

Data sharing for pediatric cancers

Last month, in a conference call held by the U.S. Department of Health and Human Services and National Institutes of Health (NIH), it was revealed that a large focus of President Trump's pledge to fund childhood cancer research will be genomic data sharing. Although the United States has only 5% of the world's pediatric cancer cases, it has disproportionately more resources and access to genomic information compared to low-income countries. We hope that the spotlight on genomic data sharing in the United States will galvanize the world's pediatric cancer community to elevate genomic data sharing to a level where its full potential can finally be realized.

Pediatric cancers are rare, affecting 50 to 200 children per million a year worldwide. Thus, with 16 different major types and many subtypes, no cancer center encounters large cohorts of patients with the same diagnosis. To advance their understanding of particular cancer subtypes, pediatric oncologists must have access to data from similar cases at other centers. Because subtypes of pediatric cancer are rare, assembling large cohorts is a limiting factor in clinical trials as well. Here, too, data sharing is the first critical step.

Typically, pediatric cancers don't have the number of mutations that make immunotherapies effective, and only a few subtypes have recurrent mutations that can be used to develop gene-targeted therapies. However, the abnormal expression level of genes gives a vivid picture of genetic misregulation, and just sharing this information would be a huge step forward. Using gene expression and mutation data, analysis of genetic misregulation in different pediatric cancer subtypes could point the way to new treatments.

A major challenge in genomic data sharing is the patient's young age, which frequently precludes an opportunity for informed consent. Compounding this, the rarity of subtypes requires the aggregation of patients from multiple jurisdictions, raising barriers to assembling large representative data sets. A greater percentage of children than adults with cancer participate in

research studies, and children often participate in multiple studies. However, this means that data collected on individual children may be found at multiple institutions, creating difficulties if there are no standards for data sharing.

To enable effective sharing of genomic and clinical data, the Global Alliance for Genomics and Health has developed the Key Implications for Data Sharing (KIDS) framework for pediatric genomics. The recommendations include involving children in the data-sharing decision-making process and imposing an ethical obligation on data generators to provide children and parents with the opportunity to share genomic and clinical information with researchers. Although KIDS guidelines are not legally binding, they could inform policy development worldwide.

To advance the sharing culture, along with the NIH, pediatric cancer foundations such as the St. Baldrick's Foundation and Alex's Lemonade Stand Foundation have incorporated genomic data-sharing requirements into their grants processes. Researchers and clinicians around the world have created dozens of pediatric cancer genomic databases and portals, but pulling these together into a larger network is problematic, especially for patients with data at more than one institution, as patient identifiers are stripped from shared

data. However, initiatives like the Children's Oncology Group's Project Every Child and the European Network for Cancer Research in Children and Adolescents' Unified Patient Identity may resolve this issue.

We urge the creators of pediatric cancer genomic resources to collaborate and build a real-time federated data-sharing system, and hope that the new U.S. initiative will inspire other countries to link databases rather than just create new siloed regional resources. The great advances in information technology and life sciences in the last decades have given us a new opportunity to save our children from the scourge of cancer. We must resolve to use them.

—Olena Morozova Vaske and David Haussler

“...analysis of genetic misregulation in different pediatric cancer subtypes could point the way to new treatments.”

Olena Morozova Vaske is an assistant professor in the Department of Molecular, Cell and Developmental Biology at the Genomics Institute of the University of California, Santa Cruz, Santa Cruz, CA, USA. olena@ucsc.edu

David Haussler is a Distinguished Professor in the Genomics Institute, in the Department of Biomolecular Engineering, and in the Howard Hughes Medical Institute at the University of California, Santa Cruz, Santa Cruz, CA, USA. haussler@ucsc.edu

Data sharing for pediatric cancers

Olena Morozova Vaske and David Hausler

Science **363** (6432), 1125.
DOI: 10.1126/science.aax2739

ARTICLE TOOLS

<http://science.sciencemag.org/content/363/6432/1125>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/363/6432/1164.full>
<http://stm.sciencemag.org/content/scitransmed/9/381/eaaf2968.full>
<http://stm.sciencemag.org/content/scitransmed/8/366/366ra161.full>
<http://stm.sciencemag.org/content/scitransmed/9/414/eaam9078.full>
<http://stm.sciencemag.org/content/scitransmed/11/483/eaau9732.full>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.