cdisc

CDISC ACADEMIC WORKSHOP JAPAN 17 November 2023

JAPAN ACADEMIC WORKSHOP

Friday, 17 November | 1:00pm -5:15pm

D4CG

Transforming Human Health Through Data
The Pediatric Cancer Data Commons

Samuel Volchenboum, MD, PhD

cdisc

A perfect storm
Rare diseases, siloed data,
manual processes, lack of standards,
difficult research questions

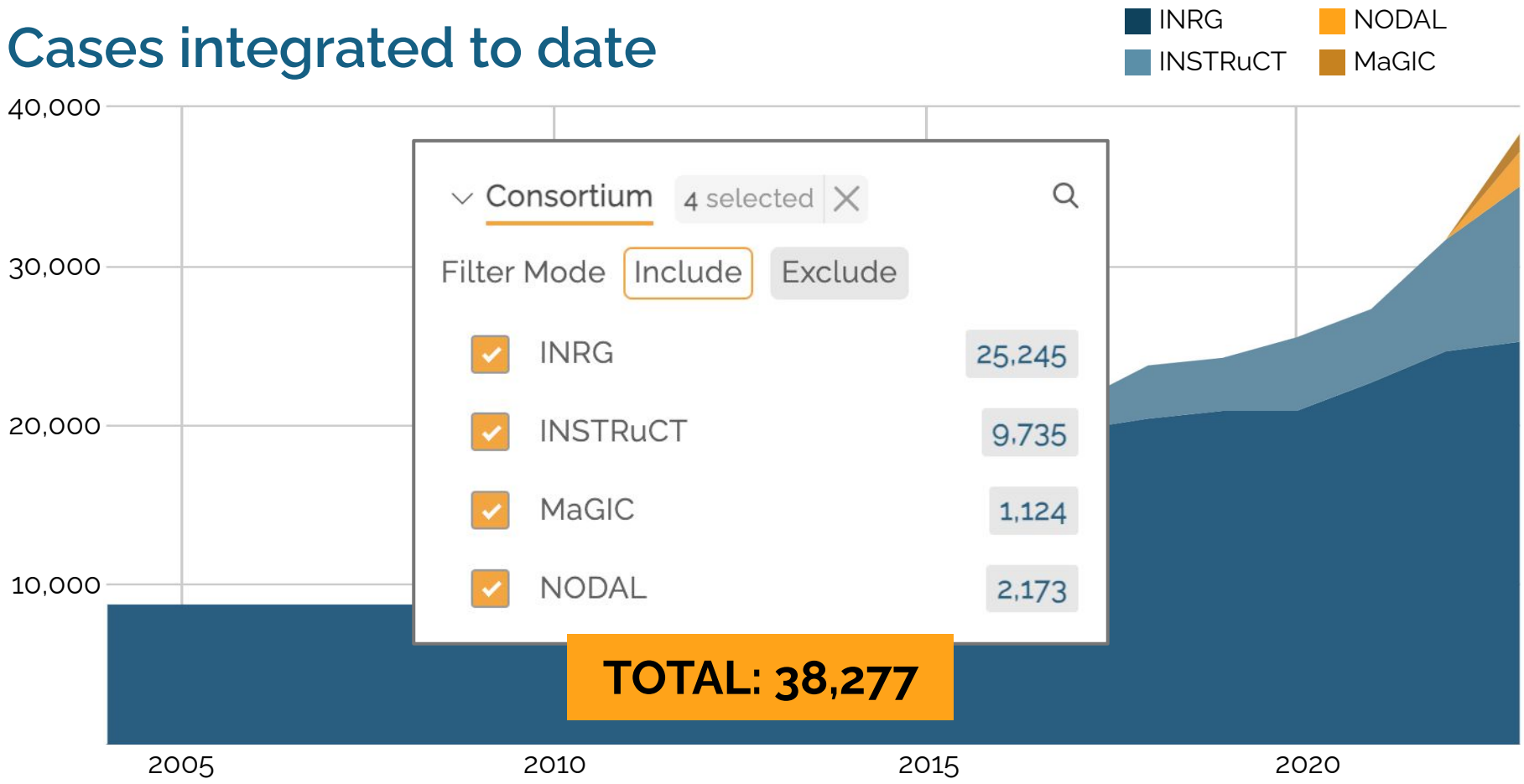
Pediatric Cancer Data Commons (PCDC)

1. Provides a hub for researchers across **most pediatric cancers**
2. Facilitates **cross-disease research**
3. Develops **common core data dictionaries**
4. Creates **common governance structures**
5. Builds on an **active international network** of researchers
6. **Educates, trains, and increases data literacy** of the community
7. Sources data from **clinical trials**, **registries**, and (soon) **EHRs**

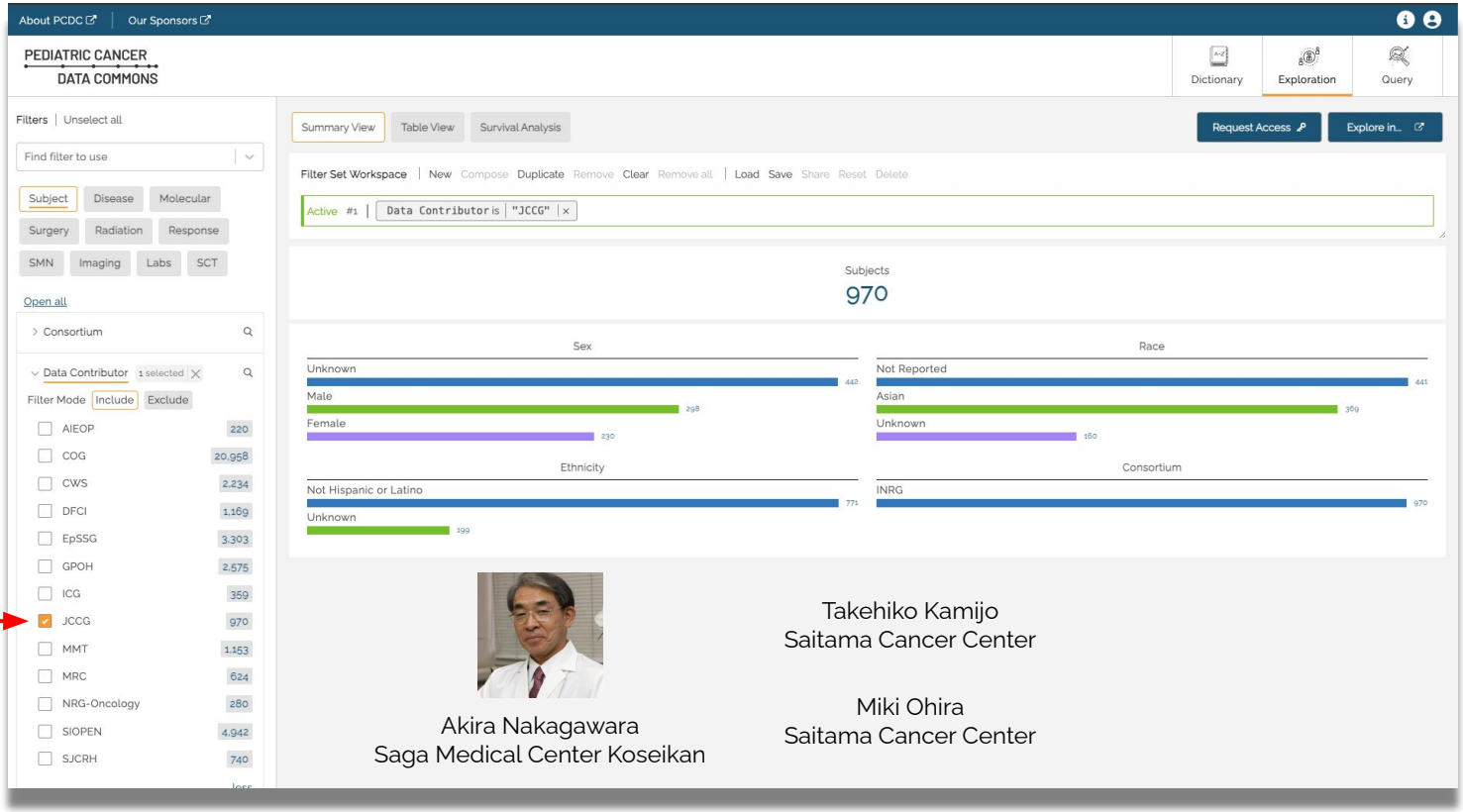
DATA FOR THE COMMON GOOD

We build communities, platforms, and ecosystems
that maximize the potential of data
to drive discovery and improve human health.

Cases integrated to date



Japan data in the PCDC



Akira Nakagawara
Saga Medical Center Koseikan

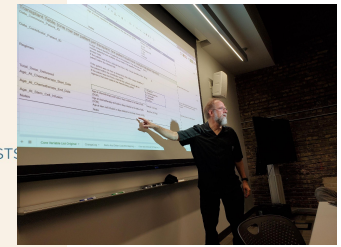
Takehiko Kamijo
Saitama Cancer Center

Miki Ohira
Saitama Cancer Center

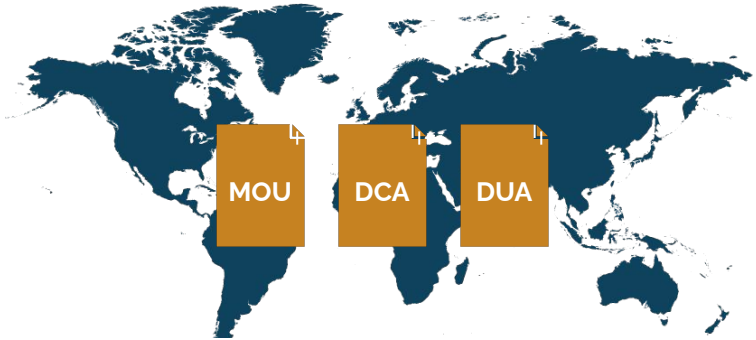
Building a data commons—the D4CG way

Step 1: Establish a **consortium**

- PCDC Master Data Dictionary
- Acute Lymphoblastic Leukemia (ALL)
- Acute Myeloid Leukemia (AML)
- Central Nervous System Tumors (CNS)
- Ewing Sarcoma (EWS)
- Germ Cell Tumors (GCT)
- Hodgkin Lymphoma (HL)
- Neuroblastoma (NBL)
- Non-Rhabdomyosarcoma Soft Tissue Sarcoma (NRST
- Osteosarcoma (OS)
- Predisposition
- Retinoblastoma (RB)
- Rhabdomyosarcoma (RMS)

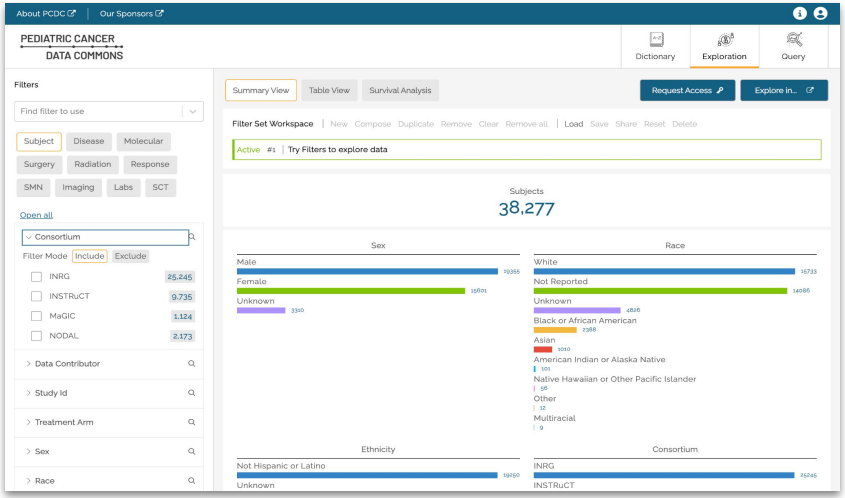


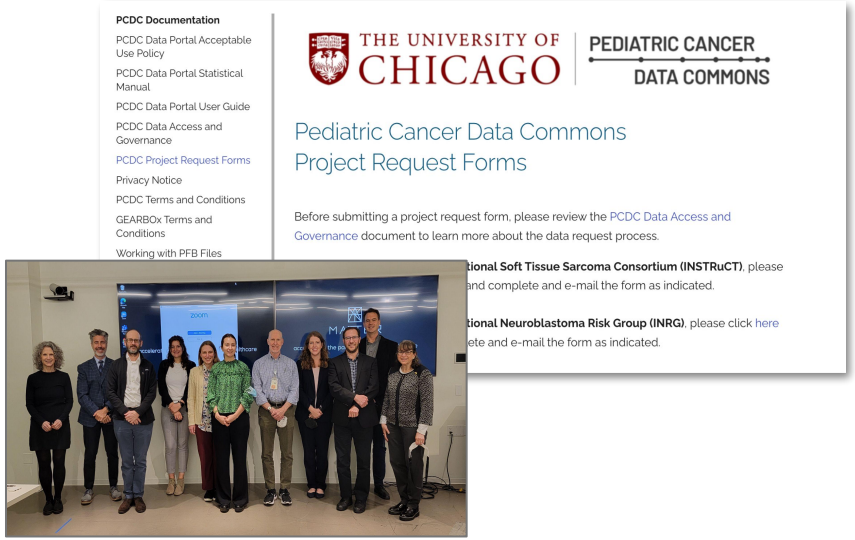
Step 2: Develop and deploy **data operations**



Step 3: Establish and implement **governance**

Building a data commons—the D4CG way





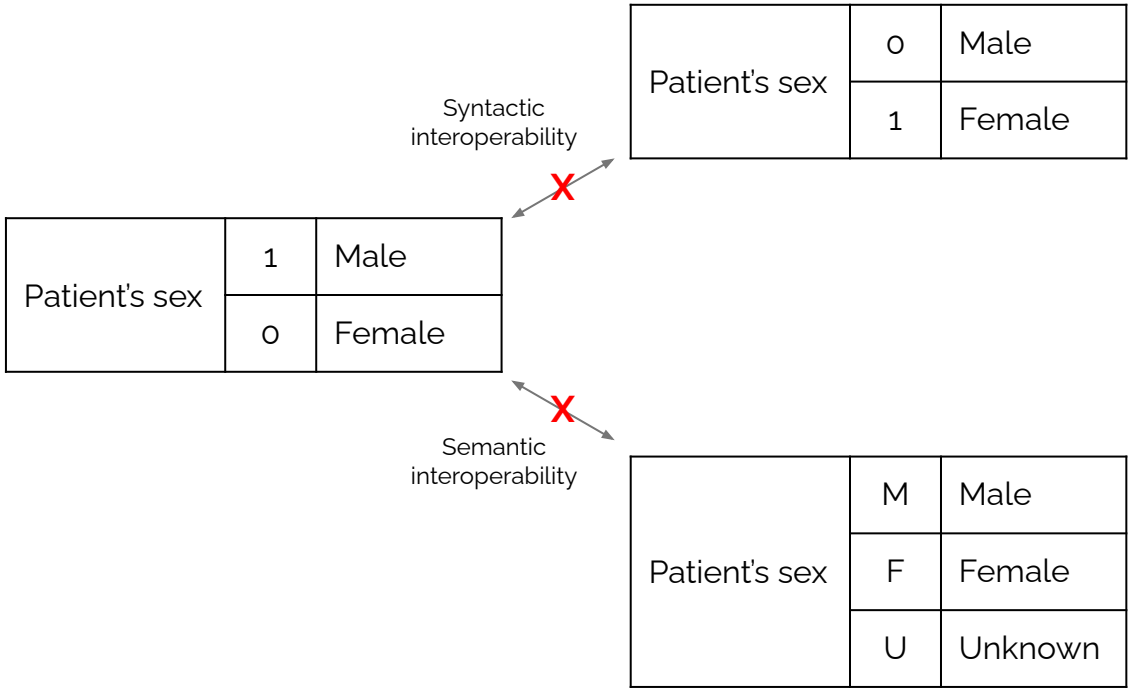
Step 4: Develop and deploy the **technical infrastructure**

Step 5: Socialize and sustain the commons

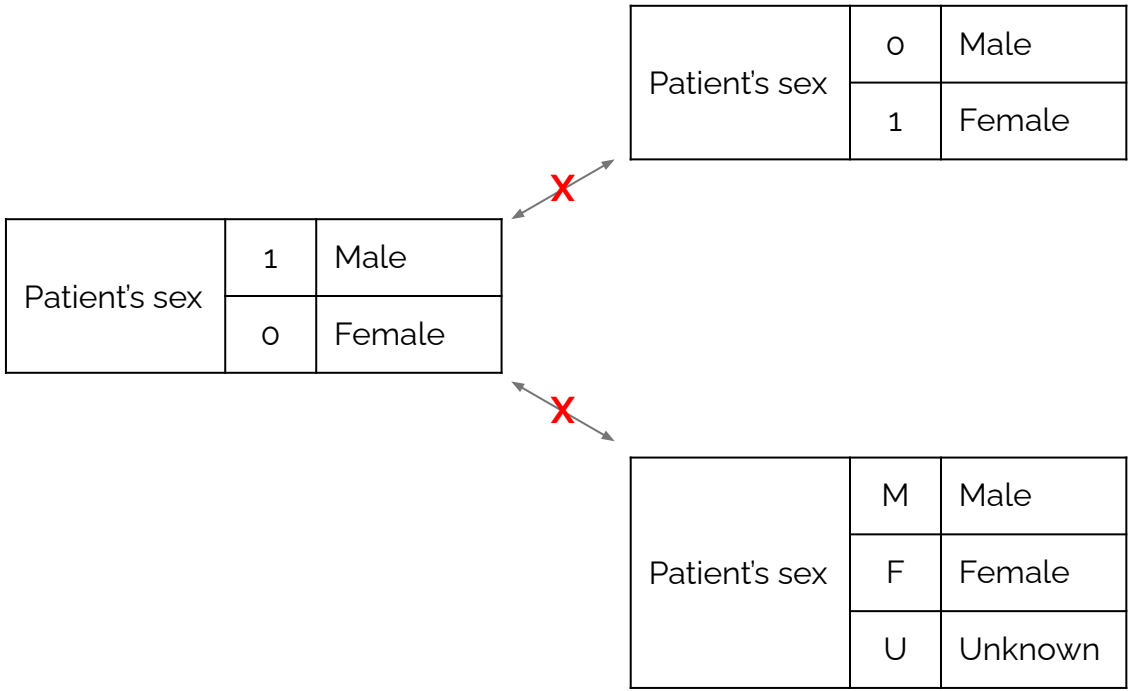
Engage stakeholders

We employ a “big tent” philosophy.

The importance of data standards



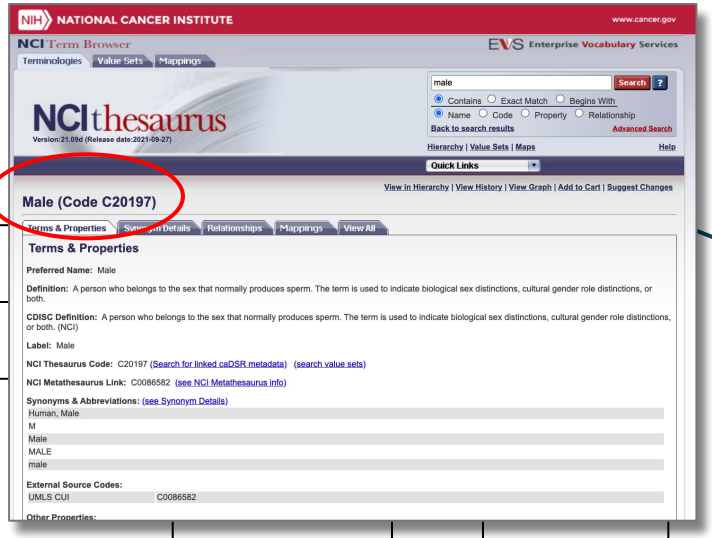
The importance of data standards



Sex	
Description	**NCIt Code**
Male	C20197
Female	C16576
Unknown	C17998

The importance of data standards

Patient's sex	1	Male
	0	Female



Patient's sex	F	Female
	U	Unknown

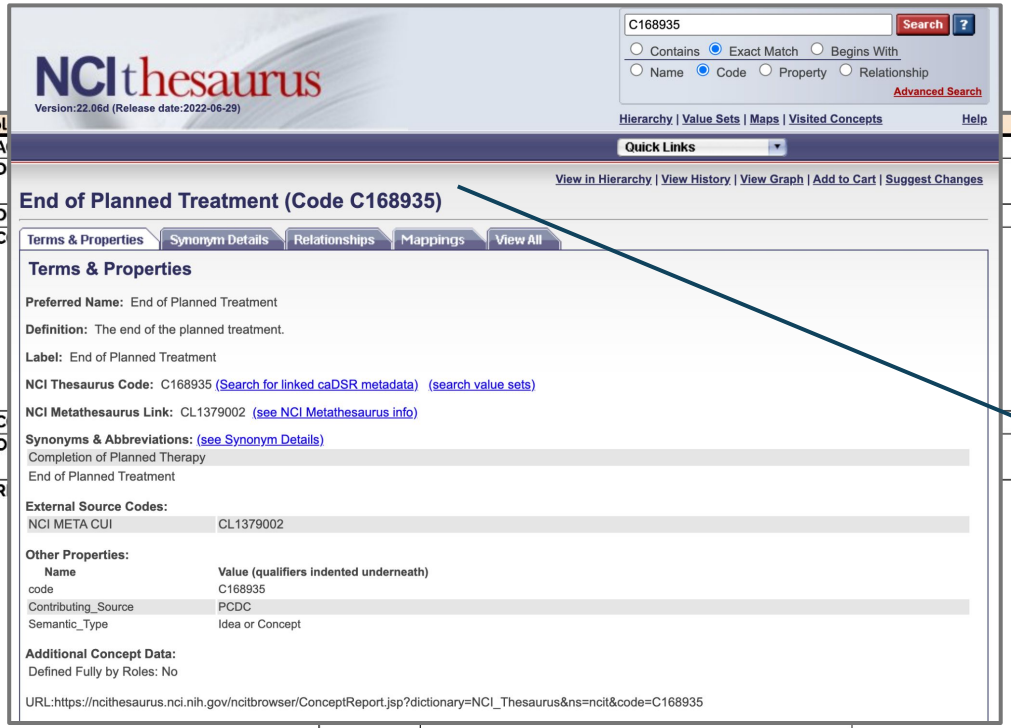
Sex	
Description	**NCIt Code**
Male	C20197
Female	C16576
Unknown	C17998

Building a consensus data dictionary

Off Protocol Therapy/Study: one row per subject per off protocol therapy/study per reason off							
	AGE_OFF	Number	Age in Days When Off Protocol Therapy or Study	C172678	Age of subject (in days) wh		
	DISEASE_PHASE	Code	Disease Phase	C168878	The stage or period of an in	Initial Diagnosis	C156813
						Relapse	C38155
	DISEASE_PHASE_NUMBER	Number	Disease Phase Number	C173258	The number of the disease		
	COURSE	Code	Protocol Treatment Course	C168807	The type of protocol treatm	Prephase	C168826
						Induction	C158876
						Intensification	C173105
						Consolidation	C15679
						Stem Cell Transplant Conditioning	C168794
						Maintenance	C15688
						Palliative Treatment	C15292
						Other	C17649
	COURSE_NUMBER	Number	Course Number	C166235	The number assigned to a		
	OFF_TYPE	Code	Off Protocol Therapy or Study	C173256	The code used to designat	Protocol Therapy	C173257
						Study	C29851
	REASON_OFF	Code	Off Protocol Therapy or Study Reason	C173519	The reason a subject went	Death	C93546
						Lost to Follow-Up	C70740
						Completion of Planned Therapy	C168935
						Physician Decision	C48250
						Withdrawal of Consent	C48271
						Subject/Guardian Refused Further Treatment	C168934
						Disease Progression	C35571
						Relapse	C38155
						Adverse Event	C41331
						Secondary Malignancy	C4968
						Other	C17649
						Unknown	C17998
						Not Reported	C43234

http://sam.am/datadictionaries

Building a consensus data dictionary



Off Protocol		
A	Age of subject (in days) wh	
D	The stage or period of an in	Initial Diagnosis C156813
		Relapse C38155
D	The number of the disease	
C	The type of protocol treatm	Prephase C168826
		Induction C158876
		Intensification C173105
		Consolidation C15679
		Stem Cell Transplant Conditioning C168794
		Maintenance C15688
		Palliative Treatment C15292
		Other C17649
C	The number assigned to a	
O	The code used to designat	Protocol Therapy C173257
		Study C29851
R	The reason a subject went	Death C93546
		Lost to Follow-Up C70740
		Completion of Planned Therapy C168935
		Physician Decision C48250
		Withdrawal of Consent C48271
		Subject/Guardian Refused Further Treatment C168934
		Disease Progression C35571
		Relapse C38155
		Adverse Event C41331
		Secondary Malignancy C4968
		Other C17649
		Unknown C17998
		Not Reported C43234

http://sam.am/datadictionaries

Data collection is highly localized

☐ Male ☐ Female ☐ Unknown *English*	☐ 男 ☐ 女性 ☐ わからない *Japanese*
☐ ذكر ☐ أنثى ☐ مجهول *Arabic*	☐ Homme ☐ Femelle ☐ Inconnue *French*
☐ Männlich ☐ Weiblich ☐ Unbekannt *German*	☐ 男性 ☐ 女性 ☐ 未知 *Chinese*

CDE mapping solves the localization problem

☐ Male
☐ Female
☐ Unknown
English

☐ 男
☐ 女性
☐ わからない
Japanese

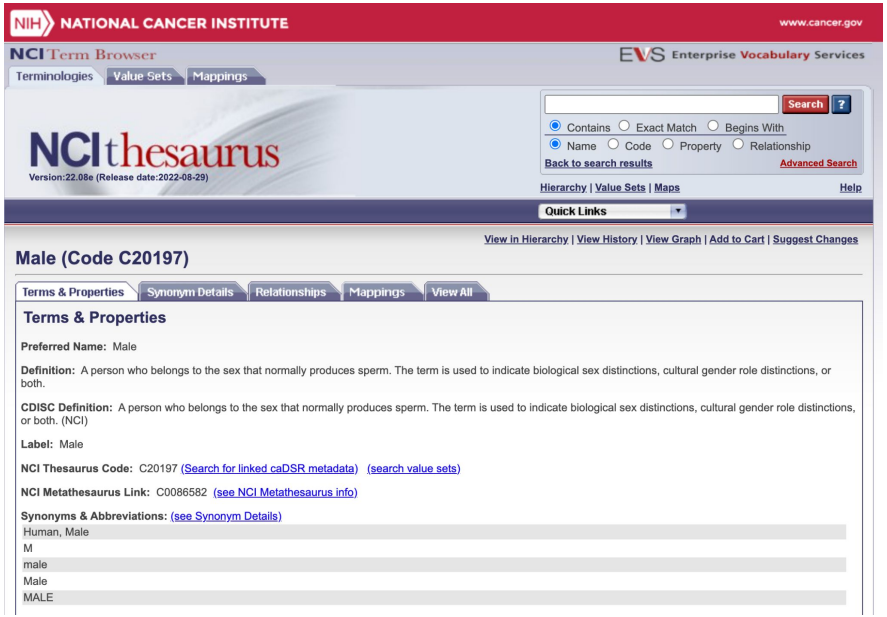
☐ ذكر
☐ أنثى
☐ مجهول
Arabic

☐ Homme
☐ Femelle
☐ Inconnue
French

☐ Männlich
☐ Weiblich
☐ Unbekannt
German

☐ 男性
☐ 女性
☐ 未知

C20197
C16576
C17998



PCDC data dictionaries

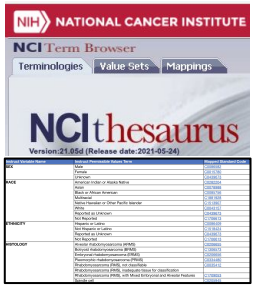
- PCDC Master Data Dictionary
- Acute Lymphoblastic Leukemia (ALL)
- Acute Myeloid Leukemia (AML)
- Central Nervous System Tumors (CNS)
- Ewing Sarcoma (EWS)
- Germ Cell Tumors (GCT)
- Hodgkin Lymphoma (HL)
- Neuroblastoma (NBL)
- Non-Rhabdomyosarcoma Soft Tissue Sarcoma (NRSTS)
- Osteosarcoma (OS)
- Retinoblastoma (RB)
- Rhabdomyosarcoma (RMS)

https://commons.cri.uchicago.edu/data-dictionaries/

Getting data into the PCDC

Balloted Dictionary → **Tiering** → **NCIt Coding** → **Harmonization** → **Quality Control** → **Ingestion into PCDC**

Tier 1 - Required
Tier 2 - Prioritize Inclusion
Tier 3 - Don't Prioritize Inclusion



STAKEHOLDERS ENGAGED
DATA DICTIONARY ESTABLISHED
DATA CONTRIBUTORS COMMITTED
CONSORTIUM MOU SIGNED
CASES IN COMMONS
ANALYSES IN PROGRESS
PAPERS PUBLISHED
acute lymphoblastic leukemia
2
acute myeloid leukemia
V1
4
INTERACT
bone tumors
V2
10
HIBiSCus
central nervous system tumors
V1
6
INSPiRE
Fanconi anemia
germ cell tumors
V1
8
MaGIC
1.1K
10
37
Hodgkin lymphoma
V1
2
NODAL
2.2K
nasopharyngeal carcinoma
NOBLE
neuroblastoma
V2
5
INRG
25K
8
32
oncofertility
R-HOPE
predisposition
V1
7
C3P
retinoblastoma
4
Global REACH
soft-tissue sarcoma
V2
5
INSTRuCT
9.8K
17
12

PCDC worldwide participation

Worldwide Data Use Agreements

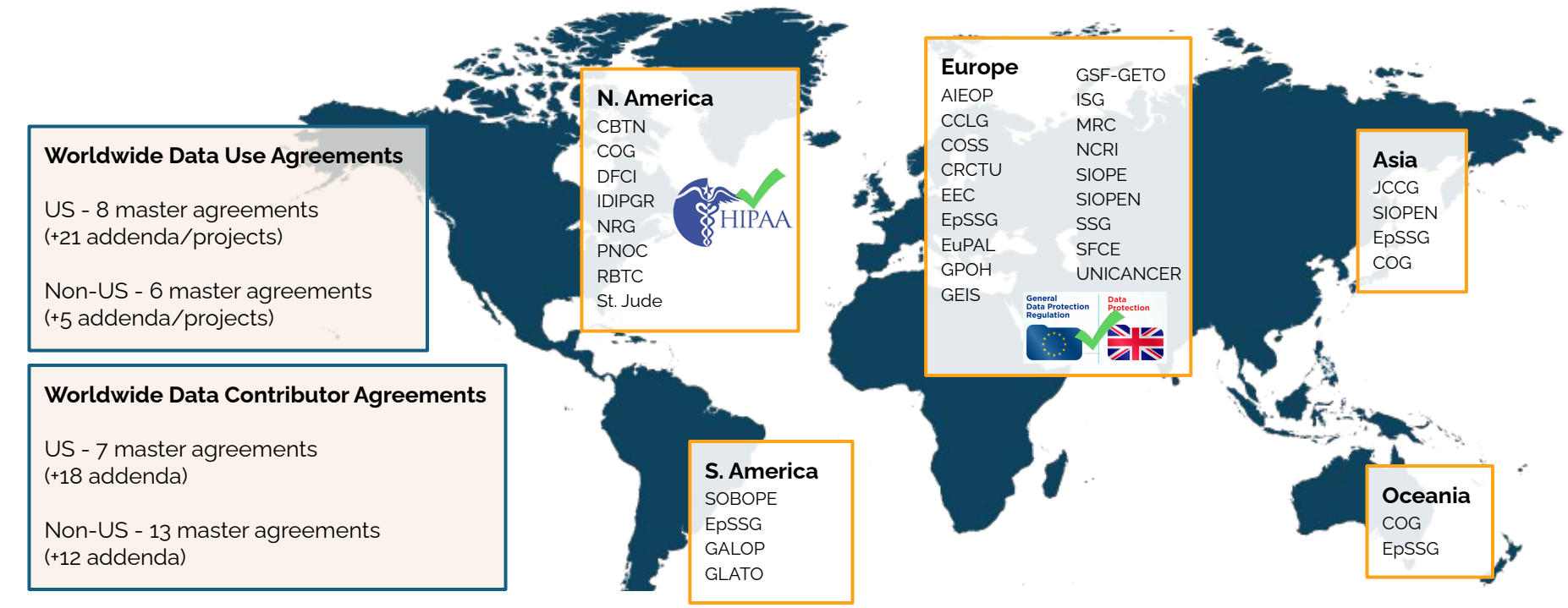
US - 8 master agreements
(+21 addenda/projects)

Non-US - 6 master agreements
(+5 addenda/projects)

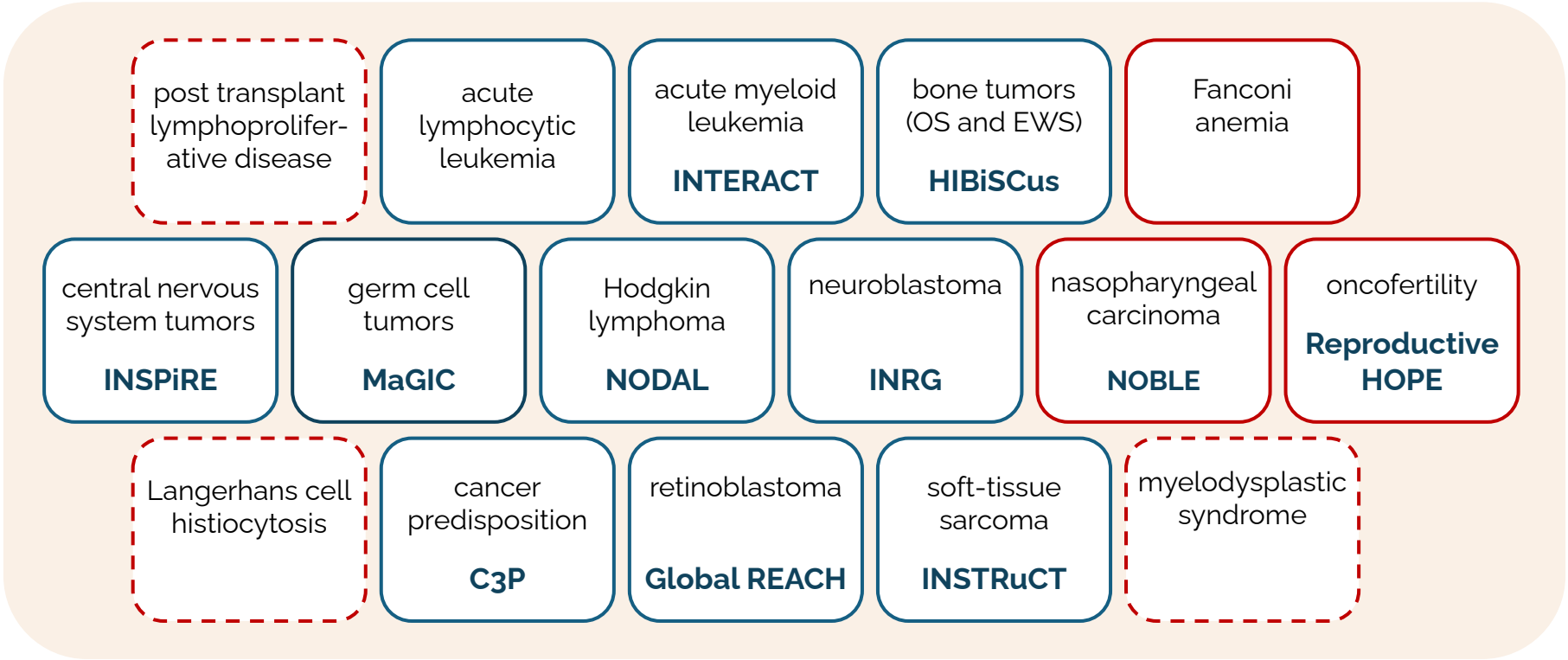
Worldwide Data Contributor Agreements

US - 7 master agreements
(+18 addenda)

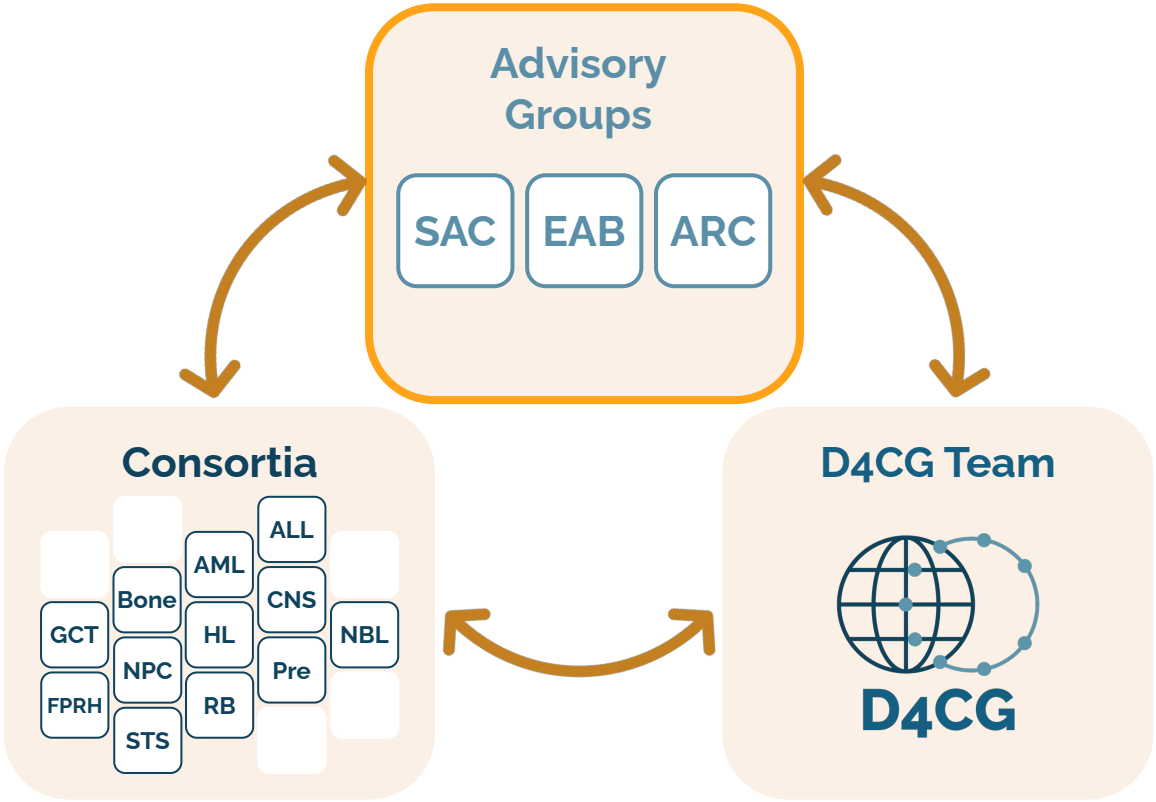
Non-US - 13 master agreements
(+12 addenda)



New cancer-related consortia



PCDC advisory groups structure



PCDC advisory groups structure

Daisuke Tomizawa
Head, Division of Leukemia and Lymphoma
National Center for Child Health and Development
Tokyo, Japan

co-chair

Advisory Groups

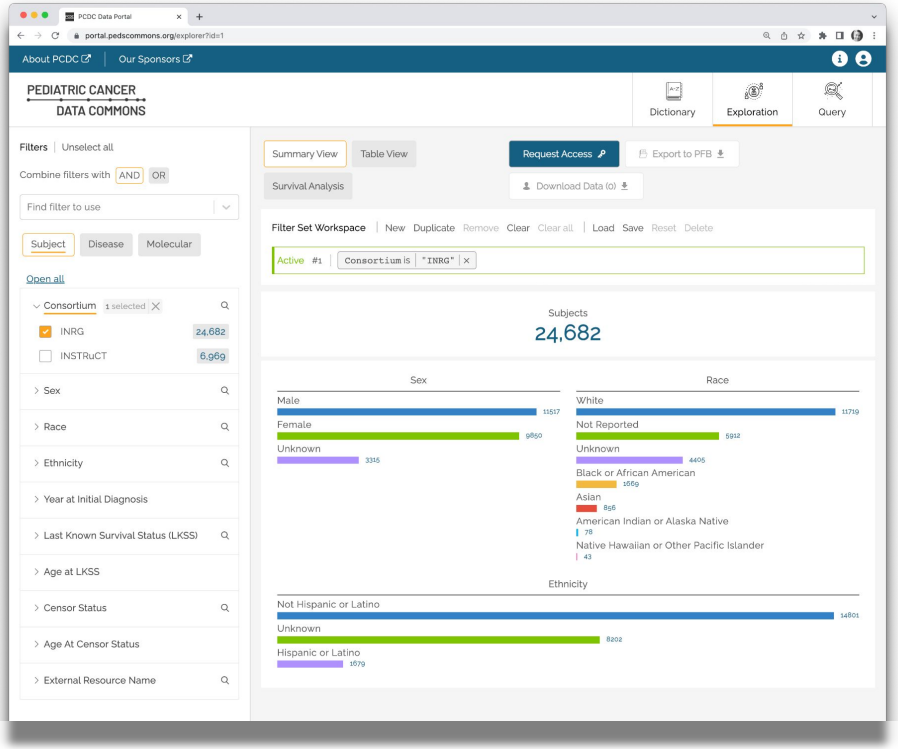
SAC EAB ARC

Consortia

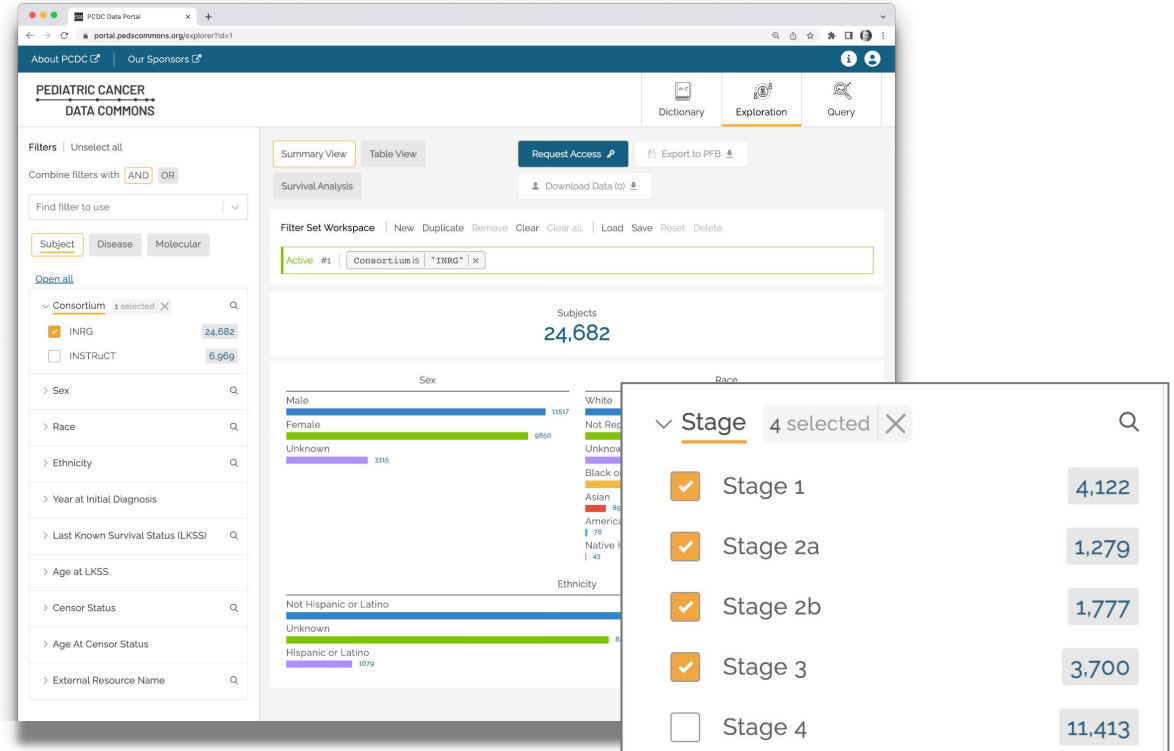
ALL AML CNS Bone HL Pre GCT NBL NPC RB FPRH STS

D4CG Team

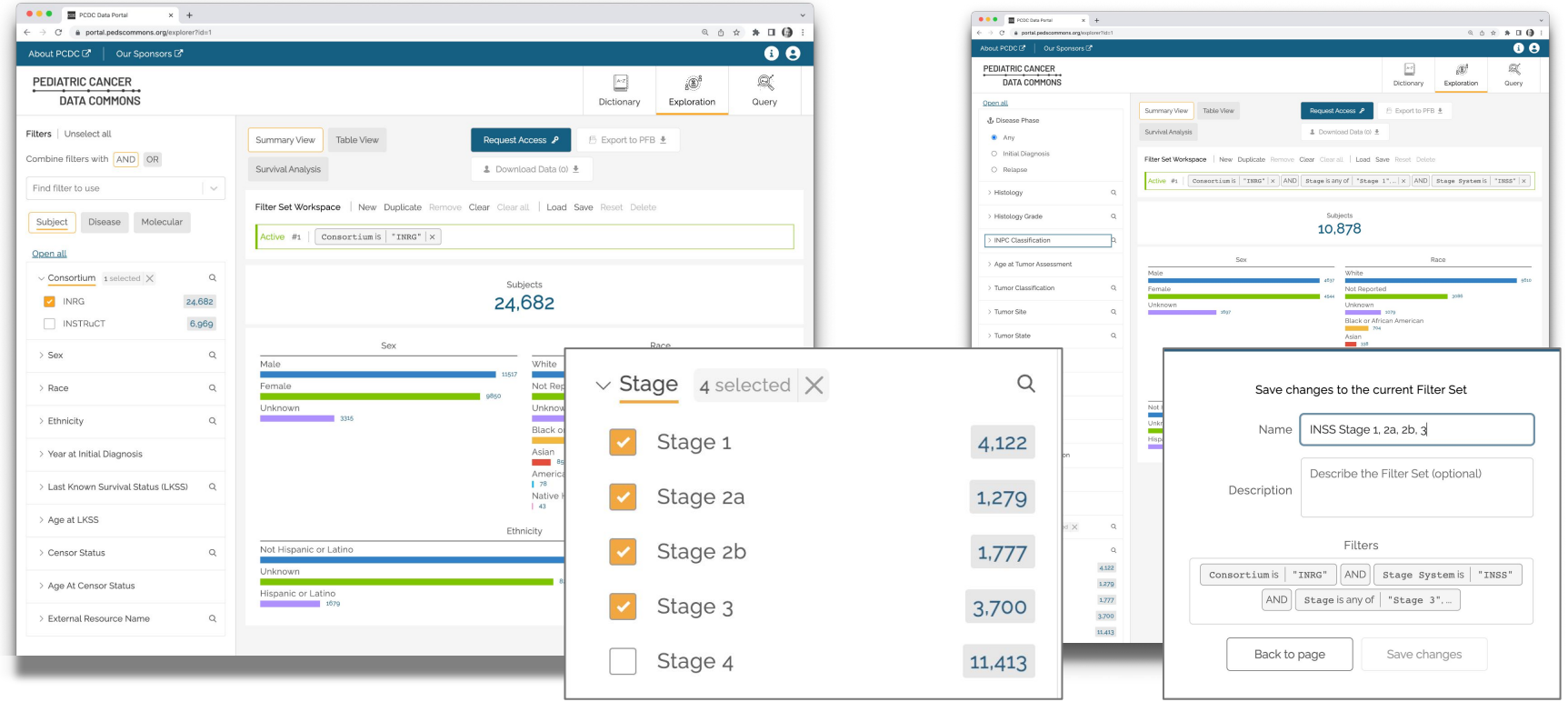
D4CG



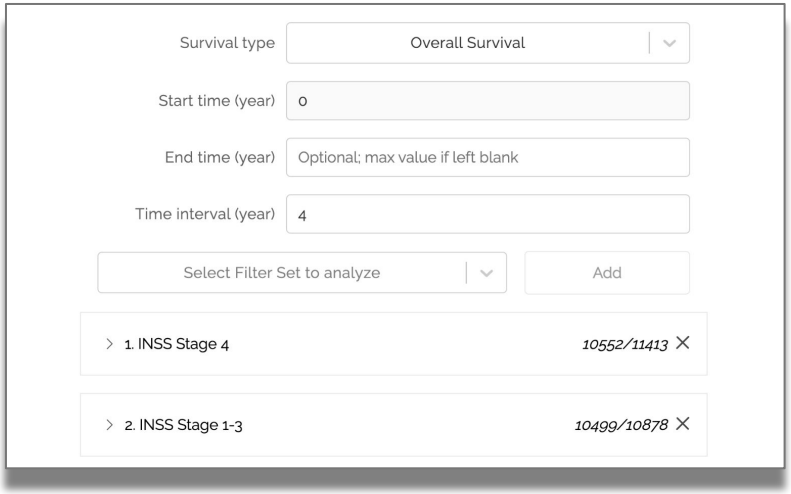
http://portal.pedscommons.org

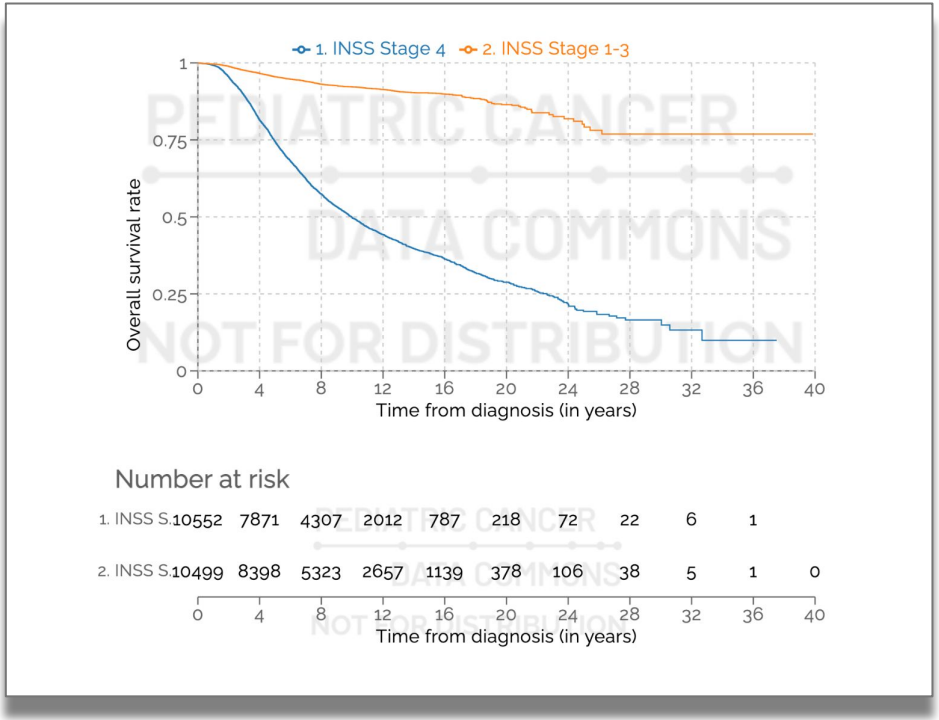


http://portal.pedscommons.org



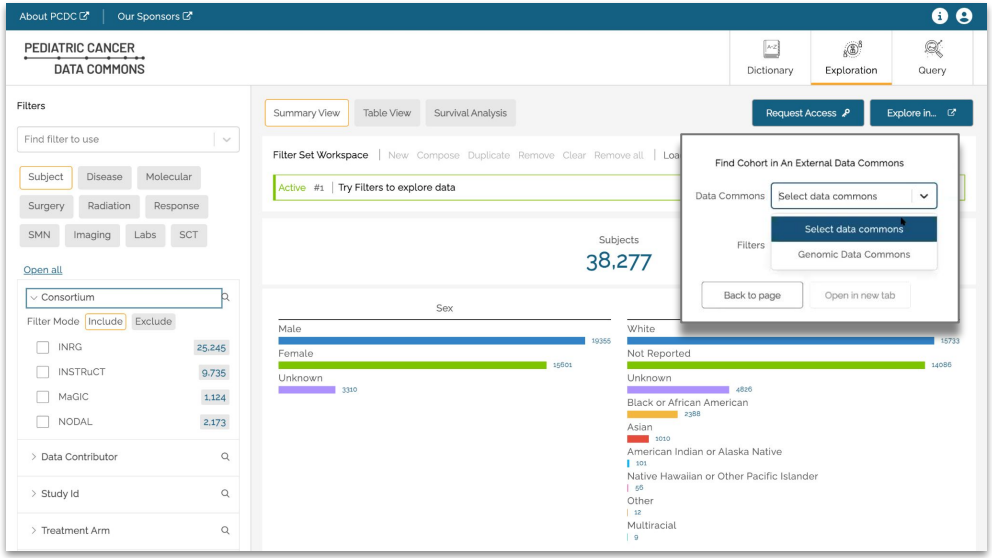
http://portal.pedscommons.org

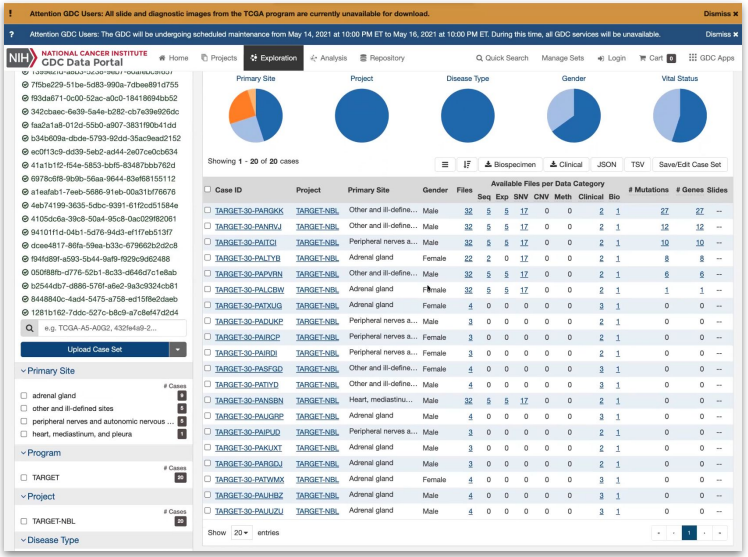




http://portal.pedscommons.org

Cohort discovery - linking to Genomic Data Commons





Direct connection to other data commons
(Genomic Data Commons, Gabriella Miller Kids First)

http://portal.pedscommons.org

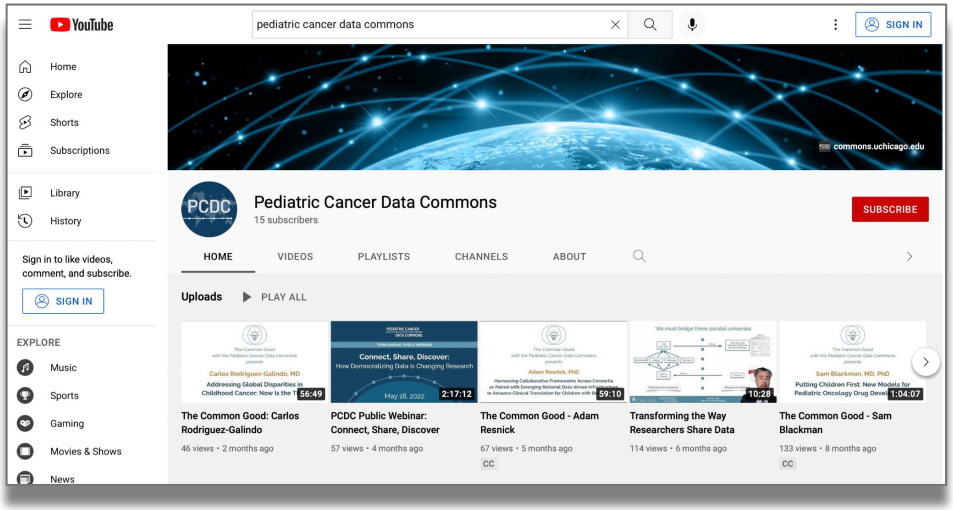
Engagement / Education

Education

- Promoting data literacy through training
- Young / early stage investigator seminars

Engage the community

- Webinars
- Thought leadership
- Live events
- Patient and family involvement
- International advisory board



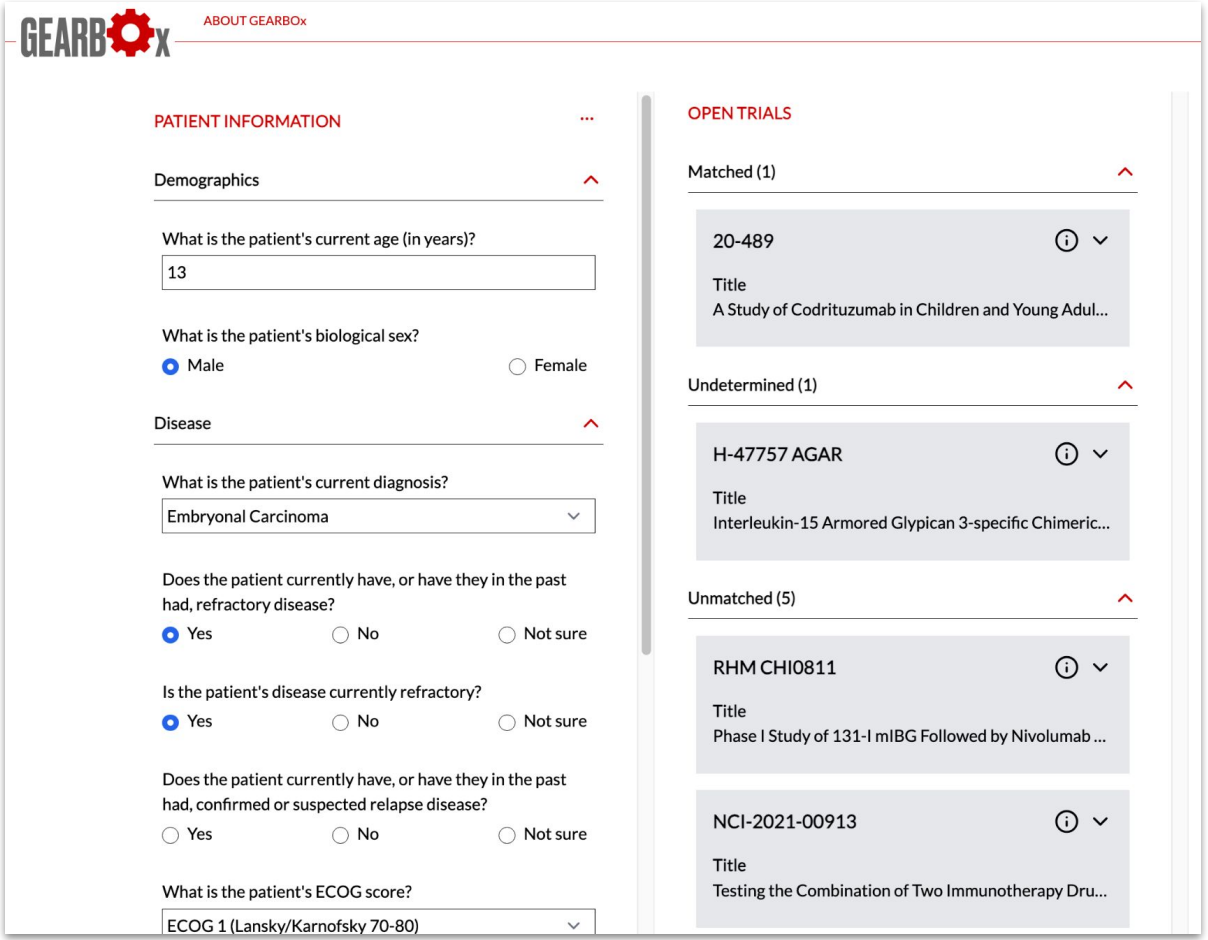
sam.am/pcdcresearch

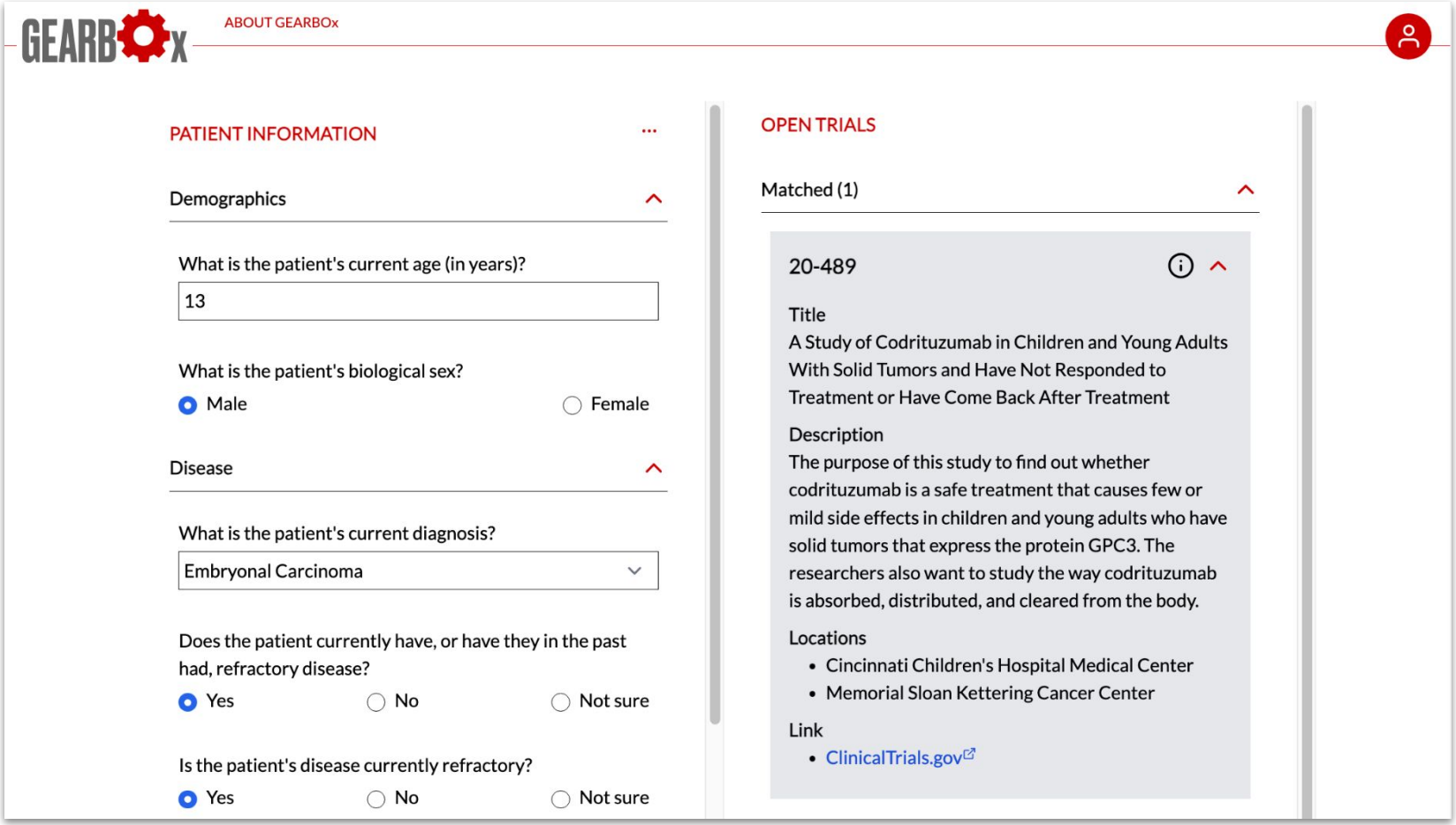
INRG Research

Click on a description to view the original project proposal.

INRG #	Principal Investigator	Description	Project Type	Status	Publication	Presentation
2022-04	Mallory Taylor Thomas Cash Wendy London Julie Park Meredith Irwin	Outcomes for patients aged 12-18 months with Stage M MYCN non-amplified neuroblastoma and unfavorable biologic features ('Mixed Phenotype Toddlers')	Investigator	In Progress		Presented at ANR 2(
2022-03	Hanxaio Yu Xingda Zhan Mark Appelbaum Gudrun Schleiermacher statistician to be named	Prognostic impact of segmental chromosome alterations in high-risk neuroblastoma patients on immunotherapy: A report from the International Neuroblastoma Risk Group (INRG) project	Investigator	In Progress		
2022-02	Boris Decarolis Wendy London Susan Cohn Andrew Pearson	Survival of patients with low-, intermediate-, or high-risk neuroblastoma over a 35 year period	Investigator	In Progress		
2022-01	Wendy London Ramya Ramanujachar Kavitha Srivatsa Paola Angelini	Neuroblastoma in adolescents and adults- a study of clinical and biological features and outcomes	Investigator	In Progress		
2021-01	Kevin Campbell Pei-Chi Kao Arlene Naranjo Takehiko Kamijo Ramya Ramanujachar Wendy London Steven DuBois	Clinical and Biological Features Predictive of Survival After Relapse of Stage MS Neuroblastoma: A Report From the International Neuroblastoma Risk Group Project	Investigator	Published	Clinical and biological features prognostic of survival after relapse or progression of INRGSS stage MS pattern neuroblastoma: A report from the International Neuroblastoma Risk Group (INRG) project. *Pediatr Blood Cancer.* 2023 Feb;70(2):e30054. Epub 2022 Oct 31. doi: 10.1002/pbc.30054.	
2020-03	Riyue Bao Stefani Spranger Kyle Hernandez Yuanyuan Zha Peter Pytel Jason Luke Thomas Gajewski Samuel Volchenboum Susan Cohn	Validation of a T-cell inflammatory signature and outcomes in patients with neuroblastoma	Investigator	Published	Immunogenomic determinants of tumor microenvironment correlate with superior survival in high-risk neuroblastoma. *J Immunother Cancer.* 2021 Jul;9(7):e002417. doi: 10.1136/jitc-2021-002417. PMID: 34272305; PMCID: PMC8287618.	

GEARBOx
Clinical trials matching

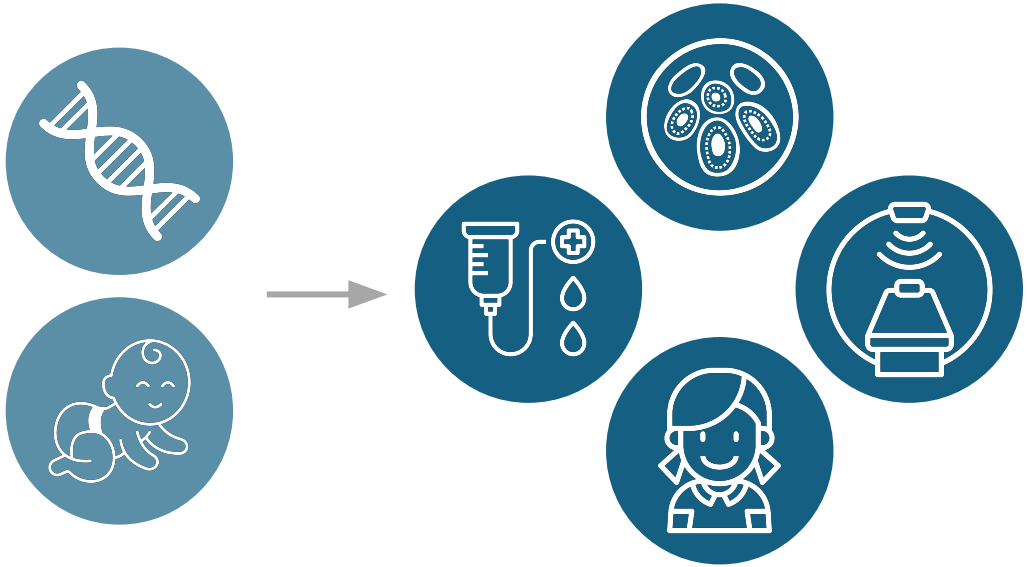



GEARBOx
ABOUT GEARBOx
PATIENT INFORMATION
Demographics
What is the patient's current age (in years)?
13
What is the patient's biological sex?
Male
Female
Disease
What is the patient's current diagnosis?
Embryonal Carcinoma
Does the patient currently have, or have they in the past had, refractory disease?
Yes
No
Not sure
Is the patient's disease currently refractory?
Yes
No
Not sure
OPEN TRIALS
Matched (1)
20-489
Title
A Study of Codrituzumab in Children and Young Adults With Solid Tumors and Have Not Responded to Treatment or Have Come Back After Treatment
Description
The purpose of this study to find out whether codrituzumab is a safe treatment that causes few or mild side effects in children and young adults who have solid tumors that express the protein GPC3. The researchers also want to study the way codrituzumab is absorbed, distributed, and cleared from the body.
Locations
Cincinnati Children's Hospital Medical Center
Memorial Sloan Kettering Cancer Center
Link
ClinicalTrials.gov

Connecting the dots across the patient's lifetime

pediatric, AYA, or
adult cancer

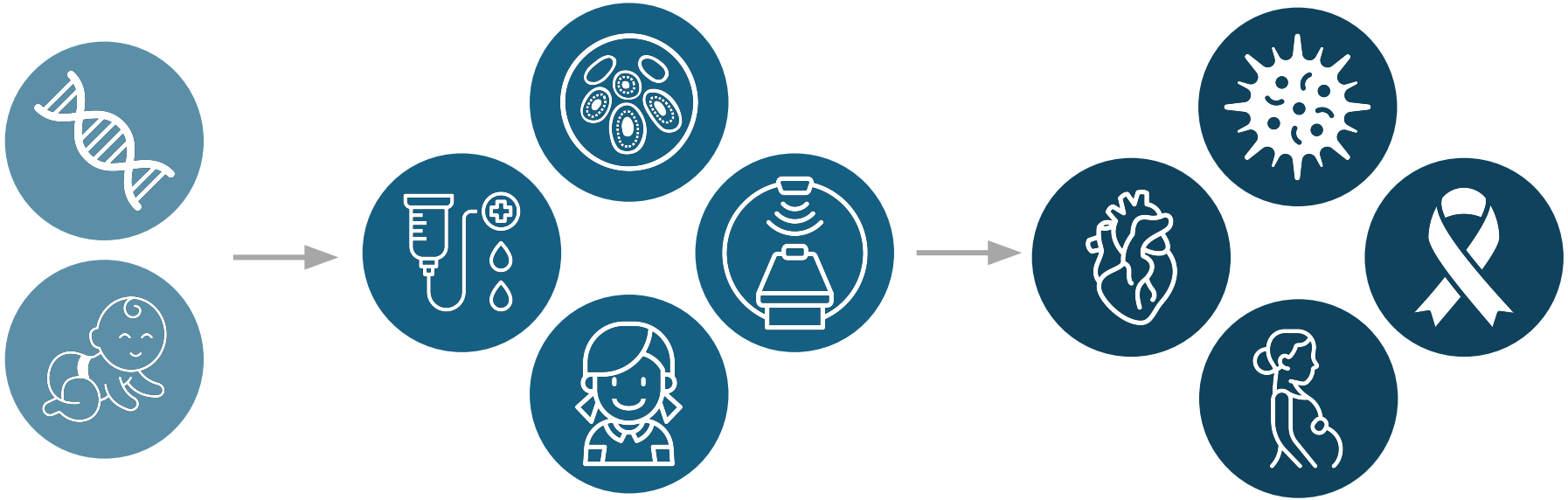
Connecting the dots across the patient's lifetime



cancer predisposition syndrome

pediatric, AYA, or adult cancer

Connecting the dots across the patient's lifetime



cancer predisposition syndrome

pediatric, AYA, or adult cancer

secondary malignancy, late effects

Why use the D4CG model for other rare diseases?

Lack of data hinders research for rare diseases.

Lack of community consensus and **deficient data standards** make data aggregation difficult or impossible.

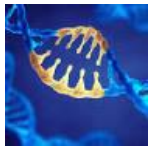
Researchers cannot **make progress and improve patient outcomes** without high-quality data.

We envision a world where **access to high-quality data** is never a barrier to **improving human health.**

food allergies

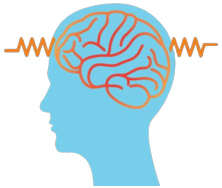
PEDIATRIC CANCER DATA COMMONS

pediatric cancer



monogenic diabetes

Crohn's disease



epilepsy

Sociome

other rare diseases

...and more

A model for transforming human health

- Studying rare diseases requires **collaboration and sharing**
- Data sharing must be built on a foundation of **trust and consensus**
- **Connecting disparate data types** and sources enriches research
- **Consensus data standards** are critical for the success of national and international data ecosystems – allowing aggregation across trials and diseases
- **Early adoption of data standards** and consideration for the **lifecycle of the data** are critical to accelerating progress and discovery

Do research with the PCDC and stay connected

portal.pedscommons.org

Explore the PCDC Data Portal

sam.am/datadictionaries

Access data dictionaries

slv@uchicago.edu

suzi@uchicago.edu

Sign up for our email list!

sam.am/D4CGnews

Sustainability is key

Thank you to these funders for supporting our work in FY23 and/or FY24.

Rally! Foundation for Childhood Cancer Research

St. Baldrick's Foundation — Conquer Childhood Cancers

Leukemia & Lymphoma Society®

Cancer Research Foundation

NIH National Cancer Institute

ICI Innovation in Cancer Informatics

Summer's Way

Children's Oncology Group Foundation

MIB Agents — Make It Better For Kids With Osteosarcoma — Because of Annaleigh

Children's Cancer Research Fund®

Foundation gray.

United States Department of the Interior

SebastianStrong Foundation

The Neuroblastoma Children's Cancer Society

BRIGHTSIDE

leidos

The Leona M. and Harry B. Helmsley Charitable Trust

Maddie's

Children's Research Foundation — "So they may live"

Seattle Children's — Hospital • Research • Foundation

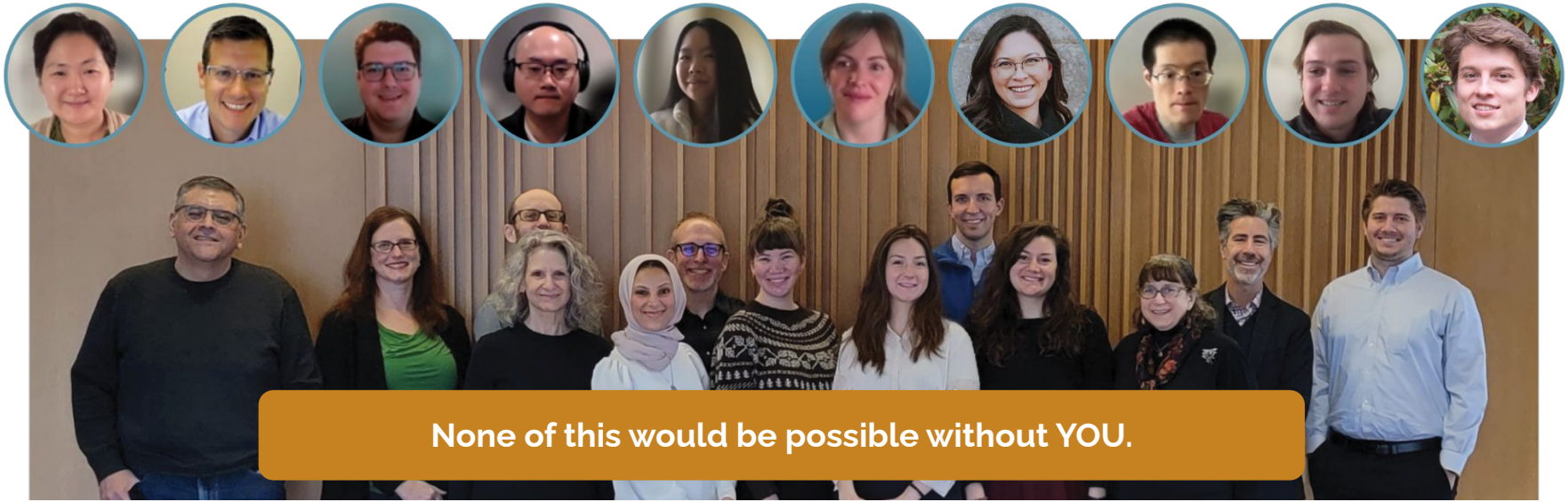
Fanconi Anemia Research Fund

Jeffrey Pride Foundation — For Pediatric Cancer Research

Aileen S. Andrew Foundation

At the Forefront of Kids' Medicine — UChicago Medicine Comer Children's — Development Board

Sarah Jane Adicoff Endowment for Research in Rhabdomyosarcoma

Thank you!

consortium members	collaborators	data portal users
advisory group members	funders	patients and families