohio State University James Cancer Hospital and Solove Research Institute, Columbus OH 3Center for Survivorship, The Ohio State University James Cancer Hospital and Solove Research Institute, Columbus OH 4Division of Medical Oncology, Department of Internal Medicine, The Ohio State University James Cancer Hospital and Solove Research Institute, Columbus OH	hematologic malignancies. Patients were provided with the PROMIS10 Global Health Questionnaire at the time of their initial visit and every three months if they were on active therapy and every six months if they were not. Briefly, this questionnaire contains a total of 10 questions in which patients self-report their physical and emotional symptoms as either 'excellent,' 'very good,' 'good,' 'fair,' and 'poor'. These responses are used to direct their psychosocial assessment at the time of their clinic visits. Results: This is a descriptive analysis of responses obtained from the PROMIS10 questionnaire. There were a total of 32 questionnaires collected from 22 patients between August 2018-April 2019. Eight patients had serial questionnaires completed at different time points in treatment. The study population contained nine males and 13 females with a median age of 28.5 years (range, 19-40 years). Patient diagnoses included acute lymphoblastic leukemia (N=9), acute myeloid leukemia (N=8) and chronic myeloid leukemia (N=5). Thirteen patients were on active treatment and nine had completed all therapy. Of the 22 patients who completed this questionnaire, only eight patients (36%) indicated 'excellent,' (N=4) or 'very good' (N=4) quality of life and health on at least five questions. Twelve patients (55%) indicated that at least one aspect of their life was 'fair,' most commonly in response to questions regarding physical health, emotional distress, relationships and level of fatigue. Conclusions: PROMIS10 is a simple assessment tool that can be easily integrated into the management of AYA patients and provides valuable information that can be more readily incorporated to assess the unique physical and psychosocial needs of this population.
71	OCTAD: an open workplace for virtually screening therapeutics targeting precise cancer pediatric patient groups using gene expression features	Billy Zeng, Benjamin S. Glicksberg, Patrick Newbury, Jing Xing, Ke Liu, Anita Wen, Caven Chow, Bin Chen Department of Pediatrics and Human Development, Michigan State University	Rapidly decreasing costs of RNA sequencing have enabled large-scale profiling of cancer tumor samples with precisely defined clinical and molecular features (e.g., low-grade IDH1 mutant glioma). Identifying drugs targeting a specific subset of cancer patients, particularly those who do not respond to conventional treatments, is critically important for translational research. Many studies have demonstrated the utility of a systems-based approach that connects cancers to efficacious drugs through gene expression signatures to prioritize drugs from a large drug library. From our previous work on liver cancer, Ewing's sarcoma, and basal cell carcinoma, we have shown that the success of this approach is made possible by critical procedures, such as quality control of tumor samples, selection of appropriate reference tissues, evaluation of disease signatures, and weighting cancer cell lines. There is a plethora of relevant datasets and analysis modules that are publicly available, yet are isolated in distinct silos, making it tedious to implement this approach in translational

Poster Number	Title	Author and Affiliations	Abstract
			research. As such, we present the current protocol OCTAD (http://octad.org/), which we envision as a best practice to prioritize drugs for further experimental evaluation. In this project, we retrieved patient tumor samples based on specified clinical and/or molecular features from the Genomic Data Commons Data Portal using an API. We then created a gene expression signature for these samples through employing normalized RNA-seq counts processed in the UCSC Xena project, where all RNA-seq samples from TCGA, TARGET, and GTEx were aligned and normalized using the same pipeline. We evaluated each disease signature via a cross-validation approach. We then created drug signatures using a similar procedure from large-scale, open-access platforms, namely the LINCS L1000 library, which consists of over 20,000 compounds. Our pipeline can then compute and assess the reversal potency between the disease signature and each drug signature. The drugs that present high reversal potency are prioritized as drug hits. We have shown that our prediction corroborates with the experimental data significantly in several pediatric cancers. We will be able to include new pediatric RNA-Seq samples into our portal.
51	Psychosocial Issues for Adolescent and Young Adult (AYA) Cancer Survivors: Grounding Future Studies in a Social Ecological Context	Brad Zebrack; Nina Jackson Levin; Steve W. Cole	Knowledge of the broader social ecological context in which AYAs live and mature will better inform science and clinical research data needs, particularly with regard to the identification and specification of predictor and outcome variables of relevance to this age-defined population. In contrast to theories of human development that have driven the first decade of AYA research, a social ecological framework for research considers: (1) Precarious labor conditions affecting AYA financial and work lives. High costs of living, lower (inflation-adjusted) wages, and the changing nature of career and employment opportunities compound physical and mental health challenges and threaten AYAs' financial security. (2) Changing timetables and priorities for developmental tasks. Young people are choosing to live longer with parents, delay marriage and children, and work multiple jobs related to their interests as opposed to striving toward a single career over the course of life. (3) Sexual and gender plurality. AYAs today live in a culture where sexual practices and gender identities are expressive and fluid, and where preferences for relationships and children cannot be assumed based upon conventional definitions of what constitutes "family." Furthermore, sex, sexuality, gender identity, and fertility are separate and distinct topics that are often confounded in AYA-related research.

Poster Number	Title	Author and Affiliations	Abstract
			(4) Expanding cultural diversity. Data representing demographic characteristics do not necessarily reflect the variation in AYAs' culturally-imbued values, preferences, or beliefs and thus limit our understanding of the AYA cancer experience across sub-populations defined by race/ethnicity, religion, or sexual orientation. (5) Social genomics. Emerging research is examining pathways by which psychological, social, cultural, economic, and other life circumstances influence the expression of human genes throughout the body, and in particular in cancer cells. (6) Technology and social media. Information circulation, interpersonal communication, and self-expression occur in a digital world for AYAs. Yet, understanding AYA media engagement and its implications for treatment and symptom management requires in-depth and unbiased investigations as to its utility and clinical benefit. As AYA oncology enters its adolescence, its knowledge base will be improved by pivoting toward a more expansive social ecological context for informing research questions and meaningful datasets for clinical care.
62	Data-driven care for children with cancer predisposition syndromes	Christopher C. Porter,1 Garrett M. Brodeur,2 David Malkin,3 Kim E. Nichols,4 Joshua D. Schiffman,5 Sharon E. Plon,6 Anita Villani3 1Children's Healthcare of Atlanta, Emory University, Atlanta, GA 2Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA 3The Hospital for Sick Children, University of Toronto, Toronto, Ontario 4St. Jude Children's Research Hospital, Memphis, TN 5Primary Children's Medical Center, University of Utah, Salt Lake City, UT 6Texas Children's Hospital, Baylor College of Medicine, Houston, TX	It is now known that genetic predisposition contributes to the development of cancer in children at much higher rates than previously appreciated. This creates a unique opportunity to develop improved methods to identify children at increased risk for cancer, prevent or detect tumors early, and improve cure rates with decreased morbidity. However, a number of challenges have impeded successful research and systematic approaches to the care of children with cancer predisposition. Individual cancer predisposition syndromes remain relatively rare and many are clinically heterogeneous. The data accumulated to date are weakened by ascertainment bias inherent in epidemiologic studies of those with the most penetrant phenotypes. This property limits our understanding of the true cancer risks, genotype/phenotype correlations and impact of genetic modifiers, leading to under-informed, management plans and little stratification of care. Further, there has been little consistency in the identification of at-risk children and lack of structured data and consistent ontologies to describe the management and surveillance practices across treatment centers for children with an underlying cancer predisposition. Concerted group efforts have emerged to begin to address these challenges, particularly in the area of pre-symptomatic tumor detection. We led development of consensus surveillance guidelines, which now demand robust, prospective evaluation. We have demonstrated in Li-Fraumeni syndrome that coordinated, multi-institutional efforts can produce impactful outcome data, and there is now a need to further enhance our ability to aggregate surveillance data across institutions and across syndromes.

Poster Number	Title	Author and Affiliations	Abstract
			Furthermore, the development and assessment of novel surveillance strategies incorporating circulating tumor DNA is an important goal. Systematic collection and integration of data from germline and somatic sequencing, methylation arrays, radiologic imaging, biochemical analyses and circulating tumor DNA present the opportunity to refine tumor surveillance and management strategies, and necessitates international collaboration with integrated databases and novel analysis approaches. We have proposed a Consortium for Childhood Cancer Predisposition to work closely with the Children's Oncology Group and international collaborators to begin to address this opportunity. With the development of an inclusive, multicenter registry, federated biorepository, and integrated databases we will address critical gaps in this field that may be applied more broadly, as well.
48	The Life Cancer Survivorship Research database: A Resource for Studying Transitional Care Status and Other Health Outcomes Among Young Adult Survivors of Childhood Cancer	David R. Freyer1,2,3,4, Amie E. Hwang2,5, Yi Juin Tan1, Maureen Cairns2, Annie Chen1,3, Joel E. Milam2,5, and Kimberly A. Miller2,5,6 1Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA 2USC Norris Comprehensive Cancer Center, Los Angeles CA Departments of 3Pediatrics, 4Medicine, 5Preventive Medicine, and Dermatology6, Keck School of Medicine, University of Southern California, Los Angeles, CA	Background Young adult survivors of childhood cancer (YASCC) are at risk for developing health problems caused by cancer treatment, necessitating transition of survivorship-focused care from pediatric to adult-focused facilities to support life-long surveillance and risk reduction strategies. Unfortunately, a large proportion of YASCC fail to complete transition, but little is known about factors influencing this disparity. The objective of this current study is to determine the efficacy of our transitional care model and the biomedical (host, disease and treatment) factors associated with transition failure. Description of Resource The LIFE Cancer Survivorship Research Database was established in 2009 to support survivorship research at Children's Hospital Los Angeles (CHLA). As of June 24, 2019, it comprised 1,655 survivors characterized for host demographics, disease characteristics, detailed cancer treatment exposures, current health problems with severity grade and attribution, and late effects risk classification (higher vs. lower, by program criteria). Higher-risk YASCC ≥ 21 years old are eligible to initiate transition from CHLA to our adult-focused survivorship clinic at USC Norris Comprehensive Cancer Center. Description of Cohort For this analysis, 742 higher-risk YASCC were transition-eligible. Of these, 408 (55.0%) were male; 431 (58.1%) were Latino. 389 (52.4%) had leukemia/lymphoma, 150 (20.2%) CNS tumors, 70 (9.4%) bone/soft tissue sarcoma, 40 (5.4%) neuroblastoma, and 81 (10.9%) other solid tumors. Age at diagnosis was < 16 years old for 656; attained age is ≥ 26 years for 430. 683 (92%) survivors were exposed to chemotherapy and 325 (43.8%) to irradiation. At initial survivorship visit, 400 had 1-5 established health problems, 138 had 6-10, 34 had 11-18, and 170 had

Poster Number	Title	Author and Affiliations	Abstract
			none. Of 366 YASCC whose transition status has been determined, 216 (59.0%) initiated transition whereas 148 (40.4%) were lost to follow-up before that point. Preliminary Conclusions and Planned Analyses Pre-transition attrition of survivors is substantial and warrants upstream interventions to improve retention. For those YASCC who do initiate transition, this database study when completed will reveal what proportion actually complete transition in the context of our transitional care model and biomedical risk factors associated with transition failure, including demographics, diagnosis, treatment exposures, and health status at time of transition.
75	Credentialing novel pediatric glioblastoma and ependymoma models with human single-cell RNA sequencing	David Rincon Fernandez Pacheco Cedars-Sinai, Los Angeles, USA Gi Bum Kim Cedars-Sinai, Los Angeles, USA David Saxon Cedars-Sinai, Los Angeles, USA Amy Yang Cedars-Sinai, Los Angeles, USA Sara Sabet Cedars-Sinai, Los Angeles, USA Hannah Park Cedars-Sinai, Los Angeles, USA Shea Chandra Cedars-Sinai, Los Angeles, USA Kristyna Sendivakova Cedars-Sinai, Los Angeles, USA Paul Linesch Cedars-Sinai, Los Angeles, USA Chintda Santiskulvong Cedars-Sinai, Los Angeles, USA Yizhou Wang Cedars-Sinai, Los Angeles, USA Jie Tang Cedars-Sinai, Los Angeles, USA Mariella Filbin Dana-Farber Boston Children's, Boston, USA Mario Suva Broad Institute of MIT and Harvard, Cambridge, USA Moise Danielpour Cedars-Sinai, Los Angeles, USA Joshua Breunig (Presentin	We have developed a simple and generalizable in vivo method for pediatric brain tumor modeling, mosaic analysis by dual recombinase-mediated cassette exchange (MADR). MADR allows for stable labeling of mutant cells expressing transgenic elements from a precisely-defined chromosomal locus. We have demonstrated the power and versatility of MADR by creating novel glioma models with mixed, reporter-defined zygosity, or with "personalized" H3. 3containing driver mutation signatures from pediatric glioma--each manipulation altering the spatiotemporal profile of resulting tumors. Further we have generated ependymoma models by employing patient-derived fusion driver mutations. Notably, each model displays divergent spatiotemporal tumor expansion profiles and cellular phenotypes. Now, we use single-cell RNA-seq to compare these models to their cells of mutation and to human datasets to elucidate the fundamental transcriptional programs and markers within and across these tumor types. This investigation will assess and credential these models against their clinical counterparts by scrutinizing the resulting datasets and validating their clinical relevance as preclinical models for therapeutic discovery and testing. This also serves as framework for assessing potential needs for clinical datasets from the Childhood Cancer Data Initiative.
24	Rapid Case Ascertainment in Childhood Cancer	David A. Siegel, Taylor Ellington, Toye Williams, Reda Wilson	Program: National cancer registry incidence data are typically not available for public analysis until 24-36 months post-diagnosis. As directed by the Caroline Pryce Walker Act (2008), the Centers for Disease Control and Prevention established the Early Case

Poster Number	Title	Author and Affiliations	Abstract
	Surveillance Data Collection	Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA	Capture of Pediatric and Young Adult Cancers (ECC) program. Nine US states during 2011-2019 were funded to capture new cases of cancer diagnosed in individuals aged 0-19 years within 30 days of diagnosis, as opposed to six months. ECC variables included patient demographics, tumor characteristics, and follow-back contact information for the patient and provider. Applicability: ECC data could be used by cancer registries or linked to clinical trial databases to ascertain diagnosis patterns of patients that could be targeted for clinical trials, and allow for changes in clinical trial recruitment efforts in a timely manner. Rapid case ascertainment could potentially help state registries identify patient needs for treatment, supportive care, or long-term follow-up. Potential for broader use: The Childhood Cancer STAR Act (2018) directs CDC to "enhance and expand infrastructure to track the epidemiology of cancer in children, adolescents, and young adults." Methods of the ECC project will inform implementation of the Childhood Cancer STAR Act. Specific goals of implementation will be to improve timeliness and completeness of rapid case ascertainment data, expand the age range of rapid case ascertainment to include young adults, and scale the data collection beyond the current registries involved. Evidence of success and critical lessons learned: Seven states were funded under ECC in 2018. These states recorded 84%, 77%, 72%, 63%, 60%, 56%, and 44% of their cases by 30 days after diagnosis, and the majority of their cases were submitted to state registries 60 days after diagnosis. Improvement in timeliness was seen during the ECC program years. Experience from the ECC program will inform implementation of the Childhood Cancer STAR Act. Electronic reporting was found to be the most timely method of data transmission. However, sole reliance on electronic pathology reporting can decrease completeness because it misses clinically or radiologically diagnosed cases, and might not capture race/ethnicity data. Specific plans could be initiated to address these limitations.
82	Informatics Methods for Population-based Early Case Capture of Pediatric and	Eric B. Durbin1,2,3, Ellen Lycan1, Toye Williams4, Vicki Benard4 1Kentucky Cancer Registry, Lexington, Kentucky 2Markey Cancer Center, University of Kentucky, Lexington, Kentucky	Background/Objectives Population-based cancer surveillance data are essential to monitoring incidence trends and the planning, implementation and evaluation of cancer prevention and control initiatives. Registry data also have the potential to play a greater role in clinical trial recruitment and cancer survivorship initiatives, particularly for pediatric and young adult cancer (PYAC) patients (ages 0-19). Rapid identification of PYAC

Poster Number	Title	Author and Affiliations	Abstract
	Young Adult Cancers	3College of Medicine, University of Kentucky, Lexington, Kentucky 4Centers for Disease Control and Prevention, Division of Cancer Prevention and Control, Atlanta, Georgia	cases is needed, yet current reporting practices require approximately 19 months to achieve complete, population-based data. We describe the development of informatics methods for early case capture (ECC) of PYAC cases within 30 days from the date of diagnosis in the state of Kentucky from 2012-2018. Design/Methods We utilized a prospective population-based study design with support from the U.S. Centers for Disease Control and Prevention's Pediatric and Young Adult Early Case Capture Program. Our approach relied heavily upon informatics methods and electronic reporting sources, such as electronic pathology reports, electronic health records and the novel use of a state health information exchange network. A variety of data sources, innovative software applications and registry workflows were evaluated to achieve complete reporting within the minimum amount of time. Results Our results show that the majority (>90%) of pediatric cancer cases can be successfully identified by a population-based central cancer registry within 30 days of diagnosis. Natural language processing of electronic pathology reports was demonstrated to be the most productive method to identify the greatest proportion of cases. Limitations arose from Kentucky cases diagnosed and treated out-of-state. In addition, these efforts led to the development of new data dissemination tools that have informed public policy and revealed a high burden of brain and central nervous system tumors among children in Kentucky. As a result, an additional study has been funded to better understand environmental and genetic factors that may be associated with the high rates. Conclusions Our study demonstrates that early case capture of PYAC cases is both feasible and practical for population-based cancer registries. However, significant investments in electronic reporting infrastructures are required. This approach clearly demonstrates the potential to collect and share early incidence data with clinicians, public health practitioners and researchers. Additional uses of ECC data, such as identifying patients for clinical trial recruitment, warrants further investigation.
81	Leveraging Data Sharing Infrastructures to Study Factors Associated with	Eric B. Durbin1,2,4, Tom Badgett2,3,4, Jamie Bloyd7, Therese Bocklage2, W. Jay Christian2,5, Robert F. Debski6, Chunyan He2,4, Bin Huang1,2,4, Jong Cheol Jeong2,4, Ellen Lycan1, Rachel	Background/Objectives Pediatric brain and central nervous system tumors (PBCNST) are the most common solid tumors in United States (U.S.) children and are the leading cause of disease-related death. PBCNST incidence rates in Kentucky are significantly higher than in the U.S. and even higher among Kentucky's Appalachian children. A population-based

Poster Number	Title	Author and Affiliations	Abstract
	High Incidence of Pediatric Brain and Central Nervous System Tumors	Maynard1, Janna M. Neltner4, Hilary Nickols6, Adam Resnick8, Thomas C. Tucker1,5 1Kentucky Cancer Registry, Lexington, Kentucky 2Markey Cancer Center, University of Kentucky, Lexington, Kentucky 3Kentucky Children's Hospital, University of Kentucky, Lexington, Kentucky 4College of Medicine, University of Kentucky, Lexington, Kentucky 5College of Public Health, University of Kentucky Lexington, Kentucky 6Norton Healthcare, Louisville, Kentucky 7Kentucky Pediatric Cancer Research Trust Fund, Frankfort, Kentucky 8Children's Hospital of Philadelphia, Philadelphia, Pennsylvania	study partially funded by the Kentucky Pediatric Cancer Research Trust Fund is underway to identify environmental and genetic risk factors that may be associated with the high incidence. Our study leverages the data sharing infrastructures provided by the Kentucky Cancer Registry (KCR), its Virtual Tissue Repository and the NIH Kids First Data Resource Center (DRC). Specific aims are: 1) to identify potential environmental exposures associated with Kentucky's high rates; 2) to assess whether Kentucky-specific mutations and mutational signatures are related to PBCNST, and to determine whether known genetic risk factors for PBCNST are present among Kentucky children; and 3) implement informatics infrastructures for data sharing with national PBCNST consortia. Methods The first aim includes geospatial analyses to explore geographic patterns of risk after adjustment for demographic factors to assess their relationship to environmental factors (proximity to hazardous waste, mining, and industrial sites). The second aim is undertaking a population-based analysis of the genomic mutation profiles. Formalin-fixed paraffin-embedded (FFPE) pathology specimens from historical PBCNST cases are being obtained and sequenced (both RNA and DNA) to characterize and extensively profile the Kentucky cases. The third aim involves enhancing informatics infrastructures to facilitate sharing study data with the DRC. Integrating molecular data from the Kentucky cases will allow us to explore potential differences with non-Kentucky/non-Appalachian cases in the DRC. Results/Conclusions Evidence of significant regional differences in rates of PBCNST have been confirmed. Detailed analyses with a variety of tumor types, patient demographics and potential environmental exposures are underway. A cohort of 356 cases with likely available biospecimens have also been identified and are being obtained for sequencing. This information will provide important clues about why an increased number of Kentucky children develop these tumors and also to identify potential opportunities for interventions and additional research. Our methods would be applicable to other retrospective tissue-based studies of childhood and adult cancers.
56	Factors impacting fertility preservation in children,	Erin M. Mobley, PhD, MPH; Ginny L. Ryan, MD, MA; Amy E. Sparks, PhD; Varun Monga, MD; William W. Terry, MD, MPH	Introduction Fertility preservation (FP) prior to therapy is an underutilized opportunity for those diagnosed with cancer as a child, adolescent or young adult (AYA). This study describes the process of data curation to understand the factors impacting utilization

Poster Number	Title	Author and Affiliations	Abstract
	adolescents, and young adults with cancer: A retrospective study		of FP consultations and procedures among childhood and AYA cancer patients at the University of Iowa Health Care (UIHC). Methods Patients were identified by the oncology registry at UIHC. Disease site, histology, date of diagnosis, sex, race, ethnicity, insurance, and zip code data were gathered by the registrars. Disease site and histology were categorized using International Classification of Diseases-Oncology-3 (ICD-O-3). UIHC's electronic medical record (EMR, Epic) was queried for ICD codes for FP consultation. Data from UIHC's Reproductive Endocrinology and Infertility clinical database were merged with the primary data set to capture information about patients who underwent FP. Rural-Urban Commuting Area codes were incorporated to capture a measure of rurality. Descriptive statistics and multivariate linear probability models were used to predict the probability of FP consultation and procedure. Results From 2008-2017, 3,605 children and AYAs were treated for an invasive malignancy. Of the 637 (18%) who received a FP consultation, 162 (25%) underwent a FP procedure. Multivariate analyses showed that those with public insurance or no insurance, a diagnosis of a CNS tumor, melanoma, or miscellaneous neoplasm, and age over 30 years at diagnosis had a significantly lower probability of having a consultation. The probability of undergoing a procedure was lower for female patients, those with germ cell tumor, melanoma, or carcinoma, seen by a pediatric-based provider, and diagnosed between 15-25 years of age. Conclusion This study has important implications for data aggregation, sharing, and research, particularly for childhood and AYA cancer. The combination of sample identification via cancer registrars and linkage with EMR data could be used by others. The use of federal data to incorporate rurality is frequently overlooked. The use of this type of publicly available data would help improve the robustness of childhood and AYA cancer research through the inclusion of measures of rurality, health service availability, and economic indicators at varying levels of geography.

Poster Number	Title	Author and Affiliations	Abstract
83	Charting the synthetic lethality landscape in pediatric cancer to advance whole-exome precision-based treatments	Fiorella Schischlik1*, Joo Sang Lee1*, Nirali Shah2, Rosandra N. Kaplan2, Carol J. Thiele2, Brigitte Widemann2, Eytan Ruppin1,3 1Cancer Data Science Lab, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda MD 20892, USA 2Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA 3University of Maryland Institute of Advanced Computer Science (UMIACS), University of Maryland, College Park, Maryland MD 20742, USA *Equally contributing first authors	Much of the research in adult cancers is focused on cancer drivers, aiming at their therapeutic targeting. However, pediatric cancers (PCs) are often driven by relatively few genetic alterations that are distinct from those that occur in adult cancers. Here we apply a novel data-driven approach to identify the synthetic lethality (SL) networks of PC, covering both genomic and transcriptomic alterations in whole exome. The identified SL network provides a new platform for discovering novel vulnerabilities in PCs that extend previous approaches commonly used in adult cancer. SL denotes the interaction between two genes whose combined inactivation is lethal, while their individual inactivation is not. To identify a PC-specific SL landscape, we analyze relevant pediatric cell line and patients' tumor data in the following four steps: First, we identify putative SL gene pairs from the RNAi/CRISPR-based pediatric cell line dependency map. Second, among the candidate pairs that pass the first step, we select those whose co-inactivation is under-represented, indicating that they are selected against. Third, we further prioritize candidate SL pairs whose co-inactivation is associated with better prognosis, indicating that they may hamper tumor progression. Finally, we prioritize such pairs with similar evolutionary phylogenetic profiles. Applying this approach to TARGET data, we identify the first genome-wide SL networks in five pediatric tumors, Wilms tumor, neuroblastoma, AML, ALL, and osteosarcoma. The predicted SL interactions are first validated via experimental CRISPR screens. We show that the PC-specific SL networks are predictive of drug response in pediatric but not in adult cell lines of the same tumor type, demonstrating that the SL network offers an exciting venue for developing PC-specific predictive biomarkers. Importantly, these predictions were performed in an unsupervised manner, reducing the known risk of over-fitting commonly associated with supervised methods. Notably, our analysis identifies many SL partners of key drivers of PCs including key interactions like ATRX-MAPK, MYCN-CDC6, and DNMT1-HK2. These provide novel drug combinations and selective drug targets for the tumors driven by these genes. Taken together, these results lay a basis for a new paradigm for whole exome SL-based precision treatments in pediatric oncology, complementing existing mutation and fusion based approaches.

Poster Number	Title	Author and Affiliations	Abstract
76	Clinical application of shared genomic data for pediatric cancer	Holly Beale1,2, A. Geoff Lyle1,2, Isabel Bjork2, Sofie R. Salama2,3, Avanthi Tayi Shah4, Lauren Sanders2, Jacob Pfeil2, Du Linh Lam2, Katrina Learned2, Ann Durbin2, Ellen Towle Kephart2, Rob Currie2, Yulia Newton2, Teresa Swatloski2, Duncan McColl2, John Vivian2, Jingchun Zhu2, Alex G. Lee4, Stanley G. Leung4, Aviv Spillinger4, Heng-Yi Liu4, Winnie S. Liang5, Sara A. Byron5, Michael E. Berens6, Adam Resnick7, Norman Lacayo8, Sheri L. Spunt8, Arun Rangaswami8, Van Huynh9, Lilibeth Torno9, Ashley Plant9, Ivan Kirov9, Keri Zabokrtsky9, S. Rod Rassekh10, Rebecca J. Deyell10, Janessa Laskin10, Marco A. Marra11, Leonard S. Sender9*, Sabine Mueller12*, E. Alejandro Sweet-Cordero4*, Theodore C. Goldstein2*, David Haussler2,3*,Olena Morozova Vaske1,2* 1Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, CA 2UC Santa Cruz Genomics Institute, Santa Cruz, CA, USA 3Howard Hughes Medical Institute, University of California, Santa Cruz, CA 4 Department of Pediatrics, Division of Hematology and Oncology, University of California San Francisco. San Francisco, CA 5Integrated Cancer Genomics Division, Translational Genomics Research Institute (TGen), Phoenix, AZ, USA 6 Cancer and Cell Biology Division, Translational Genomics Research Institute (TGen), Phoenix, AZ, USA 7Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA 8Stanford University School of Medicine and Stanford Cancer Institute, Stanford, CA 9Children's Hospital of Orange County, Orange, CA 10British Columbia Children's	UCSC Treehouse Childhood Cancer Initiative incorporates comparative gene expression information into the genomic analysis of difficult-to-treat pediatric cancers. In a pilot project we compared the RNA sequencing profile of 144 relapsed/refractory pediatric tumor samples to a compendium of over 11,000 uniformly analyzed tumor profiles from pediatric and adult cancer patients. Developing the compendium highlighted various roadblocks to re-using data: 1) incorrect representation of sequence data in repositories; 2) incomplete methods reporting in publications; 3) redundancy across data access applications; 4) substantial computational time required for uniformly processing data; and 5) the village of experts required to determine whether data is fully comparable in terms of clinical features, sample preparation and sequencing methods. In a proof-of-performance test, comparison of an individual patient's tumor to the compendium identified over-expressed genes and pathway. The pathways were reviewed by data analysts for potential clinical impact, then presented to the treating oncologist in a molecular tumor board-like setting at clinical genomic trials conducted by Children's Hospital of Orange County, UC San Francisco (Pacific Pediatric Neuro-Oncology Consortium), and Stanford University. We showed RNA-Seq-derived gene expression pointed to treatment options for 68% of patients, whereas DNA mutation information alone was useful for only 46% of patients. For 36% of patients with both RNA and DNA analysis, druggable targets identified by RNA analysis were not found by tumor DNA analysis. Our results demonstrate the potential clinical utility and efficacy of comparative tumor RNA-Seq-derived gene expression information within a large shared cohort, and we are now focussing on defining the incremental benefit of this approach to personalized treatment planning. Our work demonstrates how open sharing of cancer genomic data, including data from each analyzed patient, can contribute to better care for children with cancer in the future.

Poster Number	Title	Author and Affiliations	Abstract
		Hospital and British Columbia Children's Hospital Research Institute, Vancouver, British Columbia, Canada 11Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer, Vancouver, British Columbia, Canada 12UCSF Benioff Children's Hospital, San Francisco, CA	
80	refine.bio: A resource for extracting knowledge from pediatric cancer data	Jaclyn N. Taroni, Kurt G. Wheeler, Richard W. W. Jones, Deepashree Venkatesh Prasad, Ariel Rodriguez Romero, Candace L. Savonen, Casey S. Greene Childhood Cancer Data Lab, Alex's Lemonade Stand Foundation, Philadelphia, PA USA	Millions of genome-wide gene expression assays are publicly available in repositories such as Gene Expression Omnibus and Sequence Read Archive. These petabyte-scale data are a powerful tool for pediatric cancer research, but obstacles for use remain. Samples were assayed on multiple technologies and platforms and it is often unclear how they were processed, leading individual researchers to invest significant time and computing resources in data reprocessing efforts. To address this challenge, we developed refine.bio: a multi-organism collection of uniformly processed and normalized transcriptomic data obtained from publicly available repositories. refine.bio allows researchers to select from hundreds of thousands of samples and to build datasets that are tailored to their question of interest. This permits biologists and clinicians to quickly validate results in different cohorts or from model systems. Relatively few publicly available samples directly assay pediatric cancers. Other biological systems, such as cell lines or model organisms, hold rich information for discovery in these conditions. We have previously trained unsupervised models on random collections of human RNA-seq data from diverse biological contexts; these models capture more biological processes and better describe rare disease datasets, even in the absence of the disease in the training set (Taroni et al. Cell Systems. 2019.). When the backlog of samples in government-run repositories are processed and made available in refine.bio, it will be an order of magnitude larger than prior resources and, with the inclusion of multiple assay modalities, particularly well-suited to the study of rare diseases. The breadth of refine.bio is highly complementary to pediatric cancer-specific curation efforts and will be valuable in the investigation of new treatments for pediatric cancer in cases where successful match to currently available therapies is not possible. Collections of all samples from a species processed by refine.bio, termed species compendia, are well-suited for training unsupervised models that may best describe developmental processes important for understanding the biology underlying pediatric cancers.

Poster Number	Title	Author and Affiliations	Abstract
			refine.bio runs on Amazon Web Services. We have successfully processed over half a million samples, but the remote nature of cloud architectures has posed challenges. We will share what we learned processing sequencing data at this scale
6	UCSC Xena Browser to integrate and share bulk tissue and single-cell pediatric cancer genomics data	Jingchun Zhu, Mary Goldman, Brian Craft, Sofie R. Salama, Isabel Bjork, Olena Morozova Vaske, David Haussler UC Santa Cruz Genomics Institute, Santa Cruz, CA, USA	UCSC Xena (http://xena.ucsc.edu) is a web-based cancer genomics data visualization tool for both adult and pediatric data. Its web-based browser integrates and displays data from any number of private or public federated Xena data hubs. Xena data hubs are easy to install on any computer, laptop, or even in a cloud computing environment, thus enabling biologists, both with and without computational expertise, to easily view their own analysis results through the routine launch of a private Xena Hub. Xena combines data in a secure manner, protecting the privacy of user data. By standing up a public hub, Xena enables groups to easily share their data with our growing user community (in May 2019 alone, Xena Browser had 38,363 sessions from users across 73 countries). The Treehouse Childhood Cancer Initiative uses Xena (https://xena.treehouse.gi.ucsc.edu) to share its open-access RNA-Seq compendia with the scientific community. UCSC Xena complements the functionalities of a centralized data commons and cloud computing resources by providing online data visualization and analysis of data in the centralized repositories. Altogether, Xena has over 1500 multi-omic datasets across 130 cohorts, including pediatric data from TARGET, the GDC, and Treehouse. In addition to this bulk tumor data, Xena is entering the single-cell space. Besides data from the Human Cell Atlas, we have started to collect publicly available single-cell data from pediatric tumors, such as DIPG. Our tool can help pediatric cancer researchers to apply a genomic signature from one dataset to another dataset. For example, a researcher can use cell type information learned from the single cell data to better understand bulk pediatric tumor data. Xena Browser has a large array of interactive analyses and visualization capabilities, such as generating Kaplan Meier survival plots, deriving genomic signature scores, and building visual spreadsheets. It helps researchers answer questions like, “Is over-expression of geneA associated with lower survival?” or “What is the relationship between expression, mutation, copy number, for these genes?” In addition to genomics data, researchers can search for samples by querying available text-based documents such as de-identified case reports.

Poster Number	Title	Author and Affiliations	Abstract
72	Automated learning tools to extract data from co-registered digital images in childhood osteosarcoma.	Leavey, P.J.; Mishra, R.; Teo, K.W.; Xie, Y.; Guanghua, X. Sengupta, A.; Hallac, R.; Cederberg, K. and Daescu, O. Departments of Pediatrics and Population and Data Sciences, University of Texas Southwestern Medical Center Dallas and Department of Computer Science, University of Texas at Dallas.	Predictors, at the time of diagnosis, for the 45% of children with high-grade osteosarcoma who die of their disease, are limited to basic features of tumor metastasis and resectability. I hypothesize that we can improve outcome prediction using advanced imaging, digital histology and automated learning tools and my research team has completed several steps toward this goal. We have developed a novel approach to pathology preparation of surgically resected specimens with personalized 3-dimensional (3-D) tumor molds, printed from tumor-segmented magnetic resonance images (MRI), to identify the optimal plane for tumor bisection ex-vivo and thus target histology slide preparation. We have successfully developed an automated stitching tool to create a digital histology 2-D tumor map, that closely corresponds to an MRI plane and we are matching areas of interest (AOIs) on diffusion weighted (DWI) and dynamic contrast enhanced (DCE) MRI sequences to co-planar AOIs on digitized histology slides. Osteosarcoma response to pre-operative chemotherapy as measured by tumor necrosis predicts patient outcome. We have optimized and automated histology necrosis estimation using deep learning algorithms, and we are currently developing automated learning tools to interpret features within DWI + DCE MRI sequences acquired prospectively at 5 and 10-weeks of pre-operative chemotherapy. Using novel deep learning models, to classify different cell types in the tumor micro-environment, we anticipate that AOIs will allow us to build matched multi-modal scales of histological and radiological imaging. We predict this multi-modal platform will create an accurate procedure to estimate response in osteosarcoma, not-measurable by RECIST criteria, and will provide a novel predictor of outcome. We anticipate that use of digital histology and radiology images together with genomic data-sets, each of which independently predict clinical outcomes for multiple types of childhood sarcoma, will enhance predictive biomarker cassettes through the use of automated learning tools.
85	Assessment of Reproductive Late Effects in Adolescent and Young Adult Cancer Survivorship	Leslie Appiah, MD, Olanipekun Lanny Ntukidem, Dori Klemanski DNP, Lawrence Jenkins, MD, Bhavana Bhatnagar, DO, Maryam Lustberg, MD, MPH.	Introduction: Gonadotoxicity, infertility and radiation injury are among the most common adverse effects affecting 12% of childhood cancer survivors (Hudson, 2013). The clinical implications include acute ovarian and testicular failure with sequelae, difficulty achieving or siring a pregnancy, increased risk of breast cancer, genital graft-versus-host disease, and vaginal stenosis (Nieman et al, 2006), collectively termed reproductive late effects. We aimed to determine the percentage of Adolescents and Young Adult (AYA) survivors referred to the Fertility Preservation and Reproductive Health (FPRH) clinics to assess reproductive late effects in survivorship. Methods: A retrospective chart review was performed to identify AYA cancer survivors (ages 18 to 39) who completed active treatment in a newly developed AYA program at a high

Poster Number	Title	Author and Affiliations	Abstract
			volume adult comprehensive cancer center between January 2017 and April 2019. Demographic and programmatic data were collected. Results: Between 2017 and 2019, a total of 1,917AYA survivors were seen at The Ohio State University Comprehensive Cancer Center (OSUCCC) - James Cancer Hospital. The predominant cancer diagnoses amongst this population were hematologic cancers (n=944), thyroid (n=490), brain (n=451) breast (n=414) and testicular/male genitourinary (n=169). A total of 1,434 survivors received chemotherapy. Of those, 273 (19%) underwent oophorectomy, orchiectomy, hysterectomy, and/or retroperitoneal lymph node dissection. A total of 105 (7%) patients were referred to the FPRH clinics for consultative assessment and ongoing management of reproductive late effects in survivorship. Conclusions: Despite knowledge of reproductive late effects in cancer survivorship, few patients are referred for assessment. Referral patterns may be improved by enhanced tracking of new AYA diagnoses and automated processes for follow-up consultation in survivorship. An evidencebased and expert consensus developed AYA patient database will enable clinician-researchers to evaluate reproductive late effects in survivorship and assess efficacy of management and treatment interventions. The database will also streamline enrollment in AYA fertility research protocols in collaboration with the Oncofertility Consortium.
64	Subsequent neoplasm risk associated with rare variants in DNA repair and clinical radiation sensitivity syndrome genes: A report from the Childhood Cancer Survivor Study (CCSS)	Lindsay M. Morton, Danielle M. Karyadi, Stephen Hartley, Megan Frone, Joshua N. Sampson, Rebecca Howell, Joseph P. Neglia, Michael A. Arnold, Casey L. Dagnall, Belynda Hicks, Kristine Jones, Bin Zhu, Wendy M. Leisenring, Yutaka Yasui, Amy Berrington de Gonzalez, Smita Bhatia, Leslie L. Robison, Margaret A. Tucker, Gregory T. Armstrong, Stephen J. Chanock This work was supported by the National Cancer Institute, National Institutes of Health Intramural Program and grant CA55727. Support to St. Jude Children's Research Hospital is also provided by the Cancer Center Support (CORE) grant (CA21765) and the American Lebanese-Syrian Associated Charities (ALSAC).	Background: Radiotherapy for childhood cancer is associated with strikingly elevated risk for a number of long-term adverse effects. Whether some childhood cancer survivors may be genetically susceptible to radiotherapy-related adverse effects is largely unknown. We performed whole exome sequencing (WES) in CCSS, the largest cohort of childhood cancer survivors with detailed treatment data, available DNA, and systematic long-term follow-up. We investigated whether mutations in DNA repair and radiation sensitivity genes modulate risks for radiotherapy-related subsequent neoplasms (RT-SNs), a major cause of morbidity and mortality after childhood cancer. Methods: High-quality WES (mean 40x coverage) was available for 5105 childhood cancer survivors of European descent (originally diagnosed 1970-1986, mean follow-up=32.7 years). SnpEff/ClinVar identified potentially protein-damaging rare variants in 476 DNA repair or radiation sensitivity genes. Conditional logistic regression assessed RT-SN risk associated with these variants aggregated by gene or pathway, matching controls to cases on sex, childhood cancer type and diagnosis age, radiation dose to the SN site, and survival. Exact p-values were calculated by permutation. Analyses used all survivors or subsets stratified by in- versus out-of-field.

Poster Number	Title	Author and Affiliations	Abstract
			Results: A total of 1108 (21.7%) survivors developed at least one RT-SN, most commonly breast cancer, basal cell carcinoma, meningioma and thyroid cancer. Out-of-field RT-SN risk was associated with homologous recombination (HR) gene variants (cases=41/196 [20.9%], matched controls=11.0%; odds ratio [OR]=2.20, 95% confidence interval [CI] 1.52-3.18; P=1.41x10-4), most notably for FANCM (cases=3.1%, matched controls=0.5%; OR=9.91, 95%CI 3.73-26.4; P=2.85x10-4). Associations were not observed for the HR pathway for in-field RT-SN (cases=201/711 [14.4%], matched controls=12.4%, P=0.17) but were observed for two individual genes implicated in double-strand DNA break repair, EXO1 (cases=1.8%, matched controls=0.4%; OR=6.50, 95%CI 3.49-12.1; P=7.43x10-4) and NEIL3 (cases=0%, matched controls=1.0%; P=3.23x10-4). Conclusions: In this discovery study, we identified dose-specific novel associations between RT-SN risk after childhood cancer and potentially protein-damaging rare variants in genes involved in double-strand DNA break repair, particularly HR. With replication, these results could impact long-term screening of childhood cancer survivors and risk-benefit assessments of treatment approaches. Sharing of WES data from this large cohort will enable investigation of the genetic basis of a range of other long-term outcomes after childhood cancer.
74	Newborn Dried Blood Spots as a source for identifying infants with Cancer Predisposition Syndromes (CPS) and for estimating of prevalence of CPS-associated variants in children with cancer	Lisa Diller MD1,5; Richard Parad MD, MPH2,5; Andy Bhattacharjee PhD3; and Jennifer Yeh PhD4,5. 1Dana-Farber Cancer Institute, Boston MA; 2Brigham and Women's Hospital, Boston MA: 3Baebies, Inc, Durham, NC, 4Boston Children's Hospital, Boston MA, and 5Department of Pediatrics, Harvard Medical School.	Background: Population-based newborn screening (NBS) is one of the most successful public health programs in modern medicine. Over 60 actionable conditions are currently included in NBS using biologic samples from newborn "heel stick" dried blood spots (DBS). Genetically-based detection of increased risk for early onset childhood cancer risk, such as heritable retinoblastoma, is currently not included in NBS. With state public health laboratories increasingly using DNA-based tests in NBS algorithms, primary sequencing-based NBS for the early identification of newborns at high risk for developing infant or early childhood cancers may be feasible. Preliminary work ongoing: We developed a novel tNGS-based panel, designated “PERC-Seq” (Pediatric Early Risk for Cancer-Sequencing), which includes RB1 and 10 other Cancer Predisposition Syndrome (CPS) genes (RET, TP53, SMARCB1, SUFU, PTCH1, WT1, DICER1, APC, ALK, PHOX2B). Pathogenic variants in these selected genes are associated with increased risk of specific early-onset cancers whose outcomes could potentially be improved with early detection. We developed a simulation model to assess the impact of using this panel (and each gene individually) in NBS.

Poster Number	Title	Author and Affiliations	Abstract
			The Childhood Cancer Predisposition Simulation Model suggests that identification and surveillance of at-risk infants would reduce overall childhood cancer deaths by ~8%, as well as reduce the percentage of survivors living with high risk of treatment-related mortality. We worked with the Michigan Cancer Surveillance Program and the Michigan Biobank to identify DBS from children born in Michigan who went on to develop a malignancy in which the histologic tumor type was associated with variants in PERC-Seq panel genes. We identified 1803 such potential samples. The PERC-Seq Panel will be applied to these samples to estimate prevalence of pathogenic or likely pathogenic variants for each histologic subtype, and these data will be entered into the simulation model. Summary: Modeling the impact of universal NBS for CPS associated with early onset childhood cancer on detection and prevention can be helpful in evaluating new interventions. Linkage of DBS specimens to childhood cancer cases is feasible and provides a new source of biologic materials for studying cancer etiology and risk.
53	Applying an Operations Management Framework to the RIC Study Cancer Registry Linkage Process to Increase the Return on Investment of State Cancer Registry Data	Lisa Moy1, Casey Luce2, Charisma Jenkins3, Deborah Multerer4, Prachi Chavan5, Yannica Theda-Martinez6, Yolanda Prado3, Robert Greenlee4, Marilyn Kwan1* 1 Kaiser Permanente Northern California; 2 Kaiser Permanente Washington; 3 Kaiser Permanente Northwest; 4 Marshfield Clinic Health System; 5 University of California, San Francisco; 6 Kaiser Permanente Hawai'i	The cancer registry in the United States is an invaluable scientific resource that enables scientists to quantify the burden of cancer, investigate the origins of cancers, evaluate the effectiveness of interventions, and identify disparities and needs from a local to national level. A tremendous investment of resources is made at the regional-, state-, and federal- levels to ensure standardized, clinically meaningful data are reported in a timely matter with over 95% population coverage. To fully appreciate the 95% population coverage of the cancer registry, researchers must initiate linkages with individual state cancer registries, an activity which can be resource-intensive and prohibitive. Thus, the "return" on the financial and scientific investment into the development and maintenance of state cancer registries may not be fully realized by the end users (a lack of return on investment). The lack of return on investment is greater for the investigation of cancers diagnosed during childhood, since decades may pass between the exposure of interest and the development of cancer. In the intervening years, cohort members may have moved, thus requiring linkages to multiple state cancer registries. Using the Radiation-Induced Cancers (RIC) Study as a case study, we will utilize the operations management framework to identify the barriers to access cancer registry data. We will also examine the potential role of the Virtual Pooled Registry Cancer

Poster Number	Title	Author and Affiliations	Abstract
			Linkage System (VPR-CLS). In brief, the RIC Study is a cohort of 7.3 million geographically-diverse children and their mothers with the goal to determine the risk of pediatric cancers (<21 y) associated with in utero/fetal or pediatric exposure to medical imaging. It is a collaboration of 7 U.S. sites (Kaiser Permanente Northern California, Northwest, Washington, and Hawai'i, Marshfield Clinic Health System, Harvard Pilgrim Health Care, and Geisinger) and Ontario, Canada.
77	CAYACC's Online AYA Registry	Lynda K Beaupin1, Denise A Rokitka2, Jennifer Schweitzer2, Scott C Borinstein3, Damon R Reed4, Peter H Shaw1, Nicholas D Yeager5 1Johns Hopkins All Children's Hospital 2Roswell Park Comprehensive Cancer Center 3Vanderbilt University Medical Center 4Moffitt Cancer Center 5Nationwide Children's Hospital	Introduction – In 2014, 5 AYA programs (Roswell Park Comprehensive Cancer Center, Johns Hopkins All Children's Hospital [formerly Children's Hospital of Pittsburgh], Nationwide Children's Hospital, Vanderbilt University Medical Center, and Moffitt Cancer Center) established the Consortium of Adolescent and Young Adult Cancer Centers (CAYACC) to focus on offering an online research platform for AYA oncology. CAYACC's initial study was an online questionnaire capturing self-reported patient health and psychosocial data. The aim of this study was to assess the feasibility of an opt-in, secure online survey to collect data from a larger landscape of AYA patients and survivors in the US. Methods – CAYACC created a 28-question anonymous survey for cancer patients and survivors diagnosed between the ages of 18 and 39 years. Topics include diagnosis, treatment setting, clinical trial access and enrollment, insurance status, social support and fertility preservation utilization. The survey was open between April 1, 2018 and January 31, 2019. Recruitment efforts were through outreach to AYA support organizations through social media, email, and events. Results – 589 initiated the survey, 445 (76%) completed the survey. The majority (85%) were female. 58% reported receiving therapy at a cancer center versus a community oncology hospital. 50% of patients were unsure if a clinical trial was available to them. 59% had completed therapy. 47% of respondents had children. 82% rated their care very good or excellent. 43% sought more information on long-term effects of cancer treatment. 94% had health insurance at diagnosis, and 98% had health insurance at the time of the survey; however, 11% reported that they had lapses in insurance coverage during or after treatment. Discussion and Conclusions – An online patient-reported survey may be a feasible option to study AYA patients and survivors. Early results indicate that online reporting may accelerate tracking of AYA cancer populations and create additional databases of

Poster Number	Title	Author and Affiliations	Abstract
			information. Data interpretation continues. Next steps include broadening CAYACC to include additional AYA programs and to establish a national AYA database.
73	A Hospital-Based Pediatric Cancer Registry tool and Deidentified Data Warehouse to Support Data Processing From Low- and Middle-Income countries (LMIC)	Nickhill Bhakta1, Alisha Gray1, Gia Ferrara1, Joanne Aitken2, Soad Fuentes-Alabi3, Tricia Alcasabas4, Olga Aleinikova5, Miguel Bonilla1, Oleg Budanov5, Jolie Caneba4, Lane Faughnan1, Yijin Gao6, Tezer Kutluk7, Paula Naidu1, Karina Braga Ribeiro8, Danny Youlden2, Carlos Rodriguez-Galindo1. 1St. Jude Children's Research Hospital, Global Pediatric Medicine, Memphis, USA. 2Cancer Council Queensland, Australia Childhood Cancer Registry, Brisbane, Australia. 3Benjamin Bloom Children's Hospital, Oncology, San Salvador, El Salvador. 4Philippine General Hospital, Pediatric Hematology & Oncology, Manila, Philippines. 5Belarussian Center for Pediatric Oncology and Hematology, Pediatric Oncology & Hematology, Minsk, Belarus. 6Shanghai Children's Medical Center, Hematology/Oncology, Shanghai, China. 7Hacettepe University, Medicine & Cancer Institute, Ankara, Turkey. 8Santa Casa de Sao Paulo, Medical Sciences & Public Health, Sao Paulo, Brazil.	Background: Hospital-based pediatric cancer registration is incomplete to non-existent in many LMIC. Where it does exist, limited guidance is available for what types of data these cancer registries should collect. To address this critical gap in global cancer control efforts, we developed consensus standards and a cloud-based registry tool for LMIC hospitals treating childhood cancer. Design: The registry tool was created using an iterative design approach consisting of defining requirements, prototyping and testing. The underlying design principle was to capture only relevant data needed for stakeholders to improve quality and monitor outcomes. Content, software and security requirements were independently determined. Potential data elements and variables were first identified through a scoping review and subsequently finalized for prototyping following two rounds of consensus input by a panel of pediatric oncologists and registry directors. Concurrently, requirements for software platform selection and security compliance were completed with input from information technology and legal specialists. Alpha (internal) and Beta (external) testing procedures were conducted in collaboration with a globally representative panel of pediatric cancer registry experts. Deidentification and warehouse storage of hospital registry data is automated. Results: Six adaptive electronic forms capturing essential information, demographics, diagnosis, staging, treatment and follow-up were developed. The number of elements captured per form varies from 2 to 40 dependent on previous answers. Data collection standards with definitions, values and logic are publicly available (https://stjude.org/content/dam/en_US/shared/www/clinical/no-index/sjcares-data-dictionary.pdf). Security measures such as sandboxing and multifactor authentication were incorporated to ensure compliance with American, European and Chinese data privacy laws. Language support includes English, Spanish, Portuguese, French, Russian and simplified Chinese. Conclusions: We developed standards and a cloud-based pediatric cancer registry tool optimized for LMIC hospitals. All data are contractually owned by the respective hospital while also deidentified and stored in a communal centrally-managed data warehouse. The resultant deidentified datasets have the potential to be the largest repository of childhood cancer outcomes data in the world and a global

Poster Number	Title	Author and Affiliations	Abstract
			resource for research and cancer control efforts. Free access to the tool and data are available through the St. Jude Global Alliance as part of the St. Jude Global Childhood Cancer Analytics Resource and Epidemiologic Surveillance System (SJCARES).
63	Mondo Disease Ontology: harmonizing disease concepts around the world	Nicole A. Vasilevsky, Research Assistant Professor, Oregon Health & Science University Julie A. McMurry, Associate Director, Translational and Integrative Sciences Lab, Oregon State University Shahim Essaid, Research Assistant Professor, Oregon Health & Science University Nomi L. Harris, Program Manager, Lawrence Berkeley National Laboratory Peter N. Robinson, Professor, Jackson Laboratory Christopher J. Mungall, Scientist, Lawrence Berkeley National Laboratory Melissa A. Haendel, Director, Translational and Integrative Sciences Lab, Oregon State University	Standards exist for describing gene variants (e.g. HGVS), but there is not a definitive standard for encoding diseases for information exchange. Existing sources of disease definitions include the National Cancer Institute Thesaurus (NCIt), the Online Mendelian Inheritance in Man (OMIM), SNOMED CT, ICD, ICD-O, OncoTree, MedGen, and numerous others. However, these standards partially overlap and often conflict, making it difficult to align knowledge sources - for example variant interpretation or drug responsiveness. This need to integrate information has resulted in a proliferation of mappings between disease entries in different resources; these mappings lack completeness, accuracy, and precision, and are often inconsistent between resources. The UMLS provides intermediate concepts through which other resources can be mapped, but these mappings also suffer from the same challenges: they are not guaranteed to be one-to-one, especially in areas with evolving disease concepts such as rare disease. Further, the UMLS is not intended for classification, for example, it contains cycles. In order to computationally utilize our collective knowledge sources for diagnostics and to reveal underlying mechanisms of diseases, we need to understand which terms are truly equivalent across different resources. This will allow integration of associated information, such as treatments, genetics, phenotypes, etc. We therefore created the Mondo Disease Ontology to provide a logic-based structure for unifying multiple disease resources. Mondo is created by a combination of algorithmic equivalency determination using the kBOOM algorithm, and expert curation. Mondo does provide equivalence mappings to other disease resources, but in contrast to other mapping sets, Mondo precisely annotates each mapping using strict semantics, so that we know when two diseases are precisely equivalent or merely closely related - allowing computational integration of associated data. Mondo, NCIt and the Human Phenotype Ontology (HPO) are used to structure disease and phenotype descriptions in a number of resources, such as ClinGen, GARD, the Monarch Initiative, and the Gabriella Miller Kids First Data Resource, which is a curated database of clinical and genetic sequence data from pediatric patients with structural abnormalities or childhood cancers.

Poster Number	Title	Author and Affiliations	Abstract
68	Clinical Implementation of OncoTarget and OncoTreat to Identify Computationally Inferred Master Regulator Targets to Improve Pediatric Cancer Management and Outcomes	Oberg JA1, Hsiao S2, Mansukhani M2, Pavisic J1, Califano A3, Yamashiro DJ 1,2 Departments of Pediatrics1, Pathology2, Systems Biology3, Columbia University Irving Medical Center, Herbert Irving Comprehensive Cancer Center, New York, NY	In pediatric oncology, current molecular strategies that primarily use next-generation DNA sequencing have had limited success in identifying targetable oncogenic mutations and the majority of patients have not benefited from more personalized or precise therapy. To improve the identification of targetable molecular findings, the Precision in Pediatric Sequencing (PIPseq) program, the Laboratory for Personalized Genomic Medicine and the Califano Systems Biology Laboratory at Columbia University Irving Medical Center (CUIMC) have integrated novel systems biology-based algorithms, OncoTarget/OncoTreat, to the established CLIA-compliant/New York State approved whole-exome and transcriptome sequencing pipeline. The algorithms are applied to the RNA-seq data generated during clinical testing to identify master regulator (MR) protein activity (OncoTarget) and MR-reversing therapies (OncoTreat) for individual patients. Specifically, ARACNe (Algorithm for the Reconstruction of Accurate Cellular Networks) is an information theoretic algorithm for reverse engineering regulatory networks from large gene expression datasets including TCGA, TARGET, and the European Registry neuroblastoma dataset. VIPER (Virtual Inference of Protein-activity by Enriched Regulon analysis) uses an ARACNe-inferred network and individual patient tumor gene expression profiles to computationally infer protein activity on a given sample based on the enrichment of the protein's regulatory targets in the tumor's differential gene expression signature. OncoTarget uses VIPER to produce a list of targetable MRs that are most aberrantly active in the tumor. OncoTreat uses VIPER and a large drug perturbation database which includes >10,000 RNA-seq profiles to prioritize FDA-approved or late-stage clinical trial compounds by their ability to reverse an individual's MR signature. MR proteins are rarely themselves mutated or differentially expressed, but are necessary for maintaining a cancer state. Pharmacological MR inhibition causes destabilization and abrogates tumor viability. CUIMC is at the forefront of applying this methodology to pediatric patients in a clinical setting. Identifying novel tumor dependencies that can be pharmacologically targeted represents the ultimate opportunity to optimize and personalize cancer treatment. Furthermore, interrogation of these data may identify recurrent MRs associated with specific tumor subtypes and infer risk-stratification or prognostication for select groups of patients. These MR checkpoints may also reveal novel pathways and inform the design of future clinical investigations in pediatric precision cancer medicine.
69	Targeting Master Regulator	Pavisic J1, Mundi P2, Glade Bender J3, Yamashiro DJ1,4, Califano A5	Outcomes in pediatric osteosarcoma remain poor. Mutation-based precision oncology approaches are limited by lack of targetable mutations and significant

Poster Number	Title	Author and Affiliations	Abstract
	Dependencies in Pediatric Osteosarcoma	Departments of Pediatrics1, Medicine2, Pathology4, Systems Biology5, Columbia University Irving Medical Center, Herbert Irving Comprehensive Cancer Center, New York, NY Department of Pediatrics3, Memorial Sloan Kettering Cancer Center, New York, NY	genetic heterogeneity. We leveraged systems biology approaches to discover common targetable disease drivers—master regulator proteins (MRs)—in osteosarcoma to expand molecular risk stratification and treatment options. Using the metaVIPER algorithm and pan-cancer gene regulatory network, we analyzed 84 osteosarcoma gene expression samples from TARGET to identify tumor-specific MR activity based on the enrichment of the protein's transcriptional targets in the tumor's differential gene expression signature. Hierarchical clustering by the 30 most differentially active MRs across the patients identified four groups with common MR architectures. Differential protein activity was seen in proteins involved in immune regulation (MNDA, IL16, TFEC, IFNG) and chemotherapy resistance (TOP2A, THB, RUVBL2). We found significant differences in rates of relapse (Chisquare test, p=0.04) and event-free survival (Log rank test, 48% vs. 87%, p=0.02) between patients in a cluster with low immune-regulatory protein activity as compared to those in a cluster where these proteins were highly active. Evaluating directly targetable MRs, we found twenty-seven patients with high DNMT3B and TOP2A activity, while 53 patients showed aberrant activation of matrix metalloproteases and LOXL2. These proteins have been previously successfully targeted. We subsequently used the OncoTreat algorithm to prioritize drugs by their ability to reverse patient-specific MR-activity signatures using a large drug-perturbation database, including >10,000 RNASeq profiles from drug perturbations. Patients clustered by predicted drug sensitivities pointing to potential novel therapies, such as Fludarabine, which showed preclinical efficacy in an osteosarcoma PDX model ($p < 0.0001$). Using innovative computational algorithms and available gene expression datasets, we find that a small set of MR subtypes recapitulate the regulatory landscape of osteosarcoma patients. MR signatures defining these subtypes provide mechanistic insight in to molecularly-based risk stratification and therapy selection which is muchneeded in osteosarcoma. We will refine this analysis through a multi-institutional study in which we will generate a cohort of >200 patient gene expression profiles, use osteosarcoma cell lines replicating the common MR architectures to optimize OncoTreat therapy prioritization, and further validate our findings in vivo in matching PDX models.
60	Leveraging Linked Birth and Disease registries to Characterize Cancer Etiologies and Survivorship	Philip J. Lupo1, Logan G. Spector2, Jeremy M. Schraw1, Erin L. Marcotte2, Beth A. Mueller3, Michael E. Scheurer1. 1Baylor College of Medicine, Houston, TX	Need and Methodology: Childhood, adolescent, and young adult (AYA) cancers remain a significant public health problem due to high cost o care, ineffective treatments for some subtypes, and chronic health conditions experienced by survivors. Significant gaps in our understanding of these cancers limit new prevention and treatment efforts. By routinely linking multiple existing data sources to

Poster Number	Title	Author and Affiliations	Abstract
	Among Children, Adolescents, and Young A	2University of Minnesota, Minneapolis, MN 3Fred Hutchinson Cancer Research Center, Seattle, WA	population-based cancer registries, we have a timely opportunity to generate a new and robust data resource for addressing these gaps. Applicability to Childhood/AYA Cancers: Cancer registries identify all cancer cases within a defined region (typically states in the US) and provide valuable insights into cancer incidence and survival. However, by linking cancer registry data to birth records, other disease registries (e.g., birth defects registries) and administrative data (e.g., Medicaid), we can augment these data for more comprehensive assessments. For instance, rich information on early-life events can be captured in birth records and birth defects registries. Whereas, information on chronic health conditions in survivors can be obtained by linking to Medicaid and hospital discharge data. Potential for Broader Use: As a first step to promote broader use of these data, we propose that cancer registries in the US annually link to their respective state-based birth records, relevant disease registries, and selected administrative databases for all individuals diagnosed with cancer under the age of 20 years. Given the large population size and the racial and ethnic diversity of the US, these routine linkages would: 1) allow for the assessment of multiple research questions across the cancer continuum; and 2) accelerate research progress of rare cancer subtypes and understudied populations. Evidence of Success: We recently established a population-based birth cohort of 10,181,074 children by linking and pooling statewide data on births, birth defects, and cancer from four US states. Several findings emerged, including children with non-chromosomal birth defects were almost 3-times more likely to be diagnosed with cancer compared to unaffected children, potentially making these conditions one of the strongest risk factors for developing childhood/AYA cancers. These and similar studies could elucidate novel pathways underlying cancer risk and ultimately inform clinical management for these individuals.
52	Positive Predictive Value and Sensitivity of ICD9 Codes for Childhood Leukemia in Integrated Health Plans	S Weinmann1, M Francisco1, EJA Bowles2, AK Rahm3, R Greenlee4, NK Stout5, JD Pole6, LH Kushi7, R Smith-Bindman8, DL Miglioretti9, ML Kwan7* 1 Center for Health Research, Kaiser Permanente Northwest, Portland, OR 2 Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle, WA	Background Tumor registries are excellent sources for identifying leukemia cases for research, but they may be unavailable or incomplete for some health care populations. ICD9 diagnosis codes from electronic medical records may be used to identify potential leukemia cases; however, their accuracy and completeness are unknown. We calculated positive predictive value (PPV) and sensitivity of ICD9 codes for childhood leukemia to evaluate these codes for leukemia case identification. Methods In the Radiation-Induced Cancer Study, six U.S. integrated health plans examined and validated ICD9 diagnosis codes associated with leukemia for years 1996-2015.

Poster Number	Title	Author and Affiliations	Abstract
		3 Center for Health Research, Genomic Medicine Institute, Geisinger, Danville, PA 4 Marshfield Clinic Research Institute, Marshfield Clinic Health System, Marshfield, WI 5 Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA 6 Dalla Lana School of Public Health, University of Toronto, Canada 7 Division of Research, Kaiser Permanente Northern California, Oakland, CA 8 Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA 9 Department of Public Health Sciences, University of California, Davis, CA	Subjects were children born in the health plans and enrolled continuously for at least 120 days (unless they died) after the date of the first leukemia ICD9 code or tumor registry diagnosis date. We counted the number of leukemia-related codes per subject within 120 days after the first code. We classified ICD9 codes as specific leukemia codes, myelodysplastic syndrome/myeloproliferative disorder codes, and non-specific codes for leukemia and related diseases. The referent standard was the health plan tumor registry and/or manual medical record review. We calculated PPV and sensitivity by the number of ICD9 codes in the 120-day period, overall and stratified on code categories, site, and demographic factors. We also explored adding chemotherapy and/or biopsy/aspiration codes to the algorithm. Results 652 subjects had at least one ICD9 code of interest; 293 (45%) were validated as leukemia. Among validated leukemia cases, 97% had at least one specific leukemia code; 2% had codes in other categories but no specific leukemia code. Requiring four or more ICD9 codes gave good PPV and sensitivity overall and across sites. Overall, PPV for four or more codes was 88% for all codes combined and 97% for leukemia-specific codes. The sensitivity for four or more codes was 94% for all codes combined and 92% for leukemia-specific codes. Requiring a chemotherapy/biopsy code improved the PPV (95%) of all codes combined, while maintaining good sensitivity (92%). Conclusion Using leukemia-specific ICD-9 codes or a broader group of leukemia-related ICD-9 codes plus chemotherapy/biopsy procedure codes can reliably identify cases of childhood leukemia in electronic medical records of integrated health plans.
55	The effects of funding cuts on data completeness in a multicenter pediatric cancer registry	Stephanie M. Perkins1, Sara L. Gallotto2, Benjamin V. Bajaj2, Miranda P. Lawell2, Daniel Indelicato3, Arnold Paulino4, Nadia Laack5, Ralph Ermoian6, John Perentesis7, William Hartsell8 ,Victor Mangona9, Christine Hill-Kayser10, Suzanne Wolden11, Young Kwok12, Michael Confer13, J. Ben Wilkinson14, Torunn I. Yock1 1Department of Radiation Oncology, Washington University in St. Louis, 2Department of Radiation Oncology, Harvard Medical School, 3Department of	Background/objectives: The Pediatric Proton and Photon Consortium Registry (PPCR) is a consented patient registry that opened to enrollment in 2012. Its primary aim is to expedite outcomes based research in pediatric cancer patients requiring radiation therapy. The PPCR collects extensive information on late effects and long-term health outcomes related to treatment. It was jointly funded by the NCI and Massachusetts General Hospital Federal Share Funds which are derived from profits of the first 20 years of the MGH Proton Center. Reimbursement for proton radiotherapy was cut in 2015 resulting in a sudden contraction of funds available in the MGH/NCI Federal Share dedicated to radiation research. Prior to these funding cuts each participating center in the PPCR received payment for each patient they enrolled. After June 30, 2015, each site was asked to continue to enroll but was no

Poster Number	Title	Author and Affiliations	Abstract
		Radiation Oncology, University of Florida, 4Department of Radiation Oncology, MD Anderson Cancer Center, 5Department of Radiation Oncology, Mayo Clinic, 6Department of Radiation Oncology, University of Washington, 7Division of Oncology, Cincinnati Children's Hospital Medical Center, 8Department of Radiation Oncology, Northwestern Medicine Chicago Proton Center,9Department of Radiation Oncology, Texas Center for Proton Therapy, 10Department of Radiation Oncology, University of Pennsylvania, 11Department of Radiation Oncology, ProCure Proton Therapy Center, New Jersey, 12Department of Radiation Oncology, University of Maryland, 13Department of Radiation Oncology, ProCure Proton Therapy Center, Oklahoma City, 14Department of Radiation Oncology, Provision Healthcare	longer reimbursed on a per patient basis. Here we report on the level of data completeness before and after funding cuts. MGH/NCI Federal share funding ends in November 2019. Results: The rate of missing information for all data fields prior to funding cuts was 4.2% and increased to 9% (2016) and 17.6% (2017). Due to high rates of missing data, we prioritized several "Critical Fields" (CF) for data entry in 2018. Missing data on CF in 2018 were 7%, 50% of which were due to missing follow-up. Institutions that joined without funding (n=8) have maintained higher rates of CF entry (95.7%) than those who had lost funding (n=5) (87.7%). Three originally funded institutions halted accrual until funds could be obtained. Conclusion: Overall, the PPCR maintains a rate of less than 10% missing critical data despite contraction of funding. While the overall rate of missing information is low, the percent attributable to follow-up is high, and unequal among institutions, highlighting a lack of resources. Consistent patient follow-up and data entry are essential to maintaining a robust patient registry centered on long-term outcomes. We are piloting direct to patient PROs through a REDCap survey to identify which patients need medical record reviews.
59	Using electronic medical record data to understand adverse outcomes: development of the Leukemia Electronic Abstraction of Records Network (LEARN)	Tamara P. Miller, MD, MSCE, Kelly Getz, PhD, M. Monica Gramatges, MD, PhD, Karen R. Rabin, MD, PhD, Marla H. Daves, MD, Jennifer J. Wilkes, MD, MSCE, Lena Winestone MD, MSHP, Phillip Lupo, PhD, Michael Scheurer, PhD, MPH, Evanette Burrows, MPH, Brian Fisher, DO, MSCE, Robert Grundmeier, MD, Richard Aplenc, MD, PhD	National Cancer Institute-funded cooperative group clinical trials have led to improved overall survival for children with cancer over the past five decades, and have set the standards of care for treatment of pediatric oncology patients. However, prior research demonstrated under-reporting of adverse events (AEs) using the currently employed AE reporting methods on pediatric cooperative oncology group trials. Moreover, the resources currently available to cooperative groups are not sufficient to address all of the unresolved clinical research questions in pediatric oncology. To address these issues, we have developed an open source automated data collection package that extracts data from the electronic medical record using R (ExtractEHR). We have deployed ExtractEHR at three institutions (Children's Hospital of Philadelphia, Children's Healthcare of Atlanta, and Texas Children's Hospital) and, together with Seattle Children's Hospital and the University of California San Francisco (implementation underway), developed a new consortium for data sharing via ExtractEHR: the Leukemia Electronic Abstraction of Records Network (LEARN). Operationally, after manual input of identified patients of interest, ExtractEHR

Poster Number	Title	Author and Affiliations	Abstract
			automatically extracts demographic data including residential addresses, laboratory and pathology results, radiology results, vital signs, administered medications, and procedures from the electronic medical record. Centralized pipelines for processing, cleaning and grading these data at the Children's Hospital of Philadelphia have been established, and are now complete for laboratory data and are in development for other data elements. Currently, 1,786 pediatric acute leukemia patients with 477,967 laboratory panels have been extracted and processed. Automatically extracted residential address histories have been geocoded using ArcGIS and linked to various georeferenced data resources including the US Census and American Community Surveys. Using these data, work is ongoing to determine baseline AE rates in up-front pediatric acute lymphoblastic and acute myeloid leukemia treatment protocols. These data are also being used to determine the impact of neighborhood factors on laboratory abnormalities, including renal or liver dysfunction, coagulopathy, and high disease burden at presentation with leukemia. We believe ExtractEHR can be easily scaled up and are currently working to disseminate ExtractEHR to other institutions to support Children's Oncology Group trials and to support clinical epidemiology and comparative effectiveness studies.
18	The Pediatric Brain Tumor Atlas (PBTA): a Multi-Institution, Collaborative Pediatric and AYA Data Generation and Sharing Initiative to Accelerate Discovery and Clinical Translation	The Children's Brain Tumor Tissue Consortium (CBTTC) and Pacific Pediatric Neuro-Oncology Consortium (PNOC), submitted by Adam Resnick on behalf of member institutions Center for Data-Driven Discovery in Biomedicine, Children's Hospital of Philadelphia, Philadelphia, PA	Pediatric brain tumors are the leading cause of disease-related death in children and despite advances, morbidity and mortality rates remain poor. Large-scale data-generation efforts to profile these cancers in the past have been limited, with public, rapid access to data also limited. The Children's Brain Tumor Tissue Consortium (CBTTC) and Pacific Pediatric Neuro-Oncology Consortium (PNOC) comprise two consortia with overlapping membership of more than 30 national and international research institutions respectively focused on pediatric biospecimen-based data generation paired with longitudinal clinical data curation and accelerated precision-based clinical trial design and implementation via molecularly-informed therapeutic intervention in children diagnosed with brain tumors. Together, the two consortia recently launched the Pediatric Brain Tumor Atlas (PBTA), an open science initiative aimed at creating a first-in-kind multi-omic, comprehensive molecular and clinically rich dataset for brain tumors. The first dataset release of PBTA includes over 30 brain tumor histologies representing from over 1,000 subjects. The data have been released pre-publication and without embargo, and have been made available via the Kids First DRC Portal and PedcBioPortal. Data types comprise those for matched tumor/normal samples, including WGS, RNASeq with paired longitudinal clinical data, treatment/outcome data, imaging data (MRIs and radiology reports), histology slide images, operative reports, and pathology reports. Over the past year, the PBTA and

Poster Number	Title	Author and Affiliations	Abstract
			associated data-sharing infrastructure has further served as an integration hub for disease specific efforts with partnered consortia-based initiative and foundations including the Chordoma Foundation, Oligo Nation, and Project OPEN DIPG as well as the integration of data-driven initiatives across a continuum of research that includes both pediatric and adolescent and young adult patients, exemplified by the recently launched NCI-sponsored single-cell high grade glioma sequencing pilots, Project Hope and Project Care. The combination of cloud-based analytic platforms with near real-time large-scale data releases with longitudinal clinical annotation across a collaborative network of participating hospital that span pediatric and AYA patients defines a new paradigm for the acceleration of brain tumor research via collaborative discovery.as a community-based resource directed at accelerated clinical translation.
54	Use of national claims data to inform AYA oncology: focus on end of life	Wernli KJ,1 O'Meara ES,1 Delaney K, 1 Chen L,1 Farjah F,2 Beatty T, Mack JW3 1Kaiser Permanente Washington Health Research Institute, Seattle, WA 2University of Washington, Seattle, WA 3Dana Farber Cancer Institute, Boston, MA	Background: National claims data on AYA oncology patients are limited, constraining care quality and improvement for this overlooked population. Data: We acquired health insurance claims data (records of services billed and paid to health insurance) from Optum's de-identified Clinformatics® Data Mart Database, representing a large US health insurer. The dataset includes >260,000 individuals with a prevalent ICD9/ICD10 code for malignant cancer at age 15-39 years in 2004-2016. Inpatient and outpatient claims include provider type, dates, and billing codes. We identified 4,217 deaths, including month and year of death, and further linked 699 individuals to Centers for Disease Control (CDC) Vital Status for date and cause of death. Potential for broader use: Our novel linked dataset builds on the strengths of claims data (trend analysis and short-term outcomes) and adds robust death data from CDC. Evidence of success: Our analysis focused on end of life trends in AYA. The mean death age was 33.6 years, 51% were women and 60% died in 2001-2008 (vs. 2009-2016). CDC linkage identified 98.6% of known deaths. Cancer was leading cause of death. In the last 30 days of life, 16.3% had >1 ED visit, 78.4% were hospitalized, 40.1% had an ICU stay, and 4.5% received surgery; 12% received chemotherapy within 14 days of death. An increasing proportion of patients experienced >1 ED visit (p=0.01) or ICU stay (p=0.004) over time, with no difference in receipt of surgery or chemotherapy. We observed regional variation for hospitalizations (p=0.008) and ICU stays (p=0.003) but not surgery or chemotherapy. Only 30% had billing for palliative/hospice care (n=1286); of these, nearly 50% had their first claim within 30 days before death. Receipt of palliative/hospice care 30-180 days before death was significantly

Poster Number	Title	Author and Affiliations	Abstract
			associated with lower rates of ICU and >1 ED visits (last 30 days of life) and chemotherapy (last 14 days). We did not see this pattern for hospitalization. Critical lessons learned: Claims data offer opportunities to examine national trends in end-of-life care for AYA patients, an understudied and underserved population. However, claims data often lack exact date and cause of death. Linkage to CDC data allows more robust examination of care.
61	Sorting Through the Data: Creating Evaluation Metrics for an AYA Cancer Program—the University of Iowa Health Care Experience	William Terry, Justin Kahler, Zachary Pollock, George Weiner University of Iowa Health Care	University of Iowa Health Care (UIHC) includes the Holden Comprehensive Cancer Center (HCCC) and the Stead Family Children's Hospital (SFCH). In 2015, the UIHC Adolescent and Young Adult (AYA) Cancer Program was created to address the unique clinical and psycho-social needs of patients ages 13-39. A comprehensive, collaborative effort capitalizing on the expertise of multiple services across the institution, the UIHC AYA program is based equally in the SFCH and HCCC, serving as a consultative and coordination service. Metrics were sought to measure the program impact. Nationally, AYA Cancer programs are relatively new, housed in a variety of settings, and offer varying services. There were no published, standardized metrics available for AYA Cancer Programs that served as a model. Several steps were taken to identify both goals and metrics for the program. First was a gap analysis to identity the number and types of patients that would be served, as well as institutional resources. Feedback sessions were held with AYA patients and families to obtain their perspectives on program needs. A five-year strategic plan was developed based on these, in alignment with the institutional Mission, Goals and Objectives. This plan included 10 program priorities encompassing the comprehensive nature of the services offered. Once these priorities were identified, metrics were created to align with planned activities. Metrics tracking was complicated by the multiple locations where these data are housed, including the EMR, cancer registries, billing databases and AYA-program databases. Collating these data proved to be time consuming and error-prone. To streamline the process, formal AYA program metrics were narrowed to the data required for day-to-day operations and program evaluation. Ongoing efforts focus on developing an automated abstraction of these data from the EMR via a Tableau Server. In conclusion, many AYA programs are in institutions with multiple data sources. To efficiently collect data for AYA program management and evaluation, a program strategic plan can serve as a guide for data needs. Further narrowing of data points

Poster Number	Title	Author and Affiliations	Abstract
			allow for more efficient, accurate and timely data collection and analysis. Follow up is needed to determine the utility of such metrics.

Other Abstracts of Relevance

Poster Number	Title	Authors and Affiliations	Abstract
90	Development and Implementation of the Duke Teen and Young Adult Oncology Database and Psychosocial Assessment	Caroline Dorfman, PhD[1], Cheyenne Corbett, PhD[2], Mallori Thompson, MA[2], Lars Wagner, MD[3], Geoffrey Vaughn, MA[2], Gary Maslow, MD[1], David Van Mater, MD[3] [1]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center [2]Duke Cancer Institute, Duke University Medical Center [3]Department of Pediatrics, Duke University Medical Center	Cancer is the leading cause of disease-related death among teens and young adults (TYA). Unfortunately, TYAs have not seen the same improvement in cancer-related outcomes as younger children and older adults. Multiple hypotheses have been put forth to explain the lack of improvement in outcomes including differences in underlying tumor biology, low enrollment in clinical trials, delayed access to care, emerging independence/lack of compliance, and other psychosocial pressures specific to the TYA age group. The unique needs of TYA patients have resulted in a call to action to enhance care. The Duke TYA Oncology Database was borne out of a shared commitment by Duke Cancer Institute (DCI) and Duke Children's Hospital and Health Center (Duke Children's) to improve the care of oncology patients aged 15-29. The goal of the database is to capture clinical data for TYA oncology patients seen across Duke Health to support the following areas of interest: 1) assess treatment outcomes and increase clinical trial enrollment; 2) provide appropriate resources for patients as they bridge into survivorship; 3) promote better utilization of oncofertility resources; and 4) perform longitudinal assessment of psychosocial metrics. Implementation of data collection requires a coordinated effort across two distinct medical entities within Duke Health, DCI and Duke Children's. As such, we have developed an implementation plan to facilitate patient identification and enrollment with the input of key stakeholders from each environment. Patients will be approached by a dedicated patient navigator or by medical and psychosocial health providers who serve DCI and/or Duke Children's. Interested patients will have their demographic and clinical information entered into the database, which will facilitate identification of available clinical trials and allow us to track metrics such as utilization of fertility resources and administration of neurocognitive testing. Participants enrolled in the database will also complete two psychosocial assessments [e.g., psychological distress, physical symptoms, peer relationships, cognitive functioning] approximately 6 months apart to identify and track patients' needs over time with the goal of developing more effective services. Lessons learned in developing our implementation plan may be valuable to other programs looking to systematically track and better serve TYA patients.
93	Diverse supportive care needs of Adolescent and	Dori L. Klemanski, DNP, APRN-CNP Samantha Hulett, MSW, LISW-S Olanipekun Lanny Ntukidem Annie Trance, MSW, LISW-S Bhavana Bhatnagar, MD Maryam Lustberg, MD, MPH	Introduction: Despite advances in early detection and treatment survival rates and quality of life outcomes have not improved for Adolescents and Young Adult (AYA) cancer survivors. A cancer diagnosis during this life stage causes significant interruption with developmental milestones and unique challenges through

Poster Number	Title	Authors and Affiliations	Abstract
	Young Adult cancer survivors		survivorship. We aimed to assess AYA specific care needs of survivors during and after cancer treatment. Methods: A retrospective chart review was performed on AYA cancer survivors (ages 18 to 39) who completed active treatment at a large academic medical center between January 2016 and April 2019. Results: During the specified time period, 236 AYA cancer survivors were provided a Treatment Summary and Survivorship Care Plan. The cohort of AYA survivors were predominately female (78%), Caucasian (91%), and older than 30 (78%). Cancer diagnoses varied but were commonly melanoma (30%) or breast (29%). The majority were employed (67%) or students (7%). A review of health behaviors indicated that 30% of AYA survivors reported a smoking history with a subset actively smoking (9%); the majority (57%) reported alcohol intake; and use of illicit substances, usually marijuana, was noted in 10% of AYA survivors. The majority were established with primary care (84%) with common co-morbidities of depression/anxiety (36%) and obesity (mean Body Mass Index of 29.93 kg/m2). AYA survivors averaged 0.8 referrals to clinical support, which increased to 3.8 referrals if linked with the Survivorship clinic (51%). Highly utilized clinical services included physical therapy (35%), psychosocial oncology (25%), cancer genetics (24%), and gynecology/fertility (21%). Additionally, AYA survivors participated in healthy living classes and 'meet-ups' (n = 72) like 'Tacos & Mocktails' to facilitate social connections and learn about nutrition. Conclusions: Providing AYA cancers survivors with access to support services to address unmet and ongoing needs may be crucial to their quality of life and well-being. The unique needs of this population have been overlooked, as the majority of research to date combines AYAs with either adult or pediatric groupings. Additional research is needed to develop evidenced based AYA specific supportive care programs.
94	Multi-institutional Pediatric Patient Data Registry Evolving Toward Big Data Knowledgebase and Analytics	Kyung-Wook Jee[1], Sara L. Gallotto[1], Benjamin V. Bajaj[1,] Miranda P. Lawell[1,] Torunn I. Yock[1.] [1]Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA	Given that pediatric cancers comprise only 1% of all cancers, single-institutional studies are unable to provide adequate sample sizes necessary to elucidate the impact of treatment on long-term outcomes. To address this need for a collaborative multi-institutional effort, US pediatric radiation oncologists have worked together to contribute patient data to the Pediatric Proton/Photon Consortium Registry (PPCR). Over the years, the data infrastructure of the PPCR has evolved by adopting modern communication and data science tools that combine distributed sources of clinical and therapeutic data from multiple hospitals into a single federated data source. PPCR utilizes two main, HIPAA-compliant, cloud-based electronic data capture (EDC) methods; REDCap (PA Harris, Vanderbilt) for structured and unstructured clinical

Poster Number	Title	Authors and Affiliations	Abstract
			data/patient reported outcomes; and MIMcloud (MIM Software Inc.) for DICOM imaging and radiotherapy data. Cloud-based EDC significantly lowers barriers for registry participants to record and access to the main database in a traceable and controlled manner. Even at a local hospital level, aggregating data siloed across different sources, transforming and loading them to the EDC is still resource-intensive and often manual. To reduce such efforts, PPCR is building open-source software tools, called RO-Data, that performs ETL (extract-transform-load) from the two most-widely used radiation oncology information systems (ROIS): Elekta's MOSAIQ or Varian's ARIA. The goal is to provide a user-friendly method to pull data from ROIS without a need of a special data manager. The tool is open-sourced so that it can be transparently tailored to the specific build of the ROIS used by PPCR sites. The centralized data sources at PPCR will be further federalized by building a logical data layer with cancer treatment-specific ontologies. Radiation oncology specific ontology is being developed and the sematic data model will be broadened for treatment related measurements and outcomes. This is a vital step to make PPCR data findable, accessible, interoperable, re-usable (FAIR) for machines as well as other cancer registries beyond the pediatric registry. This will lead to create a high-quality knowledgebase with a large quantity of treatment data and accelerate the use of modern data analytics for outcome research and improve the quality control of pediatric cancer treatments.
92	Collecting Uniform Information on Health Care Costs, Financial Burden and Quality of Care for Adolescent and Young Adult (AYA) Cancer Research	Sapna Kaul,[1] Yong-Fang Kuo,[1] and Anne C. Kirchhoff[2] [1] Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, TX [2] Huntsman Cancer Institute and Department of Pediatrics, University of Utah, Salt Lake City, UT	Description of relevant methodology: NCI's Childhood Cancer Data Initiative (CCDI) should focus on essential yet ignored elements of AYA cancer care: institutional costs of care, and financial burden and quality of care from patients' perspective. Our poster will identify specific measures and existing data examples that CCDI consider for collecting data on these elements from participating organizations. Applicability to AYA cancers: Clinicians may lack awareness regarding costs of care and financial toxicity/burden that accrues to AYA cancer patients; elements increasingly considered important for improving clinical effectiveness/value.[1, 2] Our previous studies show AYA survivors face adverse financial outcomes (e.g., cost-related medication non-adherence, foregoing care) due to cancer treatments, [3, 4] and they are less likely to report receiving high-quality health care.[5] Even after the Affordable Care Act, young adults in the U.S. are at a higher risk of being underinsured (spending more than 10% of income on health care needs).[6] Potential for broader use in research: Limited data exists on costs of treatment, financial outcomes and care quality on a national scale for AYAs with cancer.[6, 7] In datasets frequently used for studying financial/care quality outcomes (e.g., Medical

Poster Number	Title	Authors and Affiliations	Abstract
			Expenditures Panel Survey), identifying AYAs with cancer and examining their financial outcomes is limited because of unavailability of clinical data (e.g., diagnosis age and stage). In contrast, we have well-defined Medicare data for older adults. Medicaid is also limited for studying AYA costs because of patients moving in and out and lack of consistent follow-up. At the organizational/hospital level, we lack a uniform system of data management and collection on treatment costs and financial burden/toxicity for AYA cancer patients. Moreover, there is no uniform way to assess quality of health care from AYA cancer patients' perspective. Capturing these data uniformly through the CCDI will be important for conducting future cost-effectiveness and value-based assessments of AYA cancer care through the continuum of diagnosis, treatment and survivorship. Evidence of success: Treatment costs and financial/quality outcomes data are needed to examine effectiveness in cancer care.[1, 2, 6, 7] Provision of such data from CCDI will be critical for future national assessments that ensure high-quality AYA cancer care.
91	The advantages of patient-derived orthotopic xenograft (PDOX) models for individualized osteosarcoma treatment and drug discovery	Takashi Higuchi [1, 2, 3], Kentaro Igarashi [3], Norihiko Sugisawa [1, 2], Jun Yamamoto [1, 2], Yoshihiko Tashiro [1, 2], Hiroto Nishino [1, 2], Yasunari Fukuda [1, 2], Norio Yamamoto [3], Katsuhiro Hayashi [3], Hiroaki Kimura [3], Shinji Miwa [3], Hiroyuki Tsuchiya [3] , Zhenfeng Duan [4], Francis Hornicek [4], Robert M Hoffman [1, 2] 1. AntiCancer, Inc., San Diego, California 2. Department of Surgery, University of California, San Diego, California 3. Department of Orthopaedic Surgery, Kanazawa University, Kanazawa, Japan 4. Department of Orthopaedic Surgery, University of California, Los Angeles, California	Osteosarcoma is the most common primary malignant bone tumor which occurs mainly in children and adolescents. Due to rarity, heterogeneity, metastatic potency, and poor response rates to conventional systemic therapy, individualized precision oncology and novel drug discovery in osteosarcoma are warranted. Toward this goal, our laboratory has established the patient-derived orthotopic xenograft (PDOX) model using surgical orthotopic implantation (SOI). In the case of osteosarcoma PDOX, the primary tumors are implanted in the femur and metastatic tumors are implanted in the corresponding place of nude mice. Many promising results have been obtained using the sarcoma PDOX model for identifying effective approved drugs and experimental therapeutics, as well as combinations of them for individual patients. We have demonstrated that single administration of some drugs approved for the other cancer, e.g. temozolomide or trabectedin, were effective on a chemotherapy-resistant osteosarcoma PDOX tumor. Some combinations of approved drugs, e.g. irinotecan-trabectedin, irinotecan-temozolomide, sorafenib-palbociclib, everolimus-sorafenib, and olaratumab-doxorubicin-cisplatinum, could regress chemotherapy-resistant osteosarcoma PDOX tumors. We have demonstrated efficacy of experimental therapy such as tumor-targeting Salmonella typhimurium A1-R and recombinant methioninase could eradicate cisplatinum-resistant lung metastasis in an osteosarcoma PDOX model, when combinations of these agents were administrated with cisplatinum. In an era of increasing promise of novel systemic treatment and precision oncology, PDOX models present a unique opportunity to

Poster Number	Title	Authors and Affiliations	Abstract
			provide osteosarcoma patients with specific and personalized therapy and novel treatment options.

Governance and Authorities Associated with Establishing a Pediatric Data Ecosystem

Poster Number	Title	Authors and Affiliations	Abstract
87	Utilization of newborn dried blood spots to elucidate causes of childhood cancer	Erin L. Marcotte, Philip J. Lupo, Catherine Metayer, Beth A. Mueller, Michael E. Scheurer, Logan G. Spector	Description of need: The causes of most childhood cancers are poorly understood. Etiologic research is hindered by many factors, including the rarity of childhood cancer subtypes, selection bias in case-control studies, and recall bias in questionnaire-based exposure assessments. Therefore population-based sources of data and biologic specimens collected prior to cancer onset, such as residual dried blood spots (DBS) from newborn screening, are invaluable to advancing epidemiologic research. Newborn screening is a mandatory nationwide public health program that tests all babies for selected genetic, endocrine, and metabolic disorders. Blood spots are collected from each child, and the residual DBS are stored for confirmation of positive findings or in the event that screening must be repeated. Storage duration varies by state, from 30 days to indefinite. Applicability to childhood cancers: Given the rarity of childhood cancers, cohort studies for prospective collection of data and biologic specimens suffer from low accrual of cases. Therefore, DBS constitute the only population-based source of pre-diagnostic biospecimens for cases and representative controls. Analytes detectable in DBS include amino acids, enzymes, viral DNA, antibodies, inflammation markers, steroids, metals, small molecules, protein adducts, and some drugs. DBS may also be used in genomic and epigenomic studies to assess germline genetic susceptibility and DNA methylation profiles. Potential for broader use: Ideally, DBS from all US infants would be cataloged and stored such that they could be retrieved and released with appropriate scientific justification and ethics board oversight. State policies are rapidly evolving and there is ongoing discussion regarding DBS storage and secondary research use. Currently, population-based DBS studies can be conducted in a very limited number of states. Evidence of success: DBS were utilized in groundbreaking studies establishing the *in utero* origin of childhood leukemia, as well as research on hormone concentrations and testicular germ cell tumor, neonatal cytokine profiles and acute lymphoblastic leukemia (ALL), and *in utero* cytomegalovirus (CMV) infection and ALL. We are currently utilizing DBS samples in the first genome-wide association study of hepatoblastoma. We have amassed 962 germline DNA samples of this extraordinarily rare tumor, of which 420 are from DBS.

Poster Number	Title	Authors and Affiliations	Abstract
86	A Multimodal Framework for Pediatric Brain Tumor Research Utilizing 4.3 Million Events Over 16,000 Patient Years from a Clinical Data Research Network	Felmeister, Alex S., Ph.D.[1,4]; Waanders, Angela J., MD, MPH[2]; Mason, Jennifer L.[1]; Stevens, Jeff[3]; Bailey, L. Charles, MD[1], Ph.D.; Helbig, Ingo, MD[1]; Leary, Sarah E.S. MD[3]; Hu, Xiaohua, Ph.D.[4] [1]The Children's Hospital of Philadelphia, [2]Ann & Robert H. Lurie Children's Hospital of Chicago, [3]Seattle Children's Hospital, [4]Drexel University	Introduction: While biological/molecular studies can identify underlying etiologies in large patient cohorts and can be performed on an industrial scale, the ability to analyze phenotypes at scale is lagging specifically in pediatric tumor research. Longitudinal clinical data is critical to define outcome measures for individualized therapy. However, the lack of methods to analyze this data represents a significant gap. Objectives: This research applies a robust framework to extract specific clinical data: demographics, clinical presentation, treatment, and outcomes from observational data across two institutions. We provide a generalizable framework for the systematic analysis of complicated, longitudinal clinical features in pediatric cancer. Methods: We use a threefold pipeline of exploratory data analysis, multi-modal data transformation, and predictive analytics to derive a data-driven phenotype from a sample of over 1900 confirmed brain tumor cases for ontologically-based categorization while focusing specifically on 90 Primitive neuroectodermal tumor (PNETs) cases. Results: We built pipelines to make machine learning-ready data sets based on domain ontologies ready for enumeration, vectorization, and can be used to create humanunderstandable data points from 4.3 million observational events across 16,000 patient days. Conclusion: We address the gap in phenotypic data features by utilizing large harmonized observational clinical data and identify resources and specific processes for their use in rare tumor research. Implications: Genetic testing is standardized and scalable, but analysis of associated extensive observational clinical data at the same scale is lacking. This research takes the first steps to establish repeatable processes in clinical data and apply it to pediatric cancer research.
88	The Case For Modern Data Governance to Facilitate Participant-Centric Research for Childhood Cancer	Juergen Klenk, PhD - Principal, Deloitte Consulting LLP Dina Mikdadi, MPA - Data Scientist, Deloitte Consulting LLP Jessica Lo - Systems Architect, Deloitte Consulting LLP Amina Jackson, MS - Data Engineer, Deloitte Consulting, LLP	Data governance is a challenge given rapidly evolving technology, complicated privacy regulations, and the need to manage multiple data workflows. This is especially true for childhood cancer, where data are scarce and extremely valuable. When research organizations adopt a managed framework for governing valuable data, the efficient curation and sharing of childhood cancer research is possible. To foster better research collaboration and efficiency, we propose a comprehensive data governance framework for childhood cancer that can accommodate secure and highly advanced technologies to track, manage, and analyze diverse data in a reproducible manner. Current data stewardship models, previously suited for studies with fewer data sources and less robust technologies, must be enhanced to manage heterogeneous data, a plethora of tools, and analytical code. Indeed, advanced technology platforms today offer the ability to integrate multiple data sources by:

Poster Number	Title	Authors and Affiliations	Abstract
			● Effectively tracking biospecimens in real-time ● Reliably applying the latest computational methods ● Processing clinical, genomic, EHR, and imaging data. The promise of such data integration can remain elusive, however, in the absence of a data governance framework that systematically brings together people, process, and technology. As part of this framework, data experts work collaboratively, soliciting buy-in from patients and researchers at the beginning of the data lifecycle to address data standards, access policies, and more, thereby improving the transparency of research studies. Then, a team of data managers, computational biologists, and clinical informaticians work alongside data and security engineers to facilitate Findable, Accessable, Interoperable, and Reusable (FAIR) data. Collectively, the team iteratively monitors and improves upon data governance, ensuring smooth integration across multiple work streams. A strong data governance framework of this kind allows researchers to focus on research participants. We have seen our clients at large public and private healthcare institutions benefit from such a modern approach to data management. For example, data insights were streamlined for a pharmaceutical client after centralizing data and implementing a proper data governance framework. We believe improvements like these can also address some core challenges in childhood cancer research to unlock more meaningful insights, expedite discoveries, and, in turn, create a healthier future for children around the world.