

Childhood Cancer Data Initiative (CCDI) Scientific Session Report

Tony Kerlavage, Ph.D.

Jaime M. Guidry Auvil, Ph.D.

Childhood Cancer Data Initiative: Background and Goals

Tony Kerlavage, Ph.D.

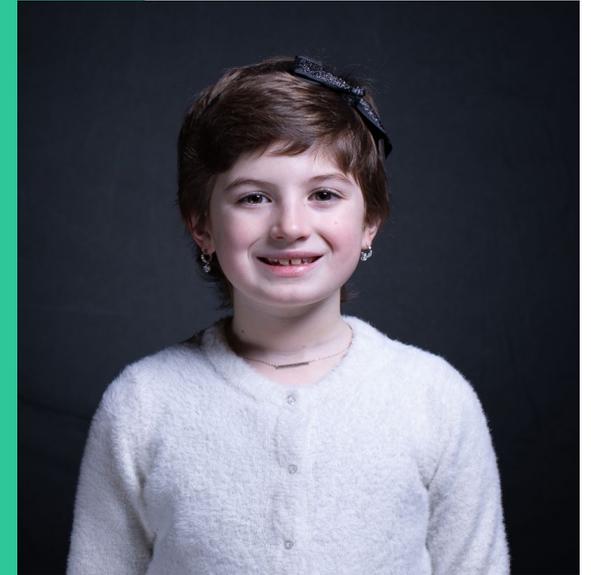
Director

*Center for Biomedical Informatics
& Information Technology*

Childhood Cancer Research Presidential Initiative

STATE
★ of the ★
UNION
2 0 1 9

- President's FY 2020 budget proposal includes \$50M
- White House has convened several stakeholder events





Childhood cancers represent about

1%

of all new cases of cancer diagnosed in the United States each year

Collaboration among children's cancer centers will continue to be essential to ensure that trials enroll sufficient numbers of participants to produce meaningful results



Answering complex scientific questions about childhood cancer requires larger datasets





Critical Importance of Pediatric Cancer Data

Even our most effective treatments don't work for all patients

Improve understanding of why some cancers develop resistance or don't respond to treatment

Short- and long-term adverse effects of cancer and its treatment

Identify less toxic treatments and strategies for management

Virtually no progress for some cancer types

Generate new ideas for interventions



Common Challenges When Sharing Data

- Unprecedented amount of data produced from basic and clinical research and care
- Different data types, sets, tools, infrastructure/ repositories and collections in silos
- Varying data collection practices, nomenclature, platforms, and policies
- Systems that don't "talk" to each other
- Privacy protections, informed consent
- Organizational culture, norms, habits, and attitudes



NCI Childhood Cancer Data Initiative Goals

- **Maximize** every opportunity to improve treatments and outcomes for children with cancer
- **Build** a connected data infrastructure to enable sharing of childhood cancer data from multiple sources
- **Identify** opportunities to make data work better for patients, clinicians, and researchers
- **Develop** and enhance tools and methods to extract knowledge from data



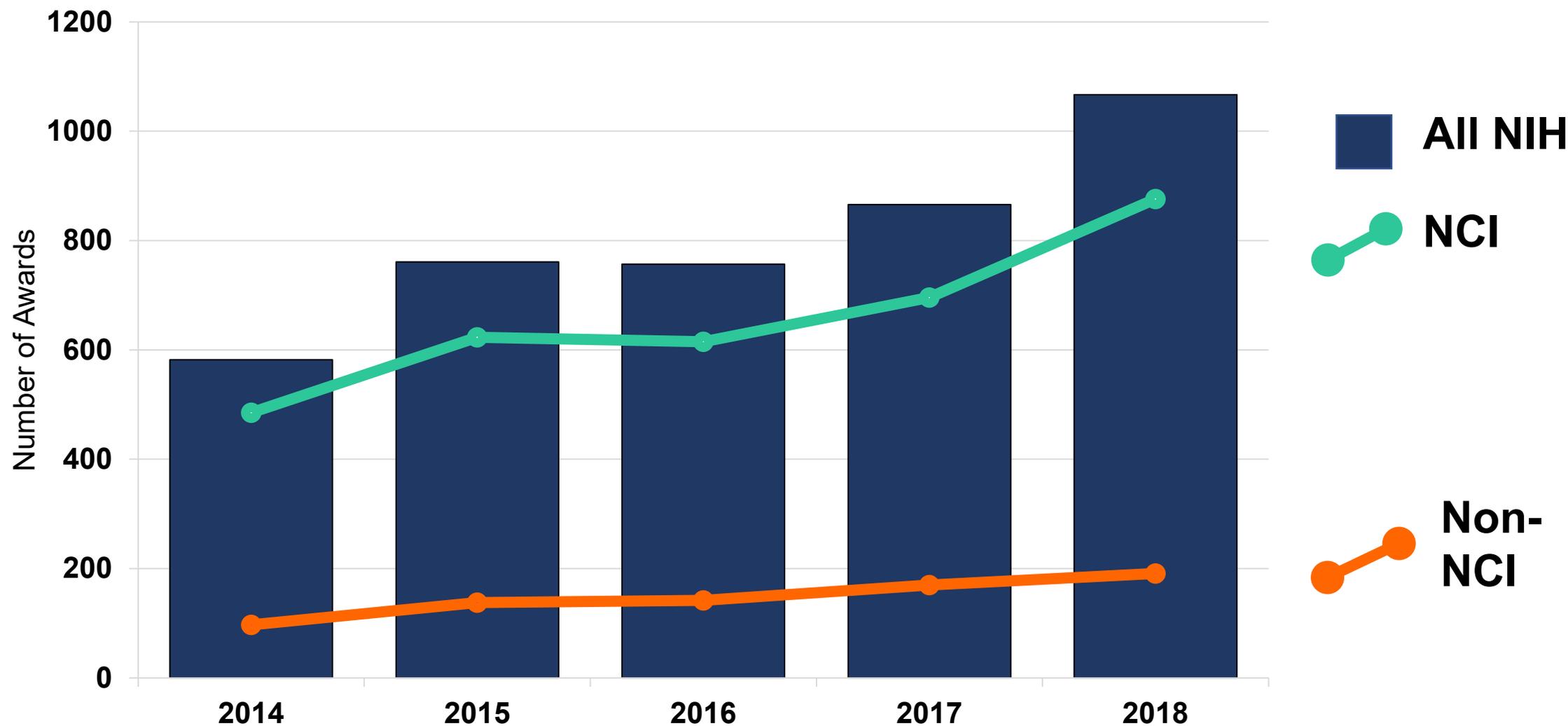


Pediatric Cancer Research: Progress and Current Activities

Jaime M. Guidry Auvil, Ph.D.

Director, Office of Data Sharing, CBIIT

Pediatric cancer research conducted or supported by NCI/NIH: 2014 - 2018



Acute Lymphocytic Leukemia

1975 - 1977

5-year relative
survival rate

54%

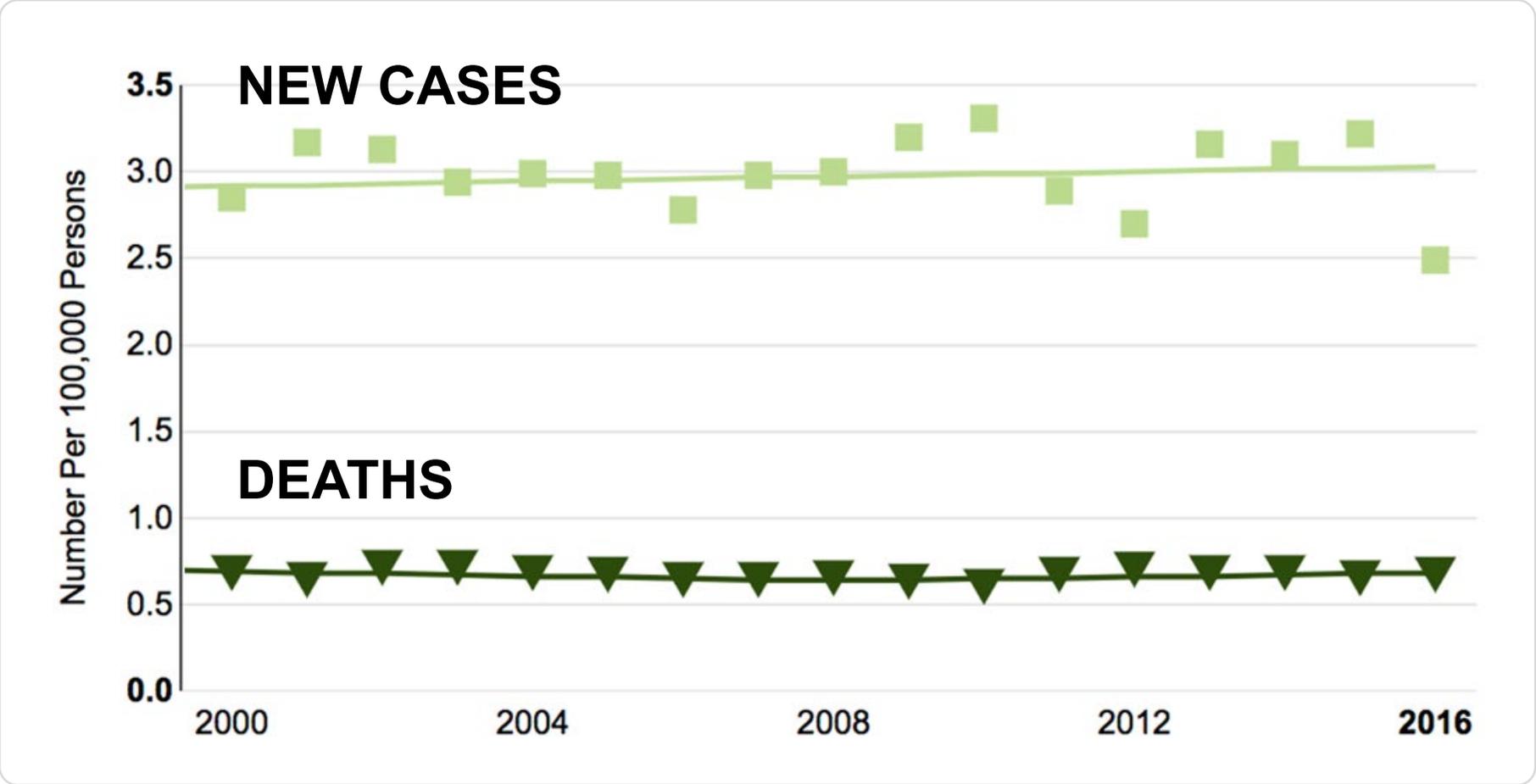


2009 - 2015

5-year relative
survival rate

90%

Childhood Brain and Other Nervous System Cancer (Ages 0-19)



**5-Year
Survival
75%**

Serious Long-term Side Effects in Childhood Cancer Survivors

May affect all major organs and systems:

- **Cardiovascular, Respiratory, Digestive, Reproductive**
- **Cognitive development and function**
- **Hearing, vision**

Secondary cancers are a serious concern for survivors of childhood cancer

60% to more than 90% develop one or more **chronic** health conditions

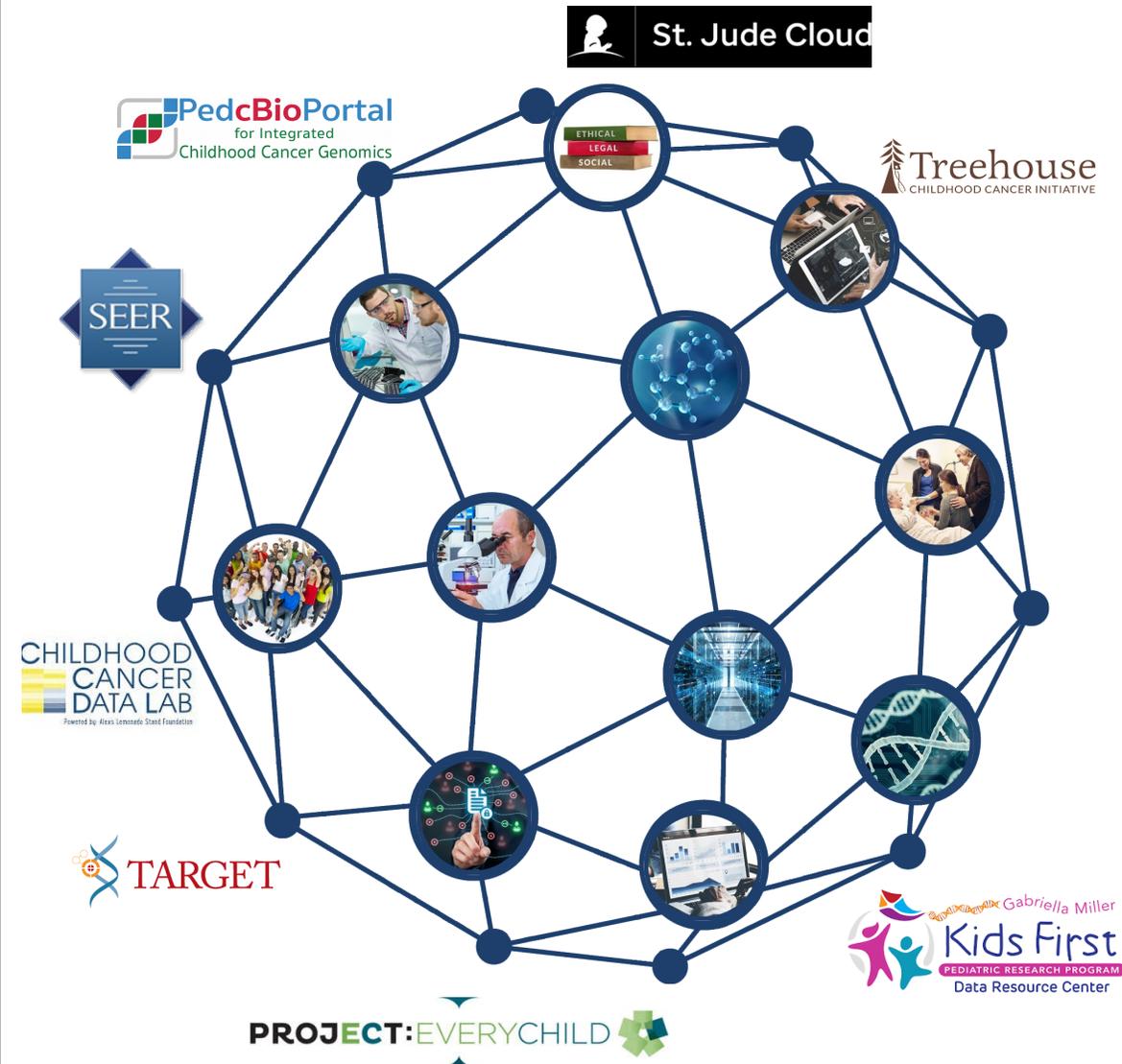
Depending on the cancer type, **20% to 80%** experience **severe** or **life-threatening** complications during adulthood



CCDI: Connecting NCI-funded Data

NCI funded data stored in various repositories: GDC, CRDC, SRA, others:

- **Childhood Cancer Survivorship Study (CCSS)** – NCBI (SRA/ dbGaP), St. Jude Cloud
- **Children's Oncology Group (COG)** – GDC, NCBI, Kids First DRC, Univ. Chicago Pediatric Data Commons, Treehouse
- **Pediatric MATCH** – GDC, CRDC, NCBI, Registries
- **Therapeutically Applicable Research to Generate Effective Treatments (TARGET)** – GDC, ICGC, NCBI, KF-DRC, Treehouse
- **Gabriella Miller Kids First Research Program** – DRC, Cavatica, NCBI
- Pediatric data in cancer registries – SEER, PHIS, VPR



CCDI: If We Build It, What Answers Will Come?

- Scientific questions that could be answered:
 - Prevention & treatment of secondary cancers, long term treatment side effects, and continuum of care across all types of childhood & AYA cancers
 - Improve understanding and discover novel therapeutic options for childhood and AYA cancers through increased power and information, including lesser known subtypes, relapsed/refractory cancers, etc.
 - Validate data to support the molecular targets list for childhood and AYA cancers (RACE Act)



Connect data from:

Basic Research
Clinical Trials
Real-world Patient Data
Population Studies
Pre-clinical Models
Biospecimen Repositories



Childhood Cancer Data Initiative: Symposium Summary

CCDI Symposium Agenda

Monday, July 29

- NCI Director's Perspective & Chat with FDA Deputy Commissioner [Videocast](#)
- Envisioning a National Cancer Data Initiative Panel Discussion [Videocast](#)
- Patient-Centered Data Management for Discovery [Videocast](#)
- Poster Session

Tuesday, July 30

- Setting the Stage for CCDI – Panel Discussions [Videocast](#)
- CCDI Focus Area Breakout Track Sessions

Wednesday, July 31

- Breakout Track Sessions Summaries/Reports Out [Videocast](#)
- Roundtable Discussion with Leadership [Videocast](#)
- Final thoughts and charge for future [Videocast](#)



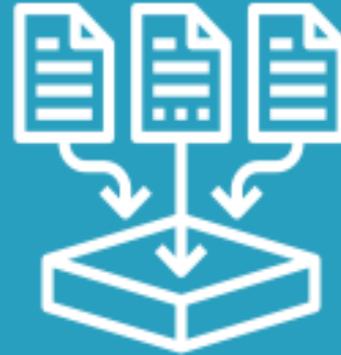


Four Areas of Focus

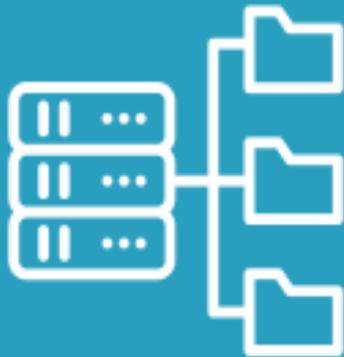
1. Prioritizing scientific and clinical research data needs for therapeutic progress



2. Creating meaningful datasets for clinical care and associated research progress



3. Infrastructure to enable federation among pediatric data repositories



4. Development of tools and methods to extract knowledge from data





Topics from CCDI Symposium Sessions

Connect Pediatric Datasets

- Identify and connect **key pediatric research & clinical data repositories**
- Aggregate **COG study data** that can be queried efficiently
- Instate **universal identifier**, issued by honest broker

Collect Pediatric Datasets

- Initiate a **National Cohort of All Pediatric Cancer Patients** – data from discovery science to clinical studies and care to surveillance
- Implementation of a **Prototype Master Protocol** with **structured data capture** from clinical to molecular to outcomes
- **Harmonize preclinical data** and develop query portal for all investigators and clinicians



Topics from CCDI Symposium Sessions

Annotate Pediatric Datasets

- Create a **central resource catalog** of available pediatric/AYA data, biospecimens, tools, analysis workflows, and other resources
- Establish standards to generate **structured phenomic data** from EMRs and clinical trials
- Develop **standards & best practices** for sharing or transferring data between repositories

Extracting Knowledge from Data

- Tools to **automate extraction of structured & unstructured data**
- Tools for **standardized “front end” data capture**
- Agree on a few **standard analysis workflows** that can be used to harmonize data across resources
- Tools to **capture Real World Data & Patient Reported Outcomes**

Childhood Cancer Data Initiative: Summary & Next Steps



Some Next Steps for CCDI

- Summary of Cancer Research Ideas community input (September 6th)
- Post-CCDI Symposium Webinar (end of September)
- NCI Position Paper on CCDI Approach (Fall 2019)
- CCDI Scientific Planning (NCI Leadership Team)
 - Create & maintain resource catalog for reference to and reuse of data, biospecimens, and tools
 - Connect pediatric data repositories and registries that already hold NCI-funded data
 - Define criteria for “ideal” datasets through connecting existing data and collecting new data

Culture Changes are Needed for Success

- NCI has lifted publication embargo on TARGET osteosarcoma data
- NCI will provide resources and the environment to make data sharing possible
- Community must think about what they will do differently to help facilitate the sharing and use of data for benefit of children with cancer

Principles for Success

Please do:

- Use Digital IDs
- Interoperate with third party authentication and authorization services
- Expose your data through an API
- Expose your data model through an API
- Interoperate with other trusted data commons with similar security and compliance
- Process authorized queries from other systems and return the results

Please don't:

- Refuse to expose any API and instead require all users to use your platform
- Bring data from other resources and platforms into your system, but don't let your data out
- Refuse to interoperate with other systems with the same or greater security and compliance

Childhood Cancer Data Initiative

Facilitate the sharing of childhood cancer data from multiple sources through a connected data infrastructure

FEDERATE



Identify opportunities to align & integrate multiple data sources to make data work better for patients, clinicians, researchers

AUGMENT



Maximize every opportunity to improve treatments and outcomes for children with cancer

USE





Questions & Comments