

Childhood Cancer Data Initiative

Written by Joe Flores-Toro

The [Childhood Cancer Data Initiative](#) (CCDI) was introduced during the 2019 President's State of the Union address as a groundbreaking 10-year, \$50 million per year federal investment aimed at advancing pediatric cancer research. Its primary objective is to facilitate both new data generation and the development of platforms for data sharing. CCDI strives to build a Data Ecosystem that connects researchers, advocates, and care communities, allowing for rapid and systematic data collection, sharing, access, and analysis. This initiative encompasses patients from infancy to age 39 diagnosed with a pediatric cancer, with the goal of accelerating progress in etiology, prevention, diagnosis, treatment, outcomes, quality of life, and survivorship discoveries for pediatric and adolescent and young adult (AYA) cancer patients.

Childhood cancers constitute a small fraction (approximately 1%) of all newly diagnosed cancer cases. However, what makes pediatric cancer unique is the existing culture of data sharing within the research community, which eases the implementation of new data sharing structures. Additionally, the relatively low number of cases compared to adult cancers makes pediatric cancer data sets more manageable for testing these new data systems. By demonstrating the success of data sharing in pediatric cancer, similar approaches can be adapted to improve outcomes for all cancer patients, including adults.

The [Molecular Characterization Initiative](#) (MCI) is a significant component of CCDI, conducted in collaboration with the Children's Oncology Group (COG), and stands as an exemplar of all the CCDI aims to accomplish. MCI provides, free of charge, molecular characterization for pediatric and AYA cancer patients diagnosed with certain cancers (with plans to expand as time goes on). These tests include tumor/normal whole exome sequencing, methylation arrays, Archer fusion panels, and deidentified clinical report, all performed in a CLIA-certified setting. The returned diagnostic molecular profiles can aid in refining diagnoses and treatment plans, help determine clinical trial eligibility, and support genetic counseling. Sharing of linked molecular and clinical data generated by the MCI enables hypothesis testing across various cancer types and histologies and empowers the research community to posit new questions for research. Currently, opportunities to expand the MCI are being evaluated to and improve the usefulness of returned data and how those data could be shared across organizations so decisions by institutional tumor boards, or outcomes of novel treatments, could prospectively inform future treatment choices and clinical investigations. Collectively, the MCI provides a means to generate a new, high value, data set with the intent to share those data and the infrastructure to accomplish this built in from the ground up. An updated version was released 8/31/23, and now includes data on 1,552 patients which be accessed for FREE via dbGaP ([phs002790](#)).

The MCI is only one CCDI resource available to the community. CCDI currently offers a many other resources for researchers and data scientists interested in pediatric cancer research. Anchoring the initiative is the [CCDI Hub](#), a pivotal portal catering to researchers, data scientists, and citizens keen on making a difference. This hub not only serves as the gateway to the CCDI's rich ecosystem but also offers a treasure trove of tools and applications meticulously aligned with the CCDI's overarching mission. It allows users to explore a diverse array of resources, each finely tuned to advance the cause of pediatric cancer research. The CCDI Hub goes beyond mere information

provision, providing in-depth descriptions of these resources, each strategically addressing specific facets of childhood cancer research. Crucially, it simplifies access by offering direct links, ensuring that users can seamlessly engage with the data and technical insights necessary for their research endeavors.

Complementing this is the [CCDI Data Catalog](#), a dynamic repository encompassing more than 250 pediatric data sets contributed by 53 participating resources. This invaluable resource houses an expansive spectrum of data, spanning biospecimens, clinical records, genomics, and more. It operates as a living repository, consistently refreshed with new data sets and revised versions to keep researchers at the cutting edge of knowledge. Notably, the CCDI Data Catalog doesn't merely present a list of data sets but empowers users with crucial contextual information. It offers direct links for accessing specific data sets, insight into participant demographics, data types, and participant numbers. Furthermore, it supports the exportation of data set lists, facilitating seamless integration into research projects. As such, the CCDI Data Catalog serves as a dynamic wellspring, enabling researchers to leverage an ever-expanding repository of data, catalyzing innovation and breakthroughs in the realm of pediatric cancer research.

The CCDI's commitment to enhancing research extends further through the [National Childhood Cancer Registry \(NCCR\) Explorer](#). The NCCR Explorer represents a pivotal development under the CCDI, leveraging existing cancer registries to identify and track childhood cancer cases across the United States. Drawing from data derived from up to 24 NCCR registries, encompassing a remarkable 69% of the U.S. population, it delivers incidence and survival statistics for cancers in children and adolescent and young adults aged 0-39, spanning the years from 1999 onwards. This invaluable resource facilitates detailed exploration, providing cancer statistics based on factors such as sex, race/ethnicity, and age while enabling cross-comparisons across various cancer sites and subsites. In doing so, the NCCR Explorer empowers researchers, clinicians, and advocates with critical insights to inform and advance childhood cancer care, ultimately contributing to better patient outcomes. With these robust resources at their disposal, the CCDI's mission to improve pediatric cancer research and treatment gains powerful momentum.

The Childhood Cancer Data Initiative is a transformative program designed to advance pediatric cancer research through data sharing and collaboration. Leveraging the existing culture of data sharing in pediatric cancer research and initiatives like the Molecular Characterization Initiative, CCDI serves as a pioneering model that can potentially improve outcomes for all cancer patients. Researchers, data scientists, and advocates can access CCDI's resources, such as the CCDI Hub, Data Catalog, and NCCR Explorer, to contribute to the mission of understanding and treating pediatric and AYA cancers more effectively. CCDI's ongoing efforts promise a brighter future for young cancer patients and their families. To learn more about CCDI, please refer to our [recent publication in the Journal of Clinical Oncology](#).