

The Gabriella Miller Kids First Pediatric Research Program: Plans for Phase 2

Childhood cancers and structural birth defects have profound, lifelong effects on children and their families. Throughout childhood, these conditions are leading causes of mortality. Moreover, a child with a birth defect is at a higher risk for childhood cancer, suggesting shared genetic pathways. The vision of Gabriela Miller Kids First (KF) program is to: *Alleviate suffering from childhood cancer and structural birth defects by fostering collaborative research to uncover the etiology of these diseases and supporting data sharing within the pediatric research community.*

Directors of NICHD, NCI, NHLBI, NHGRI, and the Common Fund provide program leadership. A Working Group of Program Officials representing many ICs along with the Common Fund provide operational oversight. In its first six years, FY15-20, Kids First pursued two broad initiatives:

1) Identified cohorts of children with childhood cancer and/or structural birth defects, and their families, for whole genome (DNA) sequencing: Cohorts have been identified through a competitive peer-reviewed application process and prioritized using criteria emphasizing the value of the data set to the pediatric research community. Thirty-nine structural birth defects or childhood cancer datasets representing over 37,000 genomes are available or in the pipeline for sequencing by Kids First-funded sequencing centers. Additional cohorts will be identified for sequencing in FY20 and FY21.

2) Built the Gabriella Miller Kids First Data Resource: The Data Resource is a large-scale data platform for sharing genomic and associated clinical and phenotypic data from Kids First participants. Researchers can access, search, analyze, and aggregate these data through the Kids First Data Resource Portal. By using integrated cloud-based workspaces in the CAVATICA analysis platform, researchers can cross-analyze multiple datasets. To date, >150 requests have been received to access Kids First data; and there have been at least 17 publications. Both categories are expected to increase. The portal enables investigators to access other relevant data, such as NCI's TARGET datasets for co-analysis with Kids First data. This interoperability greatly increases the utility of the Kids First Data Resource. Kids First has also partnered with other resources, like NHLBI's BioData Catalyst, NHGRI's AnVIL, and Common Fund's Data Ecosystem (CFCE) to develop new avenues of interoperability and strengthen this functionality.

Kids First has also been aided by collaborations including; partnership with ICs to support analyses of Kids First data, collaboration with the Common Fund's KOMP2 program developing mouse models of variants identified from Kids First data, a pilot program encouraging the use of cloud computational analysis, and storage of Kids First data through STRIDES/AWS funded by the Common Fund.

Phase 2: In FY22-24, Kids First activities will serve to increase the utility and impact of the Kids First Data Resource. Discussions have begun about relevant ICs/OD programs that might contribute to sustaining the resource after Common Fund support ends. A broad consultation was done, with input from the NIH Kids First Working Group, Kids First IC Directors, the Kids First Steering Committee, external Program Consultants, a survey distributed to Kids First and external investigators, as well as feedback from the public and patient families. Several consensus themes emerged providing direction for the next phase of the program. Within the authorized \$12.6M annual program budget, we propose three broad initiatives:

- 1) Additional generation of childhood cancer and structural birth defects-related -omics data**
- 2) Continue development and improvement of the Data Resource**
- 3) Expert-driven activities to increase the value of Kids First data**

Concerted efforts will also be made to continue and expand partnerships with other NIH data science efforts to enhance and broaden the Data Resource's interoperability. Together these initiatives will position the Kids First Data Resource as a valuable pediatric data resource - utilized by the research community, potentially sustained by multiple ICs, and serving to accelerate pediatric research to improve preventative measures, diagnostics, and therapeutic interventions.