

Authors: Gavriel Matt, Edgar Sioson, Jaimin Patel, Jian Wang, Robin Paul, Colleen Reilly, Congyu Lu, Kyla Shelton, Qi Liu, Weiyu Qiu, Cindy Im, Zhaoming Wang, Carmen L. Wilson, Nickhill Bhakta, Kirsten Ness, Gregory T. Armstrong, Melissa M. Hudson, Leslie L. Robison, Jinghui Zhang, Yutaka Yasui, Xin Zhou

St. Jude Survivorship Portal: A data portal for storing, analyzing, and sharing large and complex cancer survivorship datasets

Survivors of childhood cancer are at risk for developing various adverse health conditions as adults that are attributable to the cancer and treatments they were exposed to as children. Cancer survivorship research relies on large-scale, longitudinal studies that generate a wide range of demographic, clinical, and genetic data on cancer survivors at multiple time points. To maximize the utility of these comprehensive datasets, we must be able to store and share these datasets in a web-based environment that can be accessed by the broader survivorship research community. Furthermore, this environment should be integrated with analytical tools for performing statistical analyses on the stored data without needing to download the data and import it into third-party analytical software. To address this need, we have created the St. Jude Survivorship Portal (<https://survivorship.stjude.cloud> > Clinical Data Browser), a web-based data portal for exploring, sharing, and analyzing data from survivors of pediatric cancer. The portal hosts data from two large cohorts of pediatric cancer survivors: the St. Jude Lifetime Cohort Study and the Childhood Cancer Survivor Study. The data stored on the portal consists of demographic data, clinical data, including cancer diagnosis, cancer treatment, clinical outcomes, and patient-reported data, and genetic data, including whole-genome-sequencing-derived genotypes and published polygenic risk scores computed for >500 traits. This data is organized hierarchically in a data dictionary that can be easily explored by the user. Charts and plots of variables can be quickly created, customized, and stratified with other variables, all within the portal environment. Statistical analyses, including cumulative incidence analysis and regression analysis, may also be performed within the portal. In cumulative incidence analysis, users can analyze the incidence of a variety of CTCAE-graded adverse events (e.g., cardiovascular dysfunction, neurological disorders, subsequent neoplasms) in survivors and can also compare them across different survivor populations defined by other variables. In regression analysis, users have the option to perform either a linear, logistic, or cox regression analysis and may use any of the demographic, clinical, or genetic variables on the portal as outcome or explanatory variables in the analysis. This analysis tool allows users to assess any risk factor associations within a survivor cohort and to generate predictive models for outcomes of interest. Lastly, we also provide the user with the option to download the data on the portal for use in any future analyses. The St. Jude Survivorship Portal provides a comprehensive, powerful, and easy-to-use interface for sharing and analyzing childhood cancer survivorship data that will serve as a valuable research tool for the broader survivorship research community.

Content of abstract:

Each abstract should contain (a) an introductory sentence indicating the purposes of the study; (b) a brief description of pertinent experimental procedures; (c) a summary of the new, unpublished data; and (d) a statement of the conclusions.

Abstract length limits:

The combined length of the abstract body, title, and tables may not exceed 2,600 characters, not including spaces and the author string.