



Data Harmonization and Interoperability of Existing Data Repositories:

The Kids First Data Resource Experience

Aug 2, 2022

Symposium on Pediatric Disaster Science

Adam Resnick / d3b.center

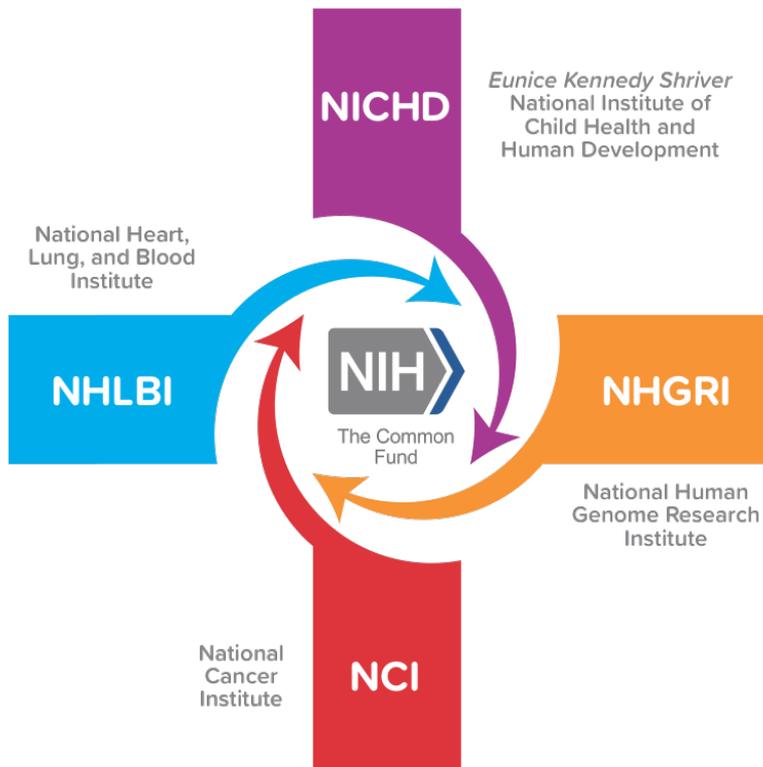


Follow @kidsfirstdrc

Visit kidsfirstdrc.org



Kids First Program Vision



Additional Working Group Representation

NIDCR	NIDA
NIAAA	NINDS
NIDDK	NIEHS
NEI	NIAMS
NIAID	NCATS
ORIP	CDC

The Kids First Program was launched via congressional action in 2015

Kids First Program Goal



The goal of the Gabriella Miller Kids First Pediatric Research Program (Kids First) is to help researchers **uncover new insights into the biology of childhood cancer and structural birth defects**, including the discovery of shared genetic pathways between these disorders.

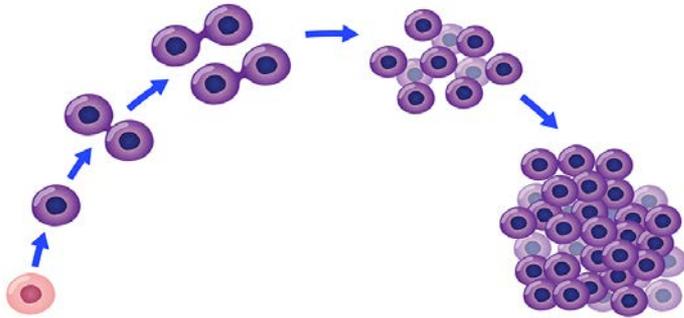


Alleviate suffering from childhood cancer and structural birth defects by fostering **collaborative** research to uncover the etiology of these diseases and supporting **data sharing** within the pediatric research **community**.

Can cross-cancer and cross-disease research empower/accelerate discovery?



Pediatric Cancer



+

Development





Kids First: By the Numbers



PI – Led Clinical Cohorts

Clinical Data
(EHR, diagnoses, demographics,
outcomes...)



Sequencing Centers

*WGS/RNAseq
Sequencing Data*



Kids First DRC *Data Harmonization*

- Sequencing Center CRAMs/FASTQ
- Match clinical data with sequencing samples and harmonize phenotypes for portal exploration



CAVATICA *Delivery & Release*

All Samples:

- Aligned Reads in CRAM
- Individual gVCF

Germline Data

- Family joint-called VCF (1 per family)

Somatic Data

- Somatic Variants, Copy Number, Structural Variants
- RNAseq - Expression/Fusions

Gabriella Miller Kids First Data Resource Center



28,285

Participants

20,222

Families

148,361

Files

1.52 PB

of Data

30 Studies Released on the Portal

- Kids First: Orofacial Cleft: African and Asian Ancestry
- Kids First: Neuroblastoma
- Kids First: Myeloid Malignancies
- Kids First: Esophageal Atresia & Tracheoesophageal Fistulas
- Kids First: Disorders of Sex Development
- Kids First: Leukemia & Heart Defects in Down Syndrome
- Kids First: Novel Cancer Susceptibility in Families (BASIC3)
- Kids First: Congenital Diaphragmatic Hernia
- Kids First: Enchondromatosis
- Kids First: Orofacial Cleft – European Ancestry
- Kids First: T Cell ALL
- Kids First: Microtia – Hispanic
- Kids First: Intersections of Cancer & SBD
- Kids First: Orofacial Cleft – Latin American
- Kids First: Familial Leukemia
- Kids First: Craniofacial Microsomia
- Kids First: Syndromic Cranial Dysinnervation
- Kids First: Kidney and Urinary Tract Defects
- Kids First: Adolescent Idiopathic Scoliosis
- Kids First: Hemangiomas (PHACE)
- Kids First: Ewing Sarcoma
- Kids First: Nonsyndromic Craniosynostosis
- Kids First: Congenital Heart Defects
- Kids First: Osteosarcoma



The Gabriella Miller Kids First Program

Big Data + Cloud Platforms + Communities = Ecosystems

Outside this room in the lobby . . .



Realities and metaphors



MOTHERLY



[PREGNANCY](#) [PARENTING](#) [LIFE](#) [SHOP](#) [CLASSES](#) [VIDEO](#) [PODCASTS](#)

[Home](#) / [Life](#) / [Motherly Stories](#)

I never thought I'd be a 'cancer mom'— here's what I want other parents to know

Cancer comes in like a tornado and tears your life apart, and very few people understand that. They tell you to be strong or worse—to stay positive. Some friends show up in the most beautiful way and others stop showing up altogether. The losses hit hard.

By [Danielle Moss](#) • July 14, 2022



Realities and metaphors



MOTHERLY

Search...



PREGNANCY PARENTING LIFE SHOP CLASSES VIDEO PODCASTS

[Home](#) / [Life](#) / [Motherly Stories](#)

I never thought I'd be a 'cancer mom'— here's what I want other parents to know

Cancer comes in like a tornado and tears your life apart, and very few people understand that. They tell you to be strong or worse—to stay positive. Some friends show up in the most beautiful way and others stop showing up altogether. The losses hit hard.

By [Danielle Moss](#) · July 14, 2022



Realities and metaphors



Science Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 372, NO. 6543 > RADIATION-RELATED GENOMIC PROFILE OF PAPILLARY THYROID CARCINOMA AFTER THE CHERNOBYL ACCIDENT

RESEARCH ARTICLE f t in g o e

Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident

LINDSAY M. MORTON , DANIELLE M. KARYADI , CHIP STEWART , TETIANA I. BOGDANOVA , ERIC T. DAWSON , MIA K. STEINBERG , JIEQIONG DAI , STEPHEN W. HARTLEY , SARA J. SCHONFELD , [...] STEPHEN J. CHANOCK +30 authors [Authors Info & Affiliations](#)

SCIENCE • 22 Apr 2021 • Vol 372, Issue 6543 • DOI:10.1126/science.aba92538

1,783 7 CHECK ACCESS

Genomics of radiation-induced damage

The potential adverse effects of exposures to radioactivity from nuclear accidents can include acute consequences such as radiation sickness, as well as long-term sequelae such as increased risk of cancer. There have been a few studies examining transgenerational risks of radiation exposure but the results have been inconclusive. Morton *et al.* analyzed papillary thyroid tumors, normal thyroid tissue, and blood from hundreds of survivors of the Chernobyl nuclear accident and compared

Realities and metaphors



Science

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 372, NO. 6543 > RADIATION-RELATED GENOMIC PROFILE OF PAPILLARY THYROID CARCINOMA AFTER THE CHERNOBYL ACCIDENT

RESEARCH ARTICLE

Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident

LINDSAY M. MORTON, DANIELLE M. KARYADI, CHIP STEWART, STEPHEN W. HARTLEY, SARA J. SCHONFELD, [...] STEPHEN W. HARTLEY

SCIENCE • 22 Apr 2021 • Vol 372, Issue 6543 • DOI:10.1126/science.1251126

1,783 7

Genomics of radiation-induced papillary thyroid carcinoma

The potential adverse effects of exposure to ionizing radiation can include acute consequences such as cancer and other sequelae such as increased risk of cardiovascular disease and transgenerational risks of radiation-induced cancer. Morton *et al.* analyzed papillary thyroid carcinoma blood from hundreds of survivors of

RARE DISEASES > TYPES

Children of Chernobyl: Birth Defects, Deformities, Ailments

Understanding short and long-term health issues caused by the nuclear meltdown

By Mary Engler, BN | Updated on July 15, 2022

Print

Fact checked by James Lacy

Table of Contents

1986 Nuclear Accident

Immediate Health Consequences

Long-term Health Consequences

What the Future Holds

In 1986, a meltdown at the Chernobyl nuclear power plant in what was then the Soviet Union showered radioactive particles all over Ukraine and neighboring countries. It resulted in the largest uncontrolled release of radiation (iodine-131, cesium-134, and cesium-137) in history.^[1]

More than 6,000 cases of thyroid cancer linked to radiation exposure in people who were under age 18 at the time were reported between 1991 and 2015.^[2] They came to be known as the Children of Chernobyl, and both researchers and the public have remained interested in their challenges.

This article discusses them and their health impacts in adulthood. It looks at Children of Chernobyl birth defects and deformities, as well as genetic mutations—and some of the more positive outcomes despite the early dire predictions.



Realities and metaphors



Science Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 372, NO. 6543 > RADIATION-RELATED GENOMIC PROFILE OF PAPILLARY THYROID CARCINOMA AFTER THE CHERNOBYL ACCIDENT

RESEARCH ARTICLE

Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident

LINDSAY M. MORTON, DANIELLE M. KARYADI, CHIP STEWART, STEPHEN W. HARTLEY, SARA J. SCHONFELD, [...] STEPHEN W. HARTLEY

SCIENCE • 22 Apr 2021 • Vol 372, Issue 6543 • DOI:10.1126/science.1254000

1,783 7

Genomics of radiation-induced papillary thyroid carcinoma

The potential adverse effects of exposure to ionizing radiation can include acute consequences such as cancer. Long-term sequelae such as increased risk of cancer and other health issues can be passed on to future generations. Morton *et al.* analyzed papillary thyroid carcinoma blood from hundreds of survivors of the Chernobyl nuclear accident.

RARE DISEASES • TYPES

Children of Chernobyl: Birth Defects, Deformities, Ailments

Understanding short and long-term health issues caused by the nuclear meltdown

By Mary Engler, RN | Updated on July 15, 2022
Fact checked by James Lacy

Table of Contents

- 1986 Nuclear Accident
- Immediate Health Consequences
- Long-term Health Consequences
- What the Future Holds

In 1986, then the neighborly radiation... More people in 2015, research... This article at Child mutation prediction...

JAMA Network

JAMA Pediatrics Search All Enter Search Term

Review

March 2017

Characterizing the Pattern of Anomalies in Congenital Zika Syndrome for Pediatric Clinicians

Cynthia A. Moore, MD, PhD; J. [...] CDC

» Author Affiliations | Article

JAMA Pediatr. 2017;171(3):288-294

Abstract

Importance Zika virus infection has provided evidence to conclude that it is a congenital infection, but the full spectrum of anomalies and how to evaluate and treat affected children with congenital Zika syndrome.

Baby with Typical Head Size Baby with Microcephaly Baby with Severe Microcephaly



GeoHealth

RESEARCH ARTICLE

10.1029/2021GH000548

Key Points:

- This study analyzed the relationships of pediatric cancer rates to an aggregate of several potentially carcinogenic metals and pesticides
- An environmental burden index (EBI) was constructed and categorized the environmental profiles of counties in Idaho as low, medium, and high
- A statistical model was developed using the EBI to estimate pediatric cancer incidence in Idaho

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. S. Kolok,
akolok@uidaho.edu

Citation:

Joseph, N., & Kolok, A. S. (2022). Assessment of pediatric cancer and its relationship to environmental contaminants: An ecological study in

Assessment of Pediatric Cancer and Its Relationship to Environmental Contaminants: An Ecological Study in Idaho



Naveen Joseph¹  and Alan S. Kolok¹ 

¹Idaho Water Resources Research Institute, University of Idaho, Moscow, ID, USA

Abstract The primary aim of this study was to determine the degree to which a multivariable principal component model based on several potentially carcinogenic metals and pesticides could explain the county-level pediatric cancer rates across Idaho. We contend that human exposure to environmental contaminants is one of the reasons for increased pediatric cancer incidence in the United States. Although several studies have been conducted to determine the relationship between environmental contaminants and carcinogenesis among children, research gaps exist in developing a meaningful association between them. For this study, pediatric cancer data was provided by the Cancer Data Registry of Idaho, concentrations of metals and metalloids in groundwater were collected from the Idaho Department of Water Resources, and pesticide use data were collected from the United States Geological Survey. Most environmental variables were significantly intercorrelated at an adjusted P -value < 0.01 (97 out of 153 comparisons). Hence, a principal component analysis was employed to summarize those variables to a smaller number of components. An environmental burden index (EBI) was constructed using these principal components, which categorized the environmental burden profiles of counties into low, medium, and high. EBI was significantly associated with pediatric cancer incidence (P -value < 0.05). The rate ratio of high EBI profile to low EBI profile for pediatric cancer incidence was estimated as 1.196, with lower and upper confidence intervals of 1.061 and 1.348, respectively. A model was also developed in the study using EBI to estimate the county-level pediatric cancer incidence in Idaho (Nash-Sutcliffe Efficiency = 0.97).



GeoHealth

RESEARCH ARTICLE

10.1029/2021GH000548

Key Points:

- This study analyzed the relationships of pediatric cancer rates to an aggregate of several potentially carcinogenic metals and pesticides
- An environmental burden index (EBI) was constructed and categorized the environmental profiles of counties in Idaho as low, medium, and high
- A statistical model was developed using the EBI to estimate pediatric cancer incidence in Idaho

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. S. Kolok,
akolok@uidaho.edu

Citation:

Joseph, N., & Kolok, A. S. (2022). Assessment of pediatric cancer and its relationship to environmental contaminants: An ecological study in

Assessment of Pediatric Cancer and Its Relationship to Environmental Contaminants: An Ecological Study in Idaho

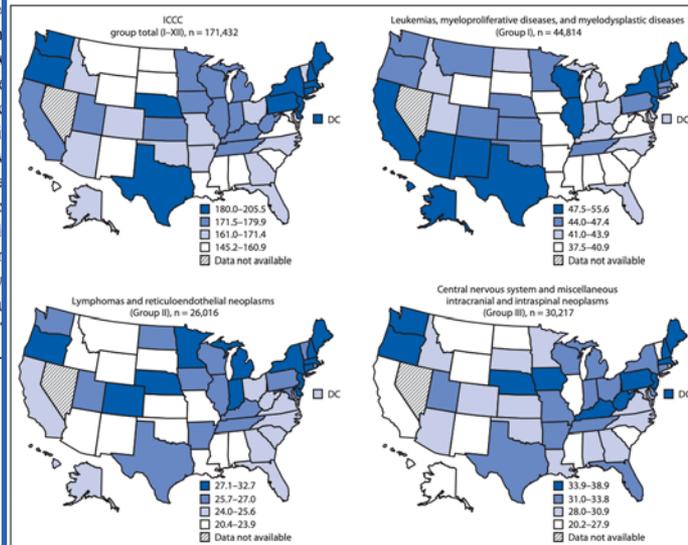
Naveen Joseph¹ and Alan S. Kolok¹

¹Idaho Water Resources Research Institute

Abstract The primary aim of this study was to assess the relationship between environmental contaminants and pediatric cancer rates across the United States. A component model based on several environmental variables was used to estimate level pediatric cancer rates across the United States. This study is one of the reasons for increased pediatric cancer rates that have been conducted to determine among children, research gaps exist. Pediatric cancer data was provided by the National Cancer Institute. Metalloids in groundwater were collected from the United States. The data were intercorrelated at an adjusted P -value analysis was employed to summarize the environmental burden index (EBI) was constructed to categorize burden profiles of counties into low, medium, and high incidence (P -value < 0.05). The relationship was estimated as 1.196, with low incidence. A statistical model was also developed in the study (Nash-Sutcliffe Efficiency = 0.91).

Geographic Variation in Pediatric Cancer Incidence — United States, 2003–2014

David A. Siegel, MD^{1,2}; Jun Li, MD, PhD²; S. Jane Henley, MSPH²; Reeda J. Wilson, MPH²; Natasha Buchanan Lunsford, PhD²; Eric Tai, MD³; Elizabeth A. Van Dyne, MD^{1,2}



See figure footnotes on next page.



GeoHealth

RESEARCH ARTICLE

10.1029/2021GH000548

Key Points:

- This study analyzed the relationships of pediatric cancer rates to an aggregate of several potentially carcinogenic metals and pesticides
- An environmental burden index (EBI) was constructed and categorized the environmental profiles of counties in Idaho as low, medium, and high
- A statistical model was developed using the EBI to estimate pediatric cancer incidence in Idaho

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. S. Kolok,
akolok@uidaho.edu

Citation:

Joseph, N., & Kolok, A. S. (2022). Assessment of pediatric cancer and its relationship to environmental contaminants: An ecological study in

Assessment of Pediatric Cancer and Its Relationship to Environmental Contaminants: An Ecological Study in Idaho

Naveen Joseph¹  and Alan S. Kolok¹ 

¹Idaho Water Resources Research Institute

Abstract The primary aim of this study was to develop a component model based on several environmental factors to estimate level pediatric cancer rates across Idaho. This study is one of the reasons for increased pediatric cancer rates that have been conducted to determine the relationship among children, research gaps exist in pediatric cancer data was provided. Heavy metals and metalloids in groundwater were collected from the United States and intercorrelated at an adjusted P -value analysis was employed to summarize the environmental burden index (EBI) was constructed and categorized the burden profiles of counties into low, medium, and high incidence (P -value < 0.05). The model was estimated as 1.196, with low incidence was also developed in the study (Nash-Sutcliffe Efficiency = 0.9).

Geographic Variation in Pediatric Cancer Incidence —



Journal of
Clinical Medicine



Article

COVID-19 in Children with Down Syndrome: Data from the Trisomy 21 Research Society Survey

David Emes^{1,†}, Anke Hüls^{2,3,†} , Nicole Baumer⁴, Mara Dierssen^{5,6,7}, Shiela Puri⁸, Lauren Russell², Stephanie L. Sherman⁹, Andre Strydom^{10,11,12}, Stefania Bargagna¹³, Ana Cláudia Brandão¹⁴, Alberto C. S. Costa¹⁵ , Patrick T. Feany², Brian Allen Chicoine¹⁶ , Sujay Ghosh¹⁷, Anne-Sophie Rebillat¹⁸, Giuseppina Sgandurra^{13,19} , Diletta Valentini²⁰, Tilman R. Rohrer²¹, Johannes Levin^{22,23,24}, Monica Lakhampaul^{25,26,27,*}  and on behalf of the Trisomy 21 Research Society COVID-19 Initiative Study Group[§]



See figure footnotes on next page.

Realities and metaphors



GeoHealth

RESEARCH ARTICLE

10.1029/2021GH000548

Key Points:

- This study analyzed the relationships of pediatric cancer rates to an aggregate of several potentially carcinogenic metals and pesticides
- An environmental burden index (EBI) was constructed and categorized the environmental profiles of counties in Idaho as low, medium, and high
- A statistical model was developed using the EBI to estimate pediatric cancer incidence in Idaho

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. S. Kolok,
akolok@uidaho.edu

Citation:

Joseph, N., & Kolok, A. S. (2022). Assessment of pediatric cancer and its relationship to environmental contaminants: An ecological study in

Assessment of Pediatric Cancer and Its Relationship to Environmental Contaminants: An Ecological Study in Idaho

Naveen Joseph¹ and Alan S. Kolok¹

¹Idaho Water Resources Research Institute

Abstract The primary aim of this study was to develop a component model based on several level pediatric cancer rates across Idaho. This study was one of the reasons for increased research in this area. Research has been conducted to determine the relationship between pediatric cancer data and environmental factors. Data were provided by the Idaho Department of Health and Welfare. Data were collected from the United States Cancer Statistics. An analysis was employed to summarize the data. A burden index (EBI) was constructed based on the profiles of counties in Idaho. The incidence (P -value < 0.05). The incidence was estimated as 1.196, with low incidence was also developed in the study (Nash-Sutcliffe Efficiency = 0.9).

Geographic Variation in Pediatric Cancer Incidence —



Journal of
Clinical Medicine

Article

COVID-19 in Children with Trisomy 21 Research Study

David Emes^{1,†}, Anke Hüls^{2,3,†}, Nicole Sherman⁴, Stephanie L. Sherman⁵, Andre Strydom⁶, Alberto C. S. Costa¹⁵, Patrick T. Feany⁷, Giuseppina Sgandurra^{13,19}, Diletta Valente⁸, Monica Lakhanpaul^{25,26,27,*} and on behalf of the Study Group[§]

See figure footnotes on next page.



COMMENTARY JOURNALS COVID-19 Science

COVID-19 is 10 times deadlier for people with Down syndrome, raising calls for early vaccination

Extra chromosome carries key genes that may work against them

15 DEC. 2020 · BY MEGHANA WISNAMA

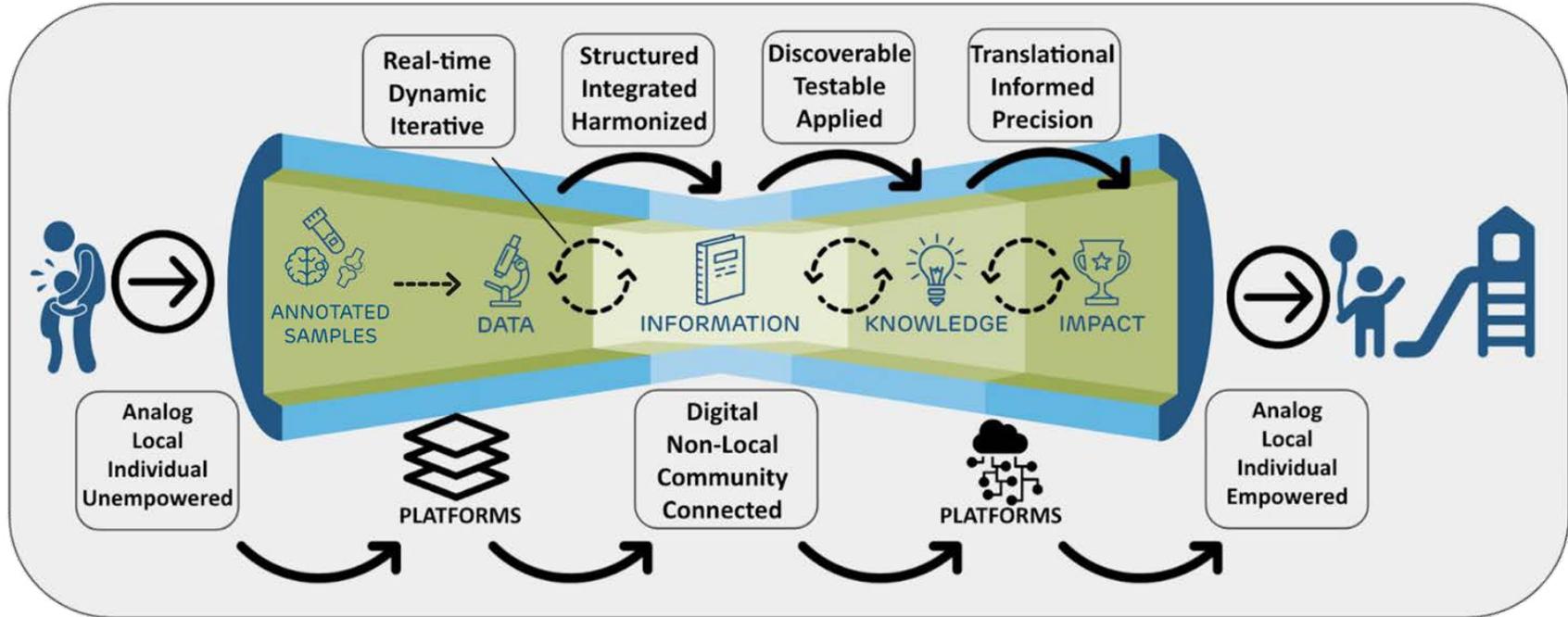
202-279
Data not available



An abstraction for shared approaches for pediatric disaster research

Multi-sourced, Multi-modal, Multi-omic, Multi-disease, Multi-User, Asynchronous Data

Gabriella Miller Kids First Data Resource Center



Platforms as a Service supporting FAIR Ecosystems



Findable



Accessible



Interoperable



Reusable

Rigor through community use:

Source

Metadata

Derived

Workflows/Tools

Big Data for Different Users (and their collaborations)



- Cross-Disease**
- Cancer
 - Birth Defects

- Cross-Modality**
- Clinical
 - WGS
 - RNA
 - Proteomics
 - Imaging



Portal/Knowledge Base



Command Line/Programming Language



Platforms as a Service supporting FAIR Ecosystems



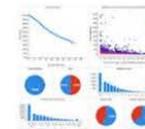
Data Resource Portal

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

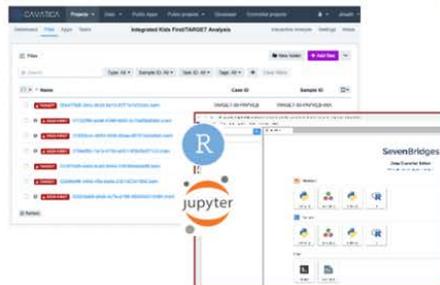


Knowledge Base Integrations (PedcBioPortal)

Integrations with existing curated/published data visualizations

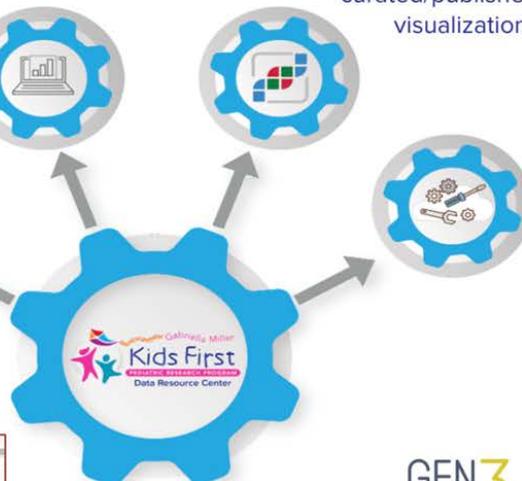


CAVATICA
Pull data from multiple sources into one workspace. Use notebooks, bring-your-own or use available workflows.



Data Services

Model clinical data in FHIR-based data services for semantic interoperability and coordination



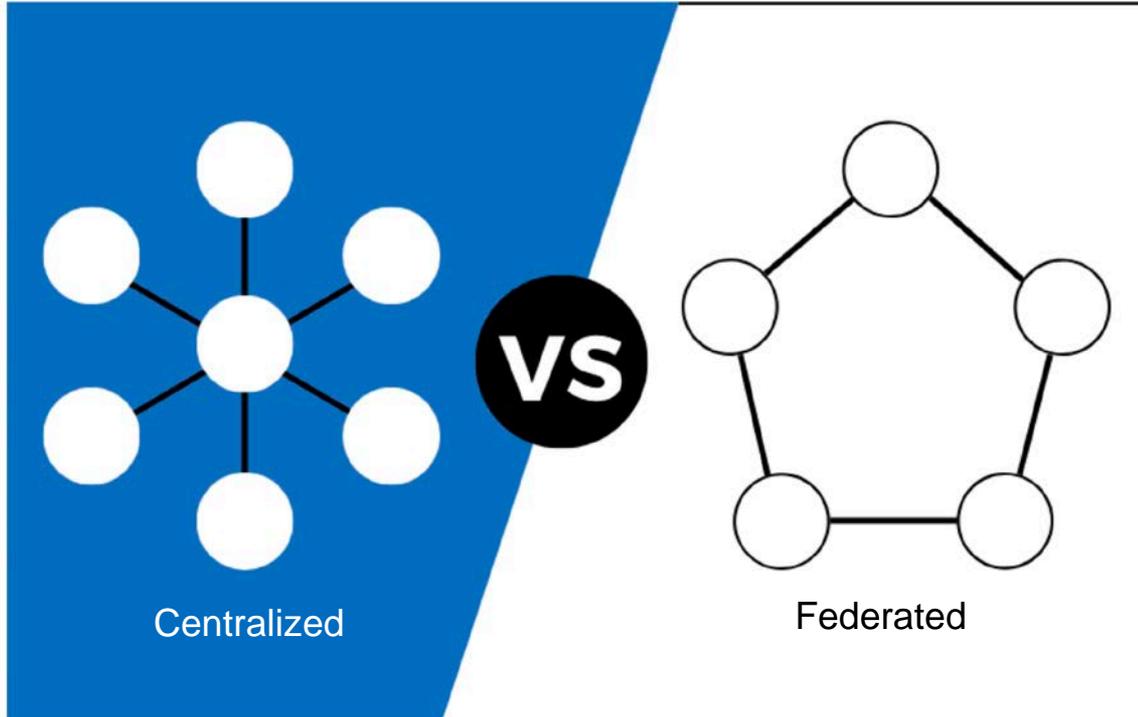
aws
STRIDES

GEN3
DATA COMMONS
Framework Services
Index and point to files in the cloud (for approved users)

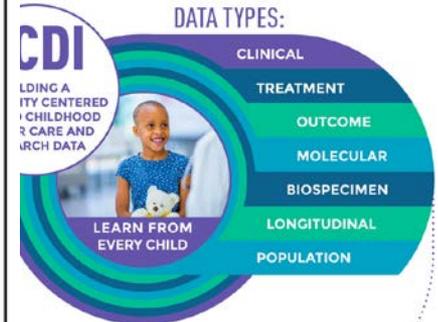
Cross-Platform Interoperability and Governance



Investigation of Co-occurring



NIH Cloud-Based Platform Interoperability (NIH Cloud-Based Platform) Empower end-user across platforms through federation & interoperability



Improving the quality, consistency, and accessibility of data to make it easier for researchers to develop new and better treatments for children with cancer.

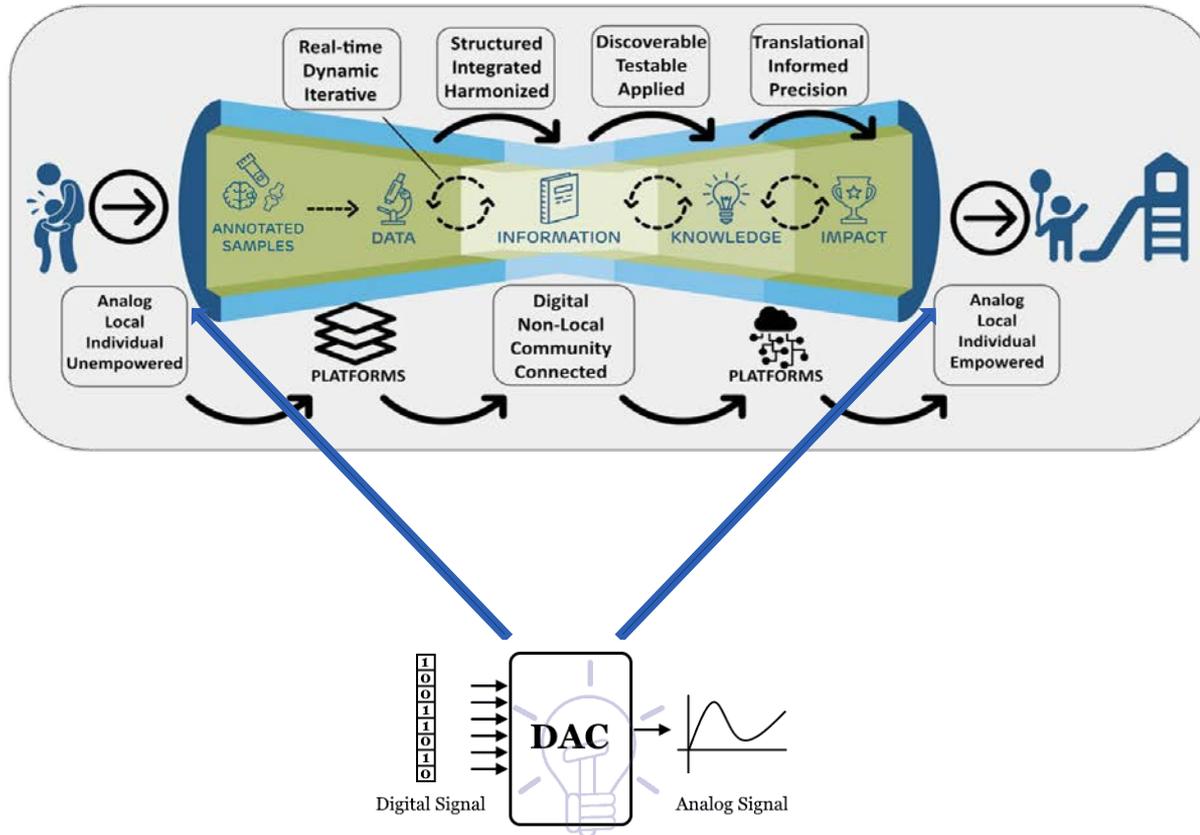




Final Thoughts

Lessons and Opportunities

The Analog to Digital Transition is the hardest part . . .



Intersection across:

1) State/federal data registries and repositories

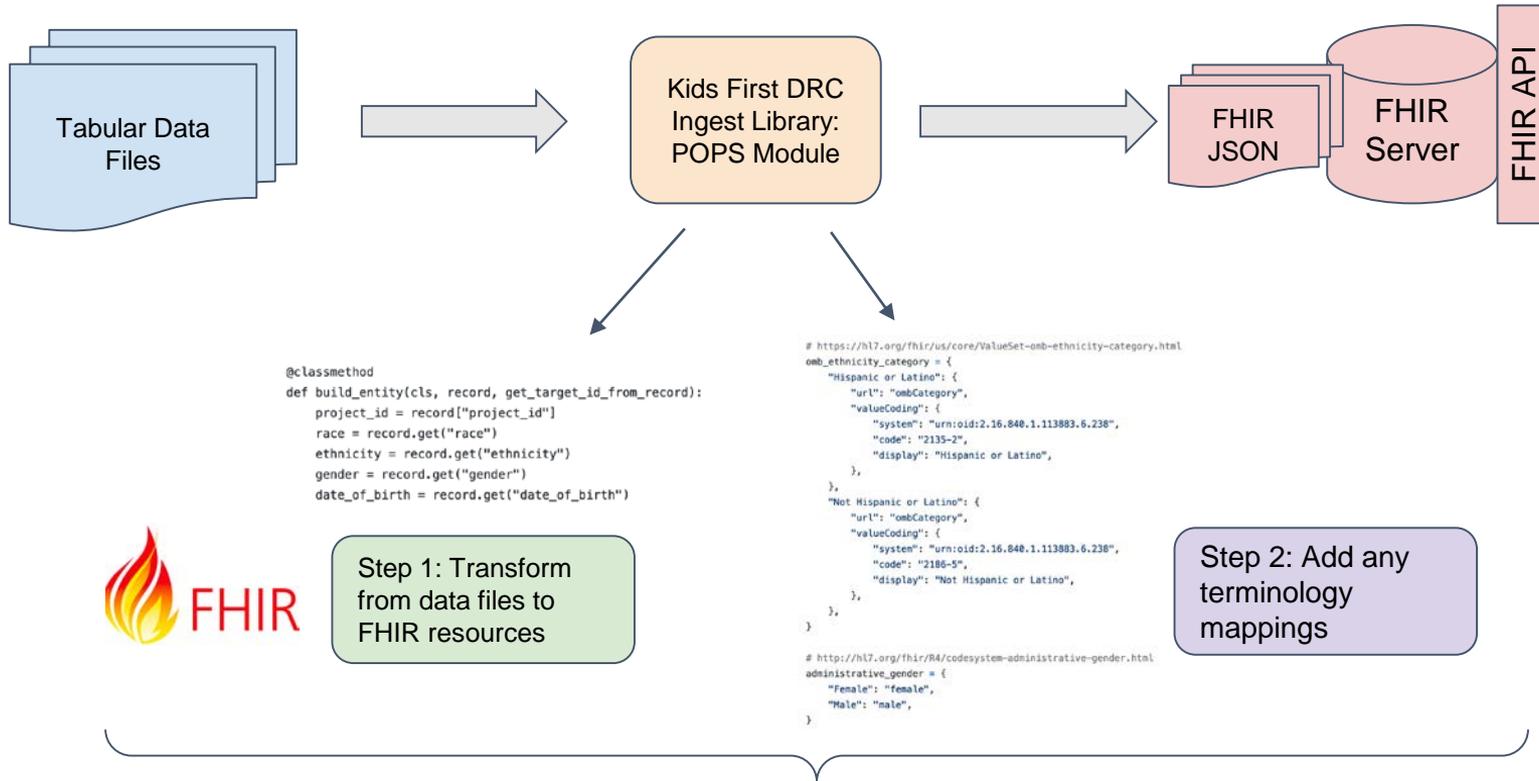
&

2) Healthcare Learning Systems

&

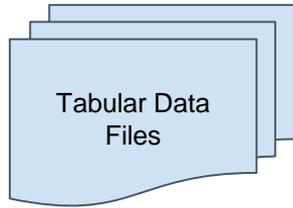
3) "Protocolized" academic research frameworks

Leveraging Community Standards as Paired with Tools



**Code once, then automated as new data arrives:
From data deposition to data “streams”**

Leveraging Community Standards as Paired with Tools



Tabular Data Files

CARING for Children with COVID Resources

FHIR & Data Resources for NIH's *Collaboration to Assess Risk and Identify LoNG-term outcomes for Children with COVID.*

CARING Data FHIR API Endpoint

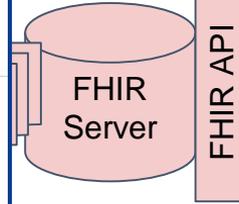
Use this to start querying the entire *CARING* dataset via the FHIR API

Kids First FHIR API Endpoint

Use this to start querying the entire *KF* dataset via the FHIR AP

Kids First FHIR Data Dashboard

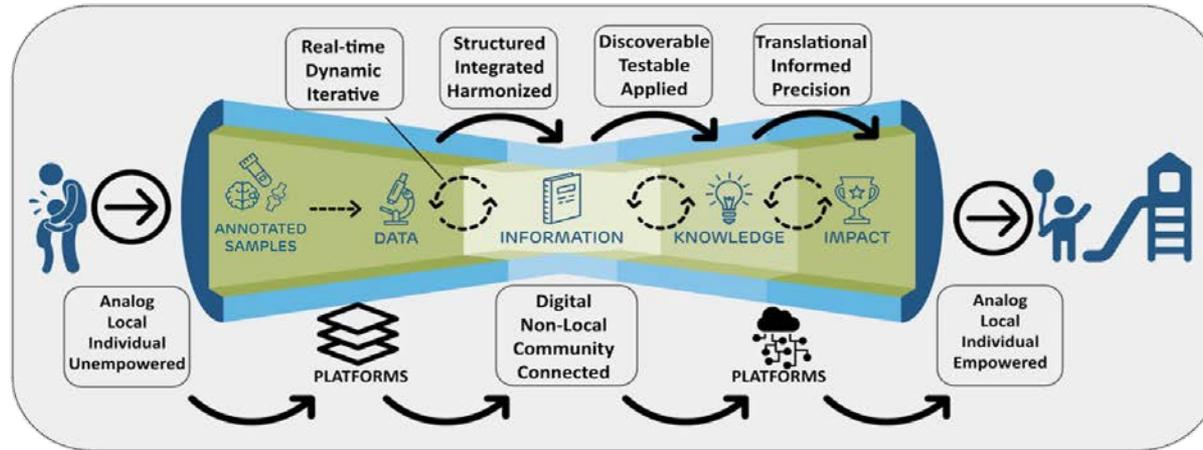
Use this to start exploring *KF* & *CARING* data via a dashboard



Add any
ogy
s



Disaster Research Infrastructure Opportunity



$$a = \frac{\Delta v}{\Delta t}$$

Leverage existing frameworks while **optimizing the ecosystem** around *ACCELERATED ACTIONABILITY* as the variable of impact is a shared need across all translational research, on behalf of all children