Welcome to the Gabriella Miller Kids First RFA-RM-21-011 Pre-Application Webinar!

- Every participant is muted upon entry.
- To ask public questions, use the Q&A bar (right side of your screen). We encourage you to save these for the question periods.
- You can also use the “chat” service to send private messages to the host or presenters throughout the webinar.
- After the webinar, additional questions can be emailed to: valerie.cotton@nih.gov

This webinar will be recorded.
We will start at 12pm (EDT)
Gabriella Miller Kids First Pediatric Research Program

*Expert-Driven Small Projects to Strengthen* Gabriella Miller Kids First Discovery, RFA-RM-21-011

March 29, 2021
12:00 pm EDT
Public Webinar: April 13th!

Register here: https://nih.webex.com/nih/onstage/g.php?MTID=e7dcd35de768586aa8c80352b1f544601

4:00PM Introduction; NIH Kids First Staff
4:05PM Detection of novel genetic bases for congenital cranial dysinnervation disorders (CCDDs) by whole genome sequencing: Elizabeth Engle, MD
4:35PM: Kids First Data Resource Center
   - New Portal Updates
   - Getting Started with Kids First
5:05PM: Proteogenomic Analysis: Breakthroughs for pediatric cancer through cross-platform collaboration; Adam Resnick, PhD, Pei Wang, PhD, Francesca Petralia, PhD
   - 5:25PM: Questions & Answers
5:35PM: NIH Program Updates
5:50PM: Questions & Answers
Birth defects are associated with increased risk of cancer among children... suggesting shared genetic pathways

From: Association Between Birth Defects and Cancer Risk Among Children and Adolescents in a Population-Based Assessment of 10 Million Live Births
Vision

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.
Kids First: Phase 1
Kids First Major Initiatives

1. Identify & sequence cohorts of children with **childhood cancer and/or structural birth defects**.
2. Build the **Gabriella Miller Kids First Data Resource** to empower discovery.
Disorders of Sex Development
Congenital Diaphragmatic Hernia
Ewing Sarcoma
Structural Heart & Other Defects
Syndromic Cranial Dysinnervation Disorders
Cancer Susceptibility
Adolescent Idiopathic Scoliosis
Neuroblastomas
Enchondromatoses
Orofacial Clefts in Caucasian, Latin American, Asian & African, Filipino populations
Osteosarcoma
Familial Leukemia
Craniofacial Microsomia
Intersection of childhood cancer & birth defects
Microtia
Esophageal Atresia and Tracheoesophageal Fistulas
Kidney and Urinary Tract Defects

Nonsyndromic Craniosynostosis
Hemangiomas, Vascular Anomalies & Overgrowth
Bladder Exstrophy
Hearing Loss
Cornelia de Lange Syndrome
Intracranial & Exacranial Germ Cell Tumors
Fetal Alcohol Spectrum Disorders
Myeloid Malignancies + overlap with Down syndrome
Congenital Heart Defects & Acute Lymphoblastic Leukemia in Children with Down Syndrome
Structural Brain Defects
Structural Defects of the Neural Tube (Spina Bifida: Myelomeningocele)
CHARGE Syndrome
Laterality Birth Defects
T-cell Acute Lymphoblastic Leukemia
Pediatric Rhabdomyosarcoma
Valvar Pulmonary Stenosis
The Kids First Data Resource for Collaborative Discovery

Data Resource Portal
Entry point. Query, search, discover, build & visualize synthetic cohorts

Cavatica
Pull data from multiple sources into one workspace. Use notebooks, bring-your-own or use available workflows.

Knowledge Base Integrations (PedcBioPortal)
Integrations with existing curated/published data visualizations

Data Services
Model clinical data in FHIR-based data services and API for sharing layers of harmonized data

Framework Services
Index and point to files in the cloud (for approved users)
Use Case: Compare genetic variants of congenital heart defects & neuroblastoma

Anyone can register & login to the portal (via ORCID, Google). User agrees to terms.

In Explore Data, user searches the terms “heart” and “neuroblastoma”. Discovers data from children with congenital heart disease (KF) & neuroblastoma (KF & NCI TARGET).

User builds a synthetic cohort based on these criteria and can view summary & deidentified individual-level clinical, demographic, and phenotypic information.

Synthetic cohort is ported to the File Repository where user selects which genomic and histology image files they want to analyze.

User pushes genomic, clinical, and image data into Cavatica for analysis & visualization.

User has or applies for dbGaP access for genomic data.

User runs statistical analyses in notebooks.

User iterates through genomic workflows.
Kids First: Phase 2
Phase 2 Strategic Planning: 7 Consensus Recommendation Themes

1. Innovation: Resource, infrastructure, or tool development.
   Activities: Data Visualization tools; other tools for clinical/phenotypic data

2. Clinical/phenotypic data extraction, harmonization, & curation.
   Activities: Collect, extract, organize, curate, harmonize, and submit deep clinical and phenotypic data; annotate variants with pathogenicity, ClinGen scores.

3. Collaborative validation and discovery.
   Activities: Building synthetic cohorts; identify structural variants; test pipelines.
   *Engage trainees in data analysis projects**Bring users to the platform*

4. Integration and interoperability of external pediatric datasets.
   Activities: Using DRC workflow and best practices to harmonize external pediatric datasets; Building tools that can operate across multiple spaces

5. Consent and data sharing.
   Activities: Re-consenting cohorts in line with our data sharing expectations

6. Validation with model organisms.
   Activities: Validating KF findings/variants, deep phenotyping of animal models

NIH Council of Councils
September 11, 2020

Common Fund Concept Clearance: Gabriella Miller Kids First Pediatric Research Program: Plans for FY22-24 (Phase 2)
James Coulombe, Ph.D.
Chief, Developmental Biology and Structural Variation Branch
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Goal of Phase 2 Initiatives: Enhance the value and impact of the Kids First Data Resource to accelerate pediatric research to improve preventative measures, diagnostics, and therapeutic interventions.

The archived videocast of the Council of Councils meeting is publicly available and can be viewed here (Kids First discussion begins at 4:48:00).
The presentation materials are available here.
Phase 2 Initiatives Approved
$12.6M/year (FY22-24)

1) Additional generation of childhood cancer and structural birth defects-related -omics data
   ➢ Add epigenomic and proteomic assays

2) Continue development & improvement of the Data Resource
   ➢ Plan for sustaining the Data Resource beyond FY24

3) Expert-driven activities to increase the value of Kids First data
   ➢ Engage Kids First & community experts in activities such as integration, curation, and/or harmonization of rich clinical and phenotypic data
RFA-RM-21-011
Expert-Driven Small Projects to Strengthen Gabriella Miller Kids First Discovery (R03 Clinical Trial Not Allowed)
Purpose

Engage **experts** in a variety of activities that will **enhance the utility of childhood cancer and/or structural birth defects datasets generated by the Kids First program and associated resources.**

*These activities should strengthen **future** analyses with the ultimate goal of improving diagnostic capabilities and therapies for children and their families affected by these conditions.*
Goals

Build on and improve Kids First:
• These projects must use Kids First data, standards, infrastructure, and resources, as much as possible …

Address challenges or gaps

Collaborate & coordinate:
• Incorporate a diversity of complementary expertise
• Coordinate with X01s, DRC, etc, to build on, but **not** duplicate, existing efforts
• Create a profile in the portal

Share data & resources
Collaborate & Coordinate

Kids First X01 investigators:
- [https://kidsfirstdrc.org/partners/partners-investigators/](https://kidsfirstdrc.org/partners/partners-investigators/)
- [https://commonfund.nih.gov/kidsfirst/x01projects](https://commonfund.nih.gov/kidsfirst/x01projects)

Kids First DRC members:
- [https://kidsfirstdrc.org/about/](https://kidsfirstdrc.org/about/)

Kids First Sequencing Centers:
- [https://kidsfirstdrc.org/partners/portal-sequence-data/](https://kidsfirstdrc.org/partners/portal-sequence-data/)

- **Sharon Pion**
  - Baylor College of Medicine
  - Research Focus: Identifying novel cancer susceptibility mutations
  - [VIEW PROFILE ➤](#)

- **Jonathan Rios**
  - UT Southwestern Medical Center
  - Research Focus: Genomics of orthopaedic disease program
  - [VIEW PROFILE ➤](#)

- **Jun Shen**
  - Brigham and Women's Hospital
  - Research Focus: Hearing loss
  - [VIEW PROFILE ➤](#)

- **Azeez Butali**
  - University of Iowa
  - Research Focus: Craniofacial genetics
  - [VIEW PROFILE ➤](#)
Collaborate & Coordinate

Create & Share a Portal Profile

New to the Kids First Data Resource Portal? JOIN NOW PORTAL LOGIN

Search other members https://portal.kidsfirstdrc.org/memberPage
Data & Resource Sharing Expectations

- It is expected that any data (including resultant raw, derived, aggregated, and summary data), tools, workflows, and/or pipelines created or used with support from this FOA will be provided to the Kids First Data Resource Center to be shared with the wider scientific community, if not already part of the Data Resource, in a timely manner that would enable other researchers to use and build on for future research efforts.

- Example: If you create a new survey instrument to collect deeper data, or a tool to extract data from EHRs, the questions, code, and other documentation must be shared.

- Example data sharing plans: https://commonfund.nih.gov/kidsfirst/FAQ
Examples: Improve Discovery by...

1. Collecting, extracting, submitting deeper data or new data types associated with Kids First datasets
2. Harmonizing or processing data to promote cross-disease or cross-species (or cross-dataset analysis)
3. Portal analysis workflows to deploy within the Kids First Data Resource (e.g., CAVATICA)
4. Creating or integrating, a new or separate tool to federate with the Kids First Data Resource
5. Consenting for broader data sharing
1. Collecting, extracting, submitting deeper data or new data types associated with Kids First datasets
2. Harmonizing or processing data to promote cross-disease or cross-species (or cross-dataset analysis)
3. Portal analysis workflows to deploy within the Kids First Data Resource (e.g., CAVATICA)
4. Creating or integrating, a new or separate tool to federate with the Kids First Data Resource
5. Consenting for broader data sharing

Examples: Improve Discovery by...
Why FHIR?

- FHIR is the bridge from hospital-based data to research
- Enough **structure** for computational approaches/tools
  - Immediate structure for EHR-based data
  - Transformations to existing data models (e.g., OMOP)
  - Facilitate ingest >> populate into tools & search portals
  - Facilitate analysis across multiple/different datasets
- But enough **flexibility** to capture rich/complex data elements that may get lost in datasets “harmonized” to Common Data Models
  - Ability to expand data models beyond EHR and/or Common Data Models (e.g., research surveys, case report forms)
In Explore Data, user searches the term “kidney” and discovers data from children with congenital kidney defects (KF data) & COVID-related kidney complications (MIS-C).

User builds a synthetic cohort based on these criteria and can view summary & deidentified individual-level clinical, demographic, and phenotypic information.

Synthetic cohort is ported to the File Repository where user selects which imaging files they want to analyze (e.g. DICOM).

User pushes imaging files into FlyWheel for analysis.
Innovation across the Phenotypic Translational Divide Webinar Series

Curation with ontologies that support heterogenous data types in Kids First

Innovation across the Phenotypic Translational Divide
Webinar Series

Part 1
- Pediatric Cardiac Genomics Consortium (PCGC) - Betsy Goldmuntz
- Down Syndrome - Congenital Heart Disease - Joaquin Espinosa
- Down Syndrome-ALL & Rhabdomyosarcoma - Phillip Lupo
- Orofacial Clefts - Mary Marazita
- Adolescent Idiopathic Scoliosis - Carole Wise
- Neuropsychological Data Harmonization - Stephanie Sherman
- Enchondromatoses and Related Malignant Tumors - Nara Sobreira

Part 2
- Kidney and Urinary Tract Defects - Ali Gharavi, Columbia University
- Structural birth defects and childhood cancer - Xiang Wang, Children’s Hospital of Philadelphia
- Neuroblastoma - Sharon Diskin, Children’s Hospital of Philadelphia
- CHARGE syndrome - Donna Martin, Michigan Medicine
- Laterality Birth Defects - Stephanie Ware, Indiana University
- Structural Brain Defects, Neural Tube Defects - Joe Gleeson, University of California, San Diego

Webinar Information:
Part 3: Cross-Species Genotype-Phenotype Analysis

Part 3 Blog

- ZFIN
- Xenbase
- KOMP2/IMPC
- Monarch Initiative
- Integrated Canine Data Commons
Examples: Improve Discovery by...

1. Collecting, extracting, submitting deeper data or new data types associated with Kids First datasets
2. Harmonizing or processing data to promote cross-disease or cross-species (or cross-dataset analysis)
3. Portal analysis workflows to deploy withing the Kids First Data Resource (e.g., CAVATICA)
4. Creating or integrating, a new or separate tool to federate with the Kids First Data Resource
5. Consenting for broader data sharing
Seven Bridges CAVATICA: Use or Create Workflows

Build, use, optimize, share new or existing workflows

Tailor a workflow with a Workflow Editor

Researchers bring multiple datasets together and run workflows over the data in secure private workspaces

SevenBridges: https://docs.sevenbridges.com/
CWL: https://www.commonwl.org/

Build, use, optimize, share new or existing workflows

Tailor a workflow with a Workflow Editor

Researchers bring multiple datasets together and run workflows over the data in secure private workspaces

SevenBridges: https://docs.sevenbridges.com/
CWL: https://www.commonwl.org/
Federate a new (or existing) tool with the Data Resource!

**Data Resource Portal**
Entry point. Query, search, discover, build & visualize synthetic cohorts

**Cavatica**
Pull data from multiple sources into one workspace. Use notebooks, bring-your-own or use available workflows.

**Knowledge Base Integrations (PedcBioPortal)**
Integrations with existing curated/published data visualizations

**Data Services**
Model clinical data in FHIR-based data services and API for sharing layers of harmonized data

**Framework Services**
Index and point to files in the cloud (for approved users)
NIH Cloud Based Platforms Interoperability (NCPI)

Goal: Empower end-user analyses across platforms through federation & interoperability
NIH Researcher Auth Services (RAS)

Simplify researcher access to NIH data through federated **authentication** (linking user identity account; “passport”) and **authorization** (claim to access specific studies/datasets; “visa”)


*Adapted from Susan Gregurick, ODSS*
Systems Interoperation WG - Technical 1st Year Vision

**Portals**

- BioData CATALYST
  - Powered by Gen3

- AnVIL

- Kids First
  - Pediatric Research Program
  - Data Resource Center

- FHIR, PFB on FHIR
  - Gabriella Miller

- NIH
  - Cancer Research Data Commons

**Workspaces**

- AnVIL
- BD Cat
- KF
- CRDC

**Data**

- DRS
- AuthN
- AuthZ
- Search
- Search result handoff
- Analysis Results

- and other workspaces

- PFB Import
- DRS Client
Key NCPI Standards to Consider

NCPI:  https://datascience.nih.gov/nih-cloud-platform-interoperability

• Fast Healthcare Interoperability Resources (FHIR; NOT-OD-19-122)
  – Tools that exchange or handle Kids First clinical, phenotypic, or meta-data should consider implementing FHIR standards to interact with the Kids First Data Resource’s FHIR server and API

• NIH Researcher Auth Service (RAS; https://auth.nih.gov/docs/RAS/)
  – Tools accessing controlled-access data should consider implementing the ability to “consume” RAS for authentication and authorization

• Data Repository Service (DRS; https://github.com/ga4gh/data-repository-service-schemas)
  – Tools handling Kids First data (which is stored in the cloud) should consider implementing the ability to “consume” DRS for referencing data objects
Examples: Improve Discovery by...

1. Collecting, extracting, submitting deeper data or new data types associated with Kids First datasets
2. Harmonizing or processing data to promote cross-disease or cross-species (or cross-dataset analysis)
3. Portal analysis workflows to deploy within the Kids First Data Resource (e.g., CAVATICA)
4. Creating or integrating, a new or separate tool to federate with the Kids First Data Resource
5. Consenting for broader data sharing
Standard Data Use Limitations (DULs)

The NIH Genomic Data Sharing Policy “expects investigators generating genomic data to seek consent from participants for future research uses and the broadest possible sharing.”

• **General Research Use** → **Broadest**
• **Health/Medical/Biomedical**
• **Disease Specific:**
  
  When data use is restricted to a specific disease area, the data *cannot* be combined with a dataset with a different disease specific data use limitation.
Institutional Certification: controlled-access data

NIH expects the submitting institution(s) to select one of the three standard Data Use Limitations (DULs) for appropriate secondary use, or, if necessary, create a customized DUL. DULs are developed based on the original informed consent of the participant(s).

**Data Use Limitations**

<table>
<thead>
<tr>
<th>General Research Use</th>
<th>GRU</th>
<th>Use of the data is limited only by the terms of the Data Use Certification: these data will be added to the dbGaP Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health/Medical/Biomedical</td>
<td>HMB</td>
<td>Use of the data is limited to health/medical/biomedical purposes, does not include the study of population origin or ancestry.</td>
</tr>
<tr>
<td>Disease-specific [list disease]</td>
<td>DS</td>
<td>Use of the data must be related to the specified disease.</td>
</tr>
<tr>
<td>Other</td>
<td>(ENTER CUSTOMIZED TEXT, IF APPLICABLE)</td>
<td></td>
</tr>
</tbody>
</table>

**Additional modifiers to the standard DULs (e.g., Not-for-profit Use Only)** are based in the informed consent from the participants or in special knowledge.

**Data Use Limitation Modifiers (Optional)**

<table>
<thead>
<tr>
<th>IRB Approval Required</th>
<th>IRB</th>
<th>Requestor must provide documentation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Required</td>
<td>PUB</td>
<td>Requestor agrees to make results of studies using the data available to the larger scientific community.</td>
</tr>
<tr>
<td>Collaboration Required</td>
<td>COL</td>
<td>Requestor must provide a letter of collaboration with the primary study investigator(s).</td>
</tr>
<tr>
<td>Not-for-profit Use Only</td>
<td>NFU</td>
<td>Use of the data is limited to not-for-profit organizations.</td>
</tr>
<tr>
<td>Methods</td>
<td>MDS</td>
<td>Use of the data includes methods development research (e.g., development and testing of software or algorithms).</td>
</tr>
<tr>
<td>Genetic Studies Only</td>
<td>GSO</td>
<td>Use of the data is limited to genetic studies only.</td>
</tr>
</tbody>
</table>

“General Research Use” with no modifiers is expected for individual-level genomic data, unless specific uses are clearly prohibited in consent.
Consent Considerations

- The NIH Genomic Data Sharing Policy “expects investigators generating genomic data to seek consent from participants for future research uses and the broadest possible sharing.”

  - “Your samples, genomic data and health information will be stored and shared with other researchers. The samples and information will be available for any research question, such as research to understand what causes certain diseases (for example heart disease, cancer, or psychiatric disorders), development of new scientific methods, or the study of where different groups of people may have come from.”

- DUL definitions and considerations:
Other R03s and Analysis Opportunities...

These activities should strengthen Kids First X01 datasets and/or enhance the functionality of the Kids First Data Resource to improve or facilitate future analyses by the broader research community, but applications do not have to address specific analyses themselves.

For applicants seeking to analyze Kids First data, see PAR-19-375. For applicants seeking to analyze other Common Fund datasets see RFA-RM-21-007.
**R03-PAR (IC funds)**

*Small Research Grants for Analyses of Gabriella Miller Kids First Pediatric Research Data*

- Support analyses of Kids First X01 datasets
- R03
- IC funds (NICHD, NCI, NHLBI, NIAAA, and NIDCR)
- PAR (no set asides)
- Deadline: Standard Receipt Dates, 3 per year
- Combined direct cost budget for the two-year project period may not exceed $200,000
- Number of awards depends on IC

**R03-RFA (KF funds)**

*Expert-Driven Small Projects to Strengthen Gabriella Miller Kids First Discovery*

- Support activities that will enhance the utility of Kids First data
- R03
- Kids First funds
- RFA (set aside)
- Deadline: June 18, 2021
- Combined direct cost budget for the two-year project period may not exceed $200,000
- 9-14 awards are anticipated from this solicitation
FOAs for Data Analyses

• Common Fund R03: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-21-007.html
• NCI: Secondary Analysis and Integration of Existing Data to Elucidate the Genetic Architecture of Cancer Risk and Related Outcomes (Contact: rotunno@mail.nih.gov)
• NIDCR: Notice of Special Interest (NOSI) of NIDCR in Supporting Discovery, Characterization, and Mechanistic Study of Genetic Variants Underlying Dental, Oral, and Craniofacial Diseases and Conditions 
FOAs for Variant Validation

- To pursue collaborations with the Knockout Mouse Phenotyping Program (KOMP2), contact: [KidsFirstKOMP@nih.gov](mailto:KidsFirstKOMP@nih.gov)
FAQ #1

**Question**: I applied for the FY21 X01 sequencing opportunity, but I would like to apply for the expert-driven R03-RFA to work with colleagues to improve the associated dataset. When will X01 decisions be announced, and should I apply to RFA-RM-21-011?

**Answer**: Can you make a compelling argument that the activity improve Kids First regardless of the X01? Then yes, apply.
FAQ #2

**Question:** I am planning to propose building and optimizing a workflow that can deployed in CAVATICA. Are cloud credits available for testing/running the tool?

**Answer:** Cloud credits are currently only available to X01s via a separate review process. However, there is no guarantee cloud credits will be approved for this purpose, so you should incorporate these cloud costs into your R03 budget.

→ Cost estimate resources: [DRC FAQs](#)
FAQ #3

**Question:** What standards have been adopted by Kids First? What other standards might be relevant?

**Answer:** See attached list (will be posted on our FAQ page soon!)
Q & A

• Use the Q&A bar (lower right of your screen) to send your questions to “All Panelists”. We will read your questions out loud and answer them.

• You can also use the “chat” service to send private messages to the host or presenters.
Anyone can register & login to the portal to filter, search, visualize datasets (build synthetic cohorts)

Submit dbGaP Data Access Requests (DARs) for controlled-access data

https://portal.kidsfirstdrc.org/