

# Collaborations and Transformative Research Opportunities using Data Available through the CCDI Ecosystem

*Breakout Session #5*

# Collaborative Opportunities to Enhance the Use of Data

- Goal: In this breakout session, participants will discuss how data available through CCDI Ecosystem could be used for discoveries and collaborations
- Facilitators: Tony Kerlavage, PhD, Sam Volchenbom, MD, PhD

# Collaborative Opportunities to Enhance the Use of Data

- Data
  - Driving scientific questions
  - What data are needed?
  - Data Harmonization
- Tools
  - Needed to support analysis, accomplish research goals, share knowledge
- Training & Communication

# Discussion Questions

- What are the burning scientific questions to be answered?
- Does the CCDI ecosystem have the necessary data to make new discoveries?
- What additional datatypes need to be added?

## Does the CCDI ecosystem have the necessary data to make new discoveries? What additional datatypes need to be added?

Feedback from F2F:

- More robust clinical (treatment & outcomes) and non-clinical
- Data for the 7 out of 10 children who are treated off-protocol
  - Actual treatment administration on clinical studies
  - RWD from EHR extraction
- Data covering the entire continuity of communities, including those suffering from financial toxicity or other barriers
- Other useful data:
  - Sociome
  - Longitudinal: PEC, NCCR, St. Jude Life
  - Patient-submitted data
    - FDA uses PROs for evaluation

## Does the CCDI ecosystem have the necessary data to make new discoveries? What additional datatypes need to be added?

Feedback from F2F:

- There are many different user types and user groups (audiences), with differing needs. The Ecosystem needs to address each of these.
  - Audiences – informatics teams, researchers, young PIs, patient and survivors, Congress, and advocacy groups.
  - Address the questions by audience, since responses will be different for each.

## Discussion Questions

- To what extent should we focus on data harmonization?
- How can we drive harmonization upstream for data aggregation?

To what extent should we focus on data harmonization?

How can we drive harmonization upstream for data aggregation?

Feedback from F2F:

- Data harmonization
  - Variant calling: everyone does this differently, so comparing results is challenging
  - Participant Index should help to bring individual patient's data together
- Rules and provenance
  - In a federated model, consent and data sharing rules per institution can be a challenge



## Discussion Questions

- Does CCDI have the right tools to support data analysis?
- What current challenges need to be addressed in accessing CCDI data and linkages to other data to accomplish research goals?
- What are the highest priorities for the CCDI infrastructure and data ecosystem?

Does CCDI have the right tools to support data analysis?

What current challenges need to be addressed in accessing CCDI data and linkages to other data to accomplish research goals?

What are the highest priorities for the CCDI infrastructure and data ecosystem?

Feedback from F2F:

- Provide a framework to take additional data into the ecosystem
  - Wearables, patient-provided data
  - Include FDA in the conversation

## Discussion Questions

- What training is needed to effectively access and analyze the data?
- What kinds of communications would be most effective?

## What training is needed to effectively access and analyze the data? What kinds of communications would be most effective?

### Feedback from F2F:

- Communications are critical and require focused attention and scientific / technical expertise
  - Do clinicians and families in underserved communities know about CCDI?
  - Do clinicians and researchers understand access to the data is “free”?
  - Researchers’ and clinicians’ needs should be the first priority and ensuring they know how to communicate back to the patients.
  - Develop data sharing language the foundations can use in their grants.
- Metrics for success need to be defined, and these could also vary by audience.
  - Training on the tools and data is required.
- Consider surveys and focus groups for clinicians to find out what they need and educate them.
- Take lessons learned from *All of Us* for patient engagement.
- Do workshops at ASCO, AACR, COG Annual Meeting, and other big scientific meetings.

# Discussion Questions

- Can CCDI data be leveraged for clinical trials?

## Can CCDI data be leveraged for clinical trials?

Feedback from F2F:

- Clinical trials take too long to get data in meaningful time
  - A game changer CCDI could deliver is the idea of synthetic cohorts, especially for rare diseases; soon all the diseases will be rare.
- CCDI is up to age 39, which also changes the dynamic - we need to think about AYA space as well.
- Potential for collaboration between advocacy and industry partners.
- Pharma is joining an alliance for sharing data for reuse/synthetic cohorts – how else can we work with them?

# CCDI Data Summary

## ❖ MSKCC: Solid tumors

- ❖ Title: Feasibility and Clinical Utility of Whole Genome Profiling in Pediatric and Young Adult Cancers (phs002620)
- ❖ 114 participants with 326 samples

## ❖ MSKCC: High-risk neuroblastoma

- ❖ Title: Clonal Evolution During Metastatic Spread in High-Risk Neuroblastoma (phs003111)
- ❖ 130 participants with 488 samples

## ❖ OHSU: Acute myeloid leukemia

- ❖ Title: NCI CCSG CCDI Supplement Additional Genomic Submission (phs002599)
- ❖ 104 participants with 152 samples

## ❖ UCSF: Juvenile myelomonocytic leukemia

- ❖ Title: Database for the Advancement of JMML - Integration of Metadata with "Omic" Data (CCDI) (phs002504)
- ❖ 189 participants with 329 samples

## ❖ CMRI/KUCC: Solid tumors & leukemias

- ❖ Title: Comprehensive Genomic Sequencing of Pediatric Cancer Cases (phs002529)
- ❖ 194 participants with 373 samples

## ❖ USC: Multiple cancer types

- ❖ Title: NGS Panel for Pediatric Malignancies (phs002518)
- ❖ 1039 participants with 2246 samples

## ❖ CCDI-COG: CNS, STS, Rare tumors

- ❖ Title: Molecular Characterization Initiative (phs002790)
- ❖ 880 participants with 2440 samples

## ❖ CHOP: Brain & other solid, hematologic malignancies

- ❖ Title: Integration of genomic and clinical data from unique rare cancer datasets to facilitate data sharing (phs002517): currently being indexed
- ❖ 1031 participants with 3255 samples

<https://datacatalog.ccdi.cancer.gov/resource/CCDI>

# CCDI Data (coming soon)

- ❖ **Children's Cancer Inst Australia: Acute lymphoblastic leukemia**
  - ❖ Title: CCDI Pediatric In Vivo Testing Program – Leukemia (phs003164)
- ❖ **Children's Hospital of Philadelphia: Neuroblastoma**
  - ❖ Title: CCDI Pediatric In Vivo Testing Program – Neuroblastoma (phs003163)
- ❖ **Children's Hospital of Chicago: Orthotopic CNS tumors**
  - ❖ Title: CCDI Pediatric In Vivo Testing Program – Orthotopic CNS Tumors (phs003162)
- ❖ **Memorial Sloan Kettering Cancer Center: Sarcomas & other solid tumors**
  - ❖ Title: CCDI Pediatric In Vivo Testing Program – Sarcomas and other Solid Tumors (phs003161)
- ❖ **UT Health Science Center San Antonio: Ewing sarcoma, rhabdomyosarcoma, kidney, and liver cancers**
  - ❖ Title: CCDI Pediatric In Vivo Testing Program – Ewing sarcoma, rhabdomyosarcoma, kidney, and liver cancers (phs003160)
- ❖ **SickKids | The Hospital for Sick Children: Bone and Soft tissue cancers**
  - ❖ Title: Discovering the Timing and Origins of Bone and Soft Tissue Cancers (phs002827)
- ❖ **St. Jude Children's Research Hospital: Childhood Cancer Survivor Study (phs001327)**
- ❖ **NCI-COG: Pediatric MATCH correlative studies: Solid tumors**
- ❖ **Dana-Farber Cancer Institute: Integrating Longitudinal Clinical, Sociodemographic and Genomic Data into the NCCR (phs002677)**
- ❖ **Baylor College of Medicine: Incorporation of Genomic Sequencing into Pediatric Cancer Care (phs001683)**
- ❖ **Univ of Michigan: Pediatric/AYA Cancer Touchstone data**
  - ❖ Title: Enhancement of Data Sharing in Pediatric, Adolescent and Young Adult Cancers (phs002431)
- ❖ **Texas Pediatric Patient Derived Xenograft (phs003215)**
- ❖ **MSKCC: Enhance Data Sharing in Pediatric, Adolescent, and Young Adult Cancers (phs002530)**
- ❖ **UCSF: Integration of genomic and clinical data from unique rare solid tumor datasets to facilitate data sharing (phs002430)**



# Data Access & Processing

## Study-level Directories

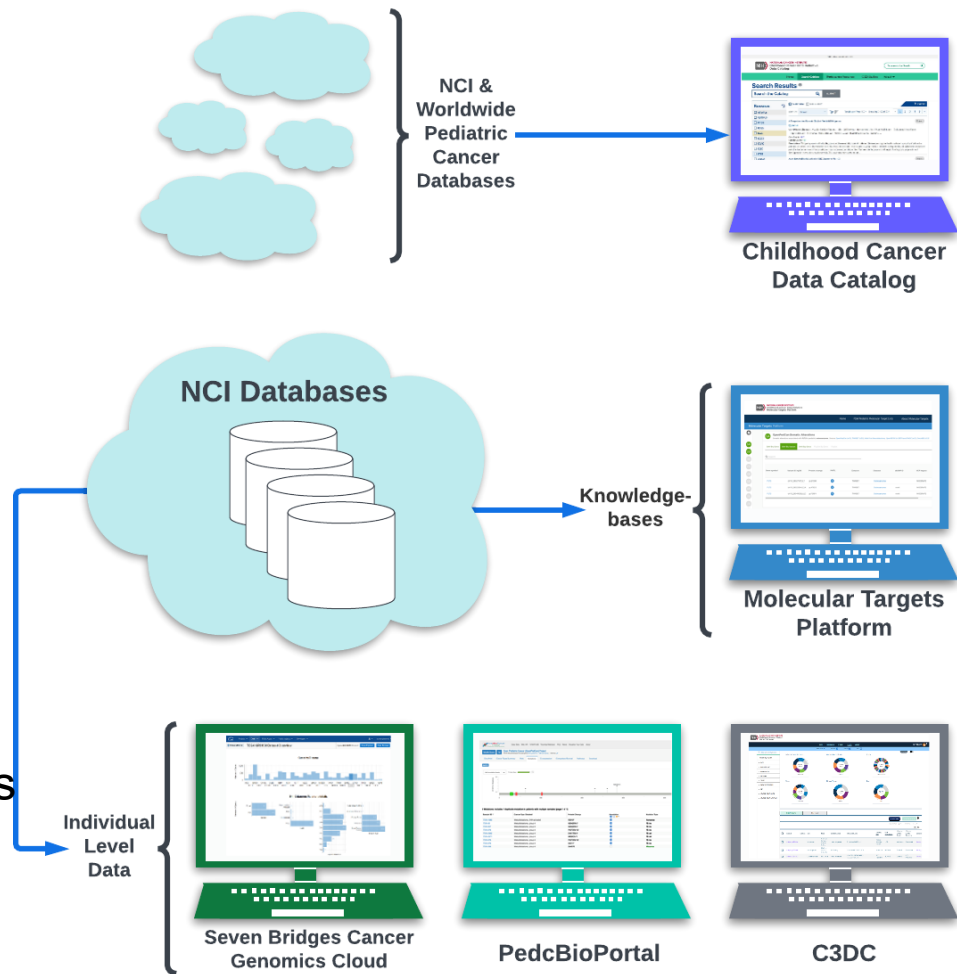
- Childhood Cancer Data Catalog

## Aggregations and knowledge bases

- Molecular Targets Platform

## Individual-level data

- Clinical: C3DC
- Genomics: PedcBioPortal
- Custom analyses: Cancer Genomics Cloud



U.S. Department of Health & Human Services  
National Institutes of Health | National Cancer Institute

[cancer.gov](https://cancer.gov)  
1-800-4-CANCER