Childhood Cancer Data Initiative: Activities To-Date and Next Steps

Joint Meeting of the NCI National Cancer Advisory Board and the NCI Board of Scientific Advisors

Warren A. Kibbe, PhD

Gregory H. Reaman, MD

NIH NATIONAL CANCER INSTITUTE

December 6, 2022

- *CCDI Overview (Structure, Goals, and Objectives)*
- *CCDI Activities in Phase I/Foundational Phase*
- *Building on the Foundation – Identified Priorities & Next Steps*

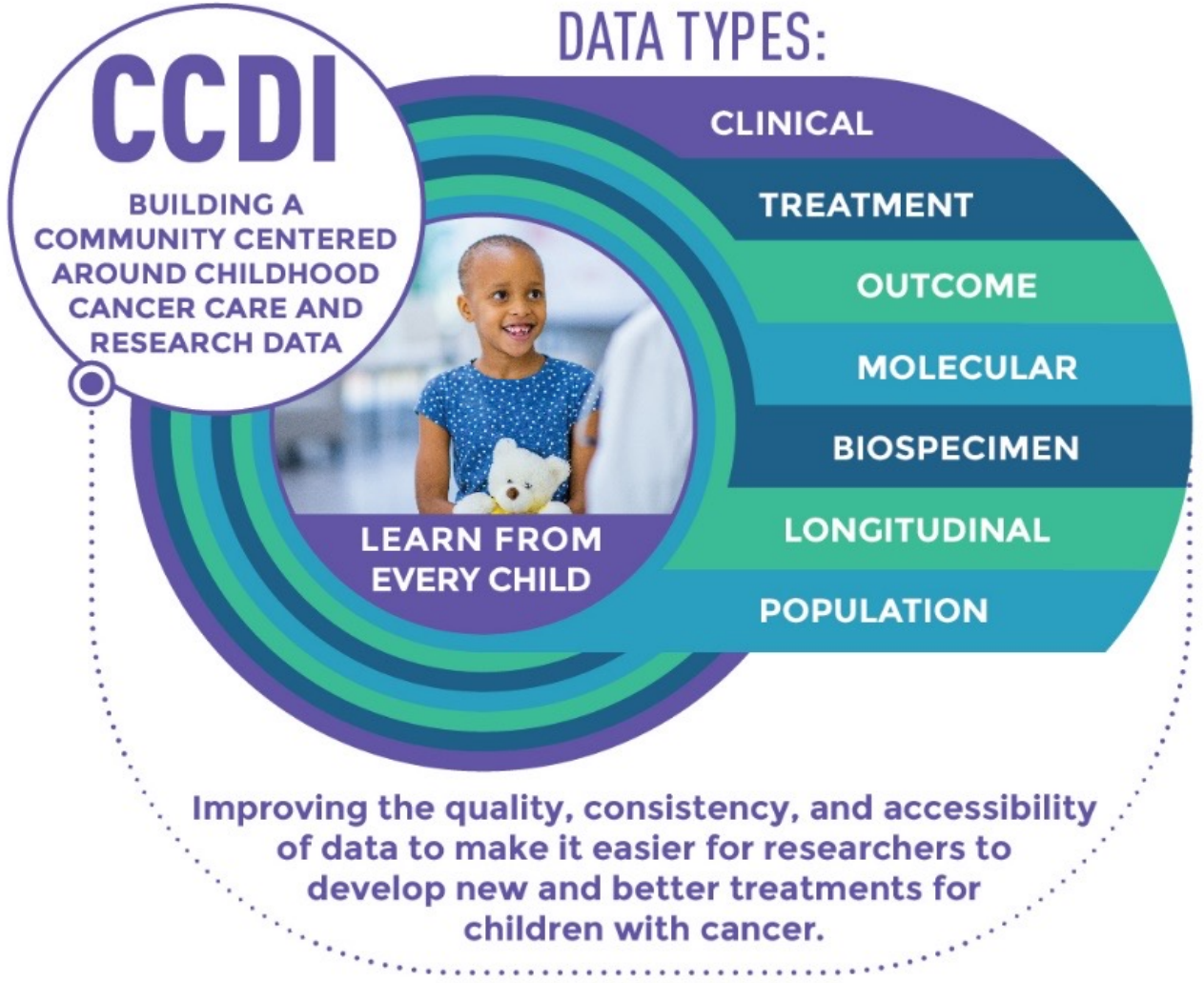
Program Overview: Structure, Goals and Objectives

NIH NATIONAL CANCER INSTITUTE

CCDI is supporting the community of pediatric cancer researchers, advocates, families, hospitals, and networks committed to generating, using and sharing data to improve treatments, quality of life, and survivorship of every child with cancer.

Foundational Goals for CCDI

- Gather data from every child, adolescent, and young adult diagnosed with a childhood cancer, regardless of where they receive their care
- Create a national strategy of appropriate clinical and molecular characterization to speed diagnosis and inform treatment for all types of childhood cancers
- Develop a platform and tools to bring together clinical care and research data that will improve preventive measures, treatment, quality of life, and survivorship for childhood cancers



BSA CCDI Working Group Report – Program Goals and Objectives

We began with priorities identified by the BSA WG

"Focus on the critical need to **collect, analyze, and share data better** in order to maximize NCI's ongoing investment in pediatric and AYA cancers and survivorship. Tissue samples from patients with cancer in this age group are critically limited and a valuable resource. The overall goal of the CCDI is not simply to generate more data, but to **build processes that transform data into knowledge that moves the field forward in meaningful ways**. The CCDI supports the wider pediatric cancer community's goal of maximizing the stated goal of ***"learning from every patient"*** so that ultimately those patients, survivors and their families can materially benefit in terms of higher cure rates and improved long-term health outcomes."

Categories of Recommendations

- Aggregate and generate broad categories of data
- Develop infrastructure
- Engage with experts
- Empower patients & families
- Ensure appropriate policy and funding
- Develop strategy for survivorship
- Ensure diverse patient representation
- Enable improved patient outcomes and treatment

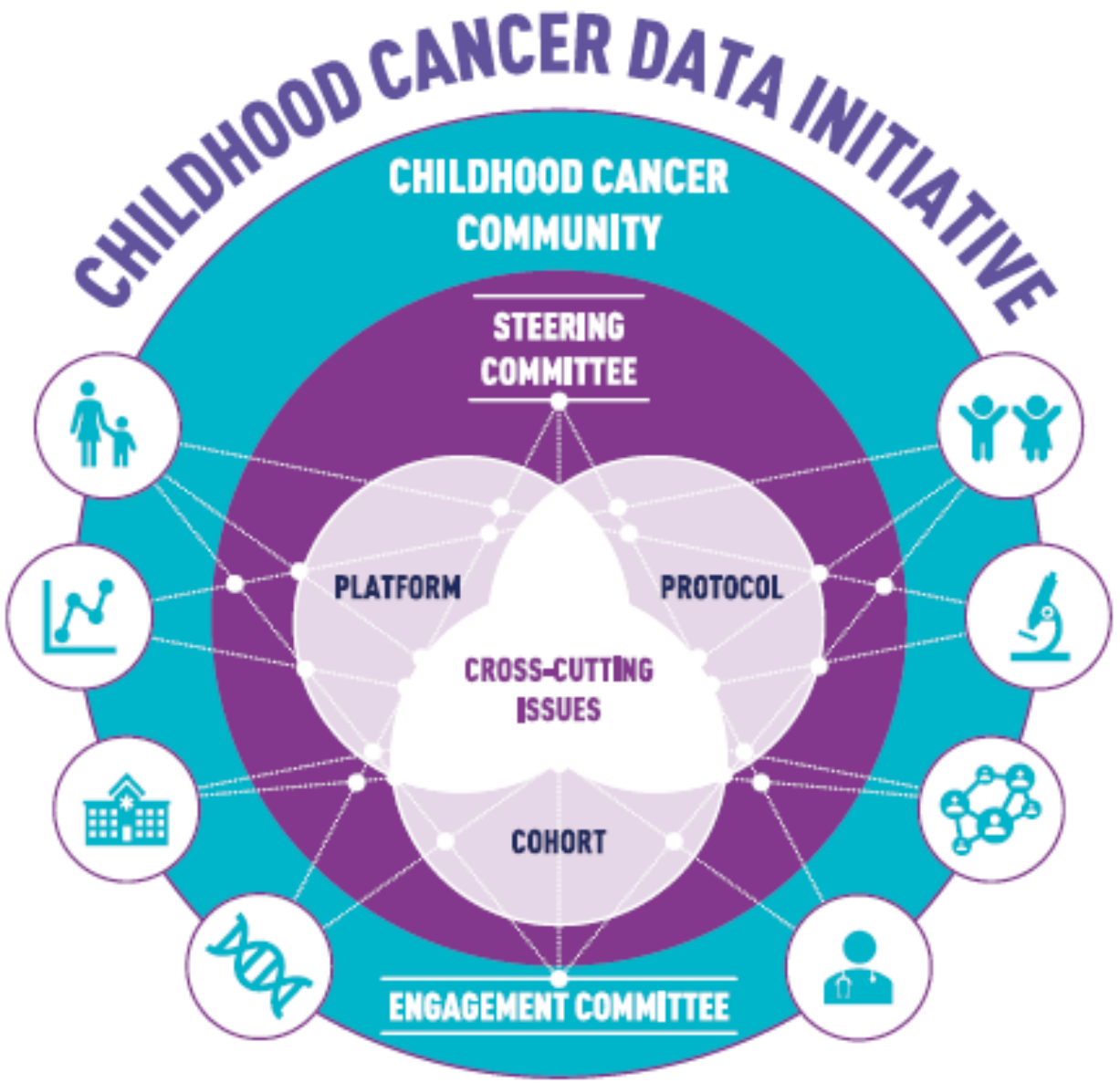
CCDI is a Unique Initiative for NCI

- Focus on how we can learn from every patient through data aggregation, harmonization, interoperability, sharing, and research use across many types of data within multiple, disparate systems
- Opportunity to align multiple ongoing and prospective activities and drive progress toward a **common vision** for the full research community by collecting, sharing and using data
  - Creating a long-term, sustainable, continually evolving and growing data ecosystem for the pediatric & AYA cancer community
  - This will be a living resource that integrates multiple pediatric programs and resources – creates a Learning Healthcare (and Research) System
- NCI CCDI program staff focus on how their activities fit within and progress towards that vision, in addition to priorities for individual programs (to ensure greatest impact for the wider community)
- NCI leaders receive regular external feedback to maintain alignment with priorities of the full pediatric and AYA research community through diverse expertise organized into CCDI working groups, as well as open meetings and workshops
  - The working groups were the primary focus of CCDI Leadership Team in second and third years

CCDI Working Framework

CCDI is establishing a **pediatric cancer data resource** for the cancer research community; teams are focused on coordinating across that long-term objective

- Building the Data Infrastructure: Establishing and evolving a data infrastructure is a foundational goal for the program and comprised the first set of activities
- Data Collection, Generation, and Aggregation: Bringing together existing and novel data to fill gaps
- Learning from the Data: As data become accessible and interoperable, new discoveries can be made and new cohorts developed.
- Working Groups (clinicians, research and data scientists, advocates) established to provide insight and alignment in the community



CCDI Leadership Structure

Strategic Direction and Program Coordination

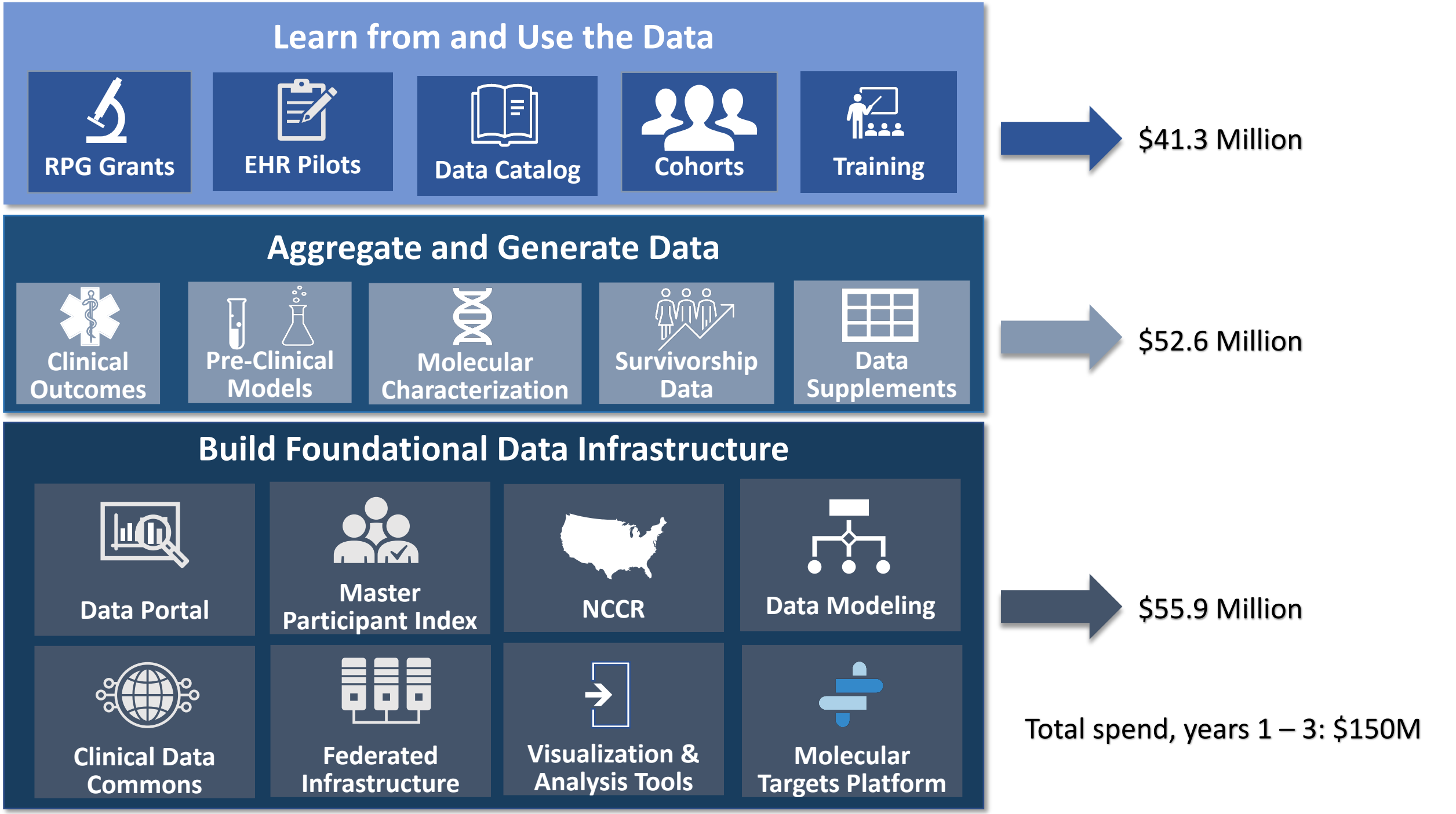
- Steering Committee Co-Chairs – Dr. James Doroshow (Deputy Director Clinical & Translational Research, DCTD) and Dr. Warren Kibbe (Chief Data Officer, Duke Cancer Institute; NCI Data Advisor)
- Dr. Gregory Reaman, Scientific Director CCDI
- Dr. Jaime Guidry Auvil, Director Office of Data Sharing (Cross-Cutting Committee Co-Chair and Program Coordination)
- Dr. Tony Kerlavage, Director Center Biomedical Informatics and Information Technology (Data Platform Co-Chair, NCI Data Ecosystem Infrastructure Lead)
- Dr. Brigitte Widemann, Special Advisor to the NCI Director for Childhood Cancer (Engagement Committee Co-Chair, CCR)

CCDI Leadership Team: CCDI Strategic Leads, NCI Director & Deputies, and Senior Program Leads

- Dr. Stephen Chanock (Cancer Cohort Co-Chair, DCEG)
- Dr. Lynne Penberthy (Cancer Cohort Co-Chair, DCCPS)
- Dr. Malcolm Smith (Molecular Characterization Co-Chair, DCTD)
- Dr. Jack Shern (Molecular Characterization Co-Chair, CCR)
- Dr. Jason Levine (Data Platform Co-Chair, CCR)
- Dr. Monica Bertagnolli (OD)
- Dr. Doug Lowy (OD)
- Dr. Dinah Singer (OD)
- Dr. Louis Staudt (CCR, CCG)
- Dr. Jean Claude Zenklusen (CCG)

CCDI Activities in Phase I: Foundational Phase

NIH NATIONAL CANCER INSTITUTE


Learn from and Use the Data
RPG Grants
EHR Pilots
Data Catalog
Cohorts
Training
$41.3 Million
Aggregate and Generate Data
Clinical Outcomes
Pre-Clinical Models
Molecular Characterization
Survivorship Data
Data Supplements
$52.6 Million
Build Foundational Data Infrastructure
Data Portal
Master Participant Index
NCCR
Data Modeling
Clinical Data Commons
Federated Infrastructure
Visualization & Analysis Tools
Molecular Targets Platform
$55.9 Million
Total spend, years 1 – 3: $150M

	Build Data Infrastructure	Aggregate & Generate Data	Learn from & Use Data
January 2020 BSA WG Report	• **National Childhood Cancer Registry (NCCR)** • Index of NCI studies • **Molecular Targets Platform** • Data Harmonization & Modeling	• Cancer Center Data Supplements (Registries, Omics, Imaging) • HCMI Supplements (organoids) • Research Sequencing Supplements (subsets: CCSS, PedMATCH, PDX & Cell Line) – *CNS & sarcomas in FY22*	• **Childhood Cancer Data Catalog** • Proton-Photon Therapy Centers • Intramural/Extramural Grants (diverse areas of data focus including pediatric underserved populations) – *FY 20 – FY22*
FY 2021 BSA/NCAB Update Establishment of Working groups	• CCDI Participant Index (universal identifier) • Clinical Data Commons • Data Submission Pilot (deWrangler)	• Rare Pediatric Tumor Cell Atlas – *FY20 – FY22* • **Molecular Characterization Initiative**	• **NCCR PedsExplorer** • Training Awards – *FY21 & FY22*
FY 2022 WG Priorities Presentation Annual Symposium	• ITCR Supplements for Tools for Ped/AYA (analysis, visualization & CT matching) - *FY20 & FY22* • Federated APIs pilot – key Pediatric databases (St. Jude Cloud, PCDC, TreeHouse, Cavatica) • CCDI Data Portal	• PiVOT Preclinical Program • COG Supplements (Project EveryChild enrollment, clinical data, Phase II clinical trials to NCTN archive)	• EHR Pilots (adverse events/FDA, state registries) – *FY20 & FY22*
FY 2023 CCDI Workshops: EHR, Rare Tumor BSA/NCAB Update			

NCI actively planning next steps to align with community feedback

Years 1 – 3: Build Data Infrastructure Portfolio

- Develop a Federated Pediatric Cancer Data Ecosystem of research repositories & patient registries:
  - ✓ Establish **National Childhood Cancer Registry** (NCCR) to link clinical patient data
  - ✓ Build a Federated Data Infrastructure that includes a **Molecular Targets Platform** (launched April 2022) to analyze preclinical data that can inform/ validate FDA Relevant Molecular Targets List, a **CCDI Participant Index** to facilitate subject and biospecimen tracking, an **Index of NCI Studies** (launched Fall 2022), and a **Clinical Data Commons** for access to subject level demographic, treatment and phenomic information in standard format
  - ✓ Develop **APIs for federation** with key pediatric databases/partners (St. Jude Cloud, PCDC, TreeHouse, Cavatica) and a **Data Submission Pipeline** to simplify/automate data aggregation
- Develop or adapt **analytic tools & computational methods** using grants and contracts for use in childhood and AYA cancer research
  - ✓ Automated curation of data (e.g., natural language processing) for refining, scaling & real-world data capture
  - ✓ Interpreting pathology images & patient reports
  - ✓ Pediatric data model & terminology harmonization

Years 1 – 3: Aggregate and Generate Data Portfolio

- Develop a national strategy to develop consensus guidelines for pediatric and AYA cancers
  - ✓ **Molecular characterization initiative (MCI)** - Multiple COG supplementation, data flow, vendor contract
    - ✓ 569 clinical reports returned to physicians, as of end of November, 2022
    - ✓ First data release to CCDI Data Ecosystem – sequencing data from first 380 patients (updated monthly)
- Aggregate existing data through transfer of patient-linked clinical (phenomics, treatment) and molecular data (genomics, proteomics, imaging, preclinical) and analytic tools to NCI resources
  - ✓ **Data & tools supplements (Cancer Centers)** - registries and data repositories; **Clinical Data (CCSS)**
- Generate new cancer models and sequence data to fill gaps for key NCI initiatives; data will be submitted to NCI databases
  - ✓ Diagnostic tumors and germline samples from Pediatric MATCH, CNS tumors (CBTN) and soft tissue sarcomas; secondary cancers from CCSS
  - ✓ Pre-Clinical Data Efforts (PiVOT, Molecular Targets Platform); organoids and cell lines, including CNS tumors
- Establish a **Rare Pediatric Tumor Cell Atlas** from tissues obtained through the NCI Pediatric Rare Tumor Network

Years 1 – 3: Learn From and Use the Data Portfolio

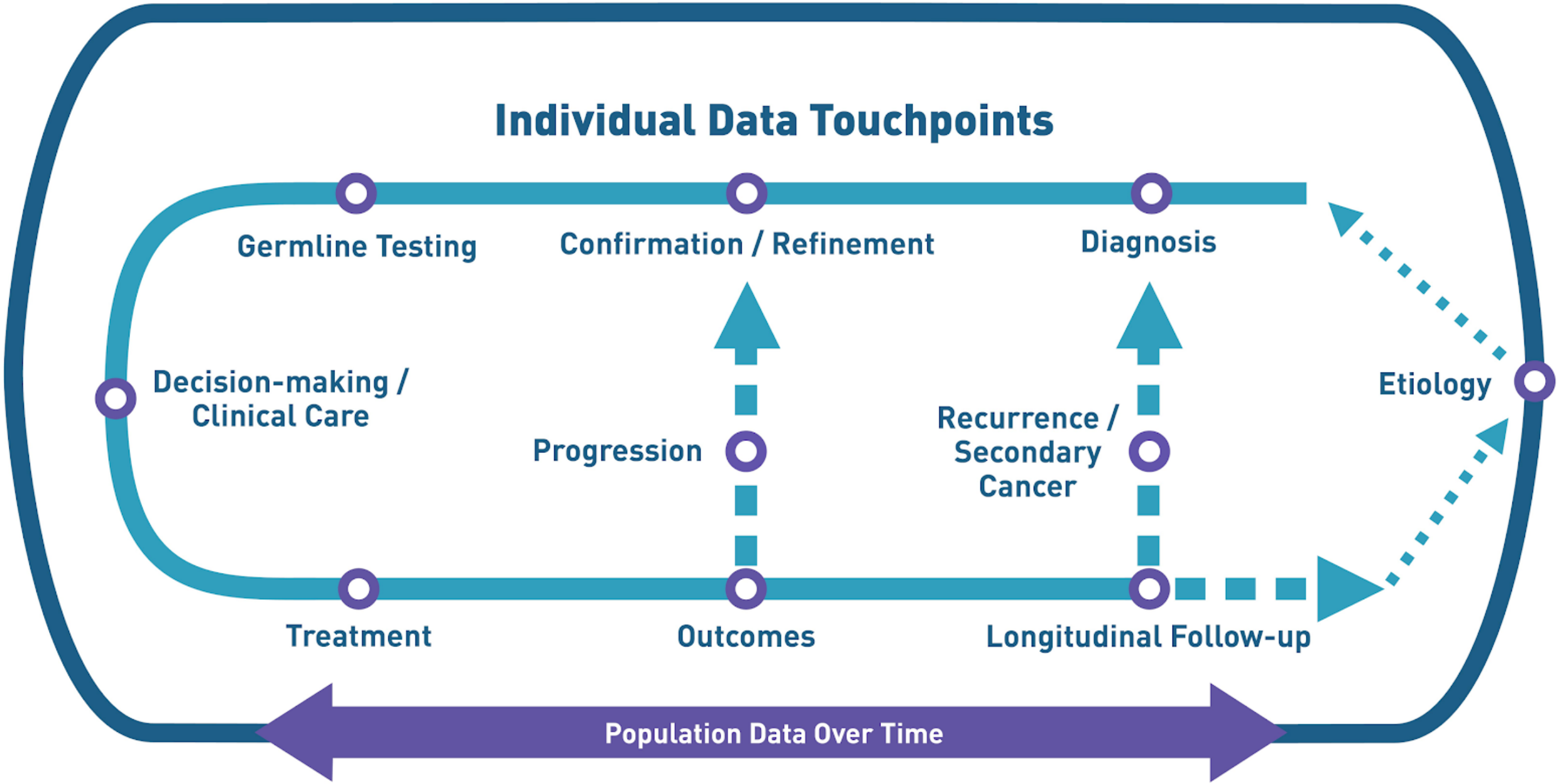
- Develop a **Childhood Cancer Data Catalog** (launched April 2022) of all available childhood cancer data registries and data repositories
- **NCCR PedsExplorer** (launched Sept. 2021)
- Supplement **intramural** and **extramural grants, contracts** supporting:
  - ✓ Etiological and clinical risk prediction and genetic susceptibility for the development of childhood malignancies
  - ✓ Patient-reported outcomes and toxicities in pediatric and AYA patients and survivors
  - ✓ Improvements to childhood clinical trials data reporting
  - ✓ Molecular pathogenesis of pediatric and AYA cancer development
- **EHR Data Extraction Pilots** to test methods for automated retrieval of data from electronic health records
  - ✓ Lab data quality and efficiency of conduct in pediatric clinical trials
  - ✓ Improve automated reporting of detailed clinical data to state registries

CCDI Phase I-Supported - Data and Tools Released

Item	Link
NCCR Pediatric Explorer	https://nccrexplorer.ccdi.cancer.gov
CIViC Data Platform	https://civicdb.org
Childhood Cancer Data Catalog	https://datacatalog.ccdi.cancer.gov/
Molecular Targets Platform	https://moleculartargets.ccdi.cancer.gov
Available Individual-level Genomics Data (dbGaP)	• phs002790.v1.p1 (CNS, STS, Rare tumors - MCI) • phs002599.v1.p1 (Acute Myeloid Leukemia - OHSU) • phs002504.v1.p1 (Juvenile myelomonocytic leukemia - UCSF) • phs002620.v1.p1 (Solid tumors - MSKCC)
MCI First Data Release	• Available through dbGaP and Seven Bridges Cloud Resource

Building on the Foundation
Identified Priorities & Next Steps

NIH NATIONAL CANCER INSTITUTE


Individual Data Touchpoints
Germline Testing
Confirmation / Refinement
Diagnosis
Decision-making / Clinical Care
Progression
Recurrence / Secondary Cancer
Etiology
Treatment
Outcomes
Longitudinal Follow-up
Population Data Over Time

CCDI High Priorities - Confirmed Across Working Groups

Priority	What CCDI Has Funded To-date (In-progress)	Future CCDI Plans
Patient Identifiers: Required to connect patients across repositories for research, while preserving patients' privacy	• CCDI Participant Index • NCCR PPRL	• Incorporate universal IDs for CCDI • Work with COG on alignment with COG identifiers
Data Models and Standards: Required to enable data federation & interoperability (API)	• Pediatric Clinical Data Commons • Data harmonization effort across data sets	• Incorporate harmonized data model into CCDI supported projects • Work to define standards across ecosystem
Consent: Consent patients early, and to recontact or to opt-out at age of majority; power in the hands of the patients and families	• Updated for Molecular Characterization Initiative (Project:EveryChild)	• Develop computable consent, with consistent language that allows for research use • Incorporate consents for clinical and research use into CCDI protocols
Baseline Data Collection: Collect more clinical data early, and identify high-value data elements for research (cohorts)	• Collection of additional data elements in CCDI data sets (MCI, others)	• Working group to identify key data elements • Collect these data as part of CCDI studies (Rare Tumor Protocol)

Next Steps (1)

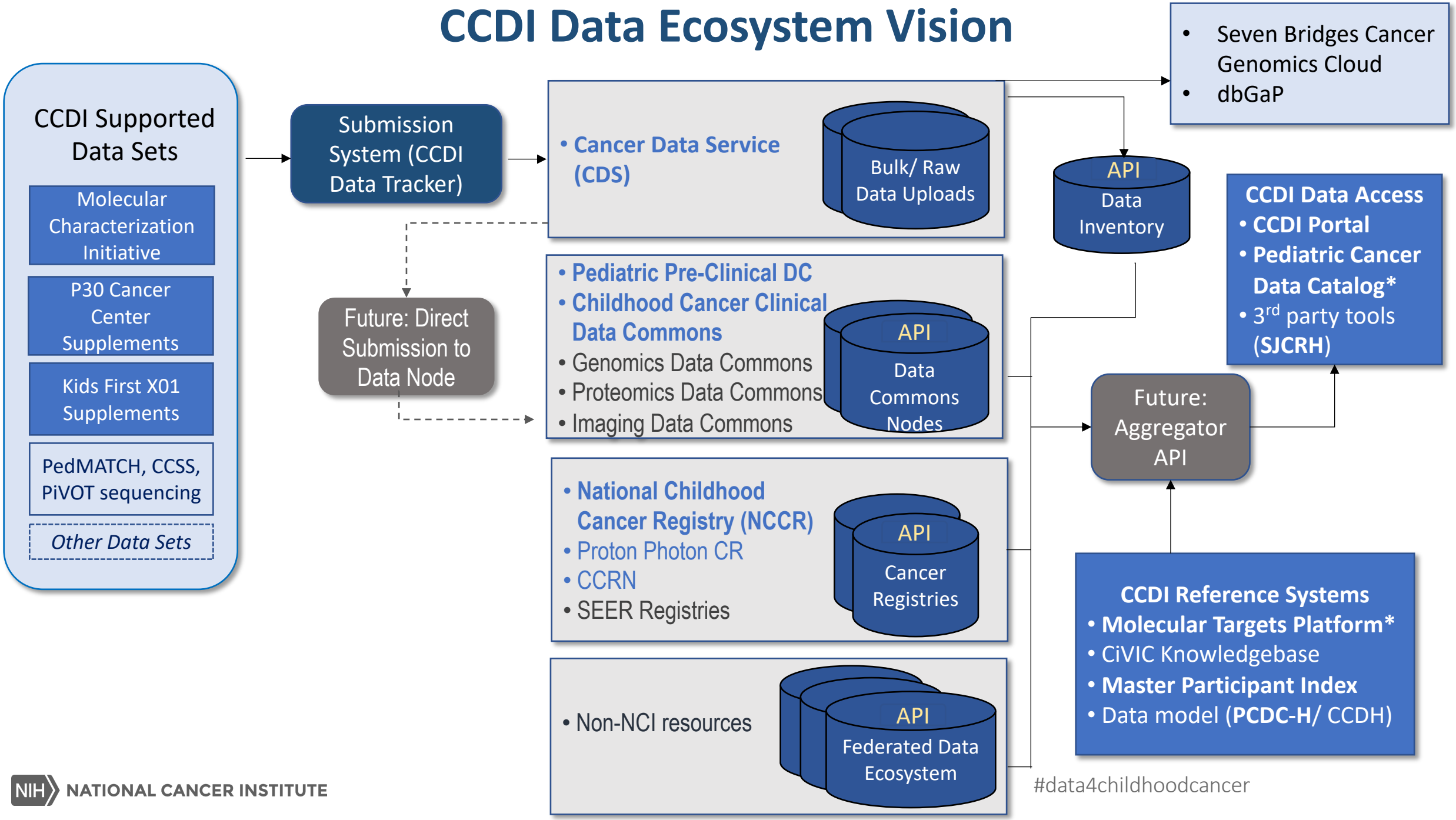
Build Data Infrastructure

Continue to expand data ecosystem capabilities with focus on tools and portal for access and broad use

Foundational Phase of CCDI (2020 – 2022) – Develop a framework of critical activities that will fill major areas of need in the pediatric research community and support future efforts

Discovery and Expansion phases of CCDI (2023 – 2026….2029) – Establish opportunities to expand foundational efforts to make them work well together and create feasibility studies in the wider community

CCDI Data Ecosystem Vision



Next Steps (2)

Building Data Infrastructure

Continue to expand data ecosystem capabilities with focus on tools and portal for access and broad use

Aggregate and Generate Data

Establish CCDI Ultra Rare Tumor Protocol that includes comprehensive clinical and molecular characterization, collected over time

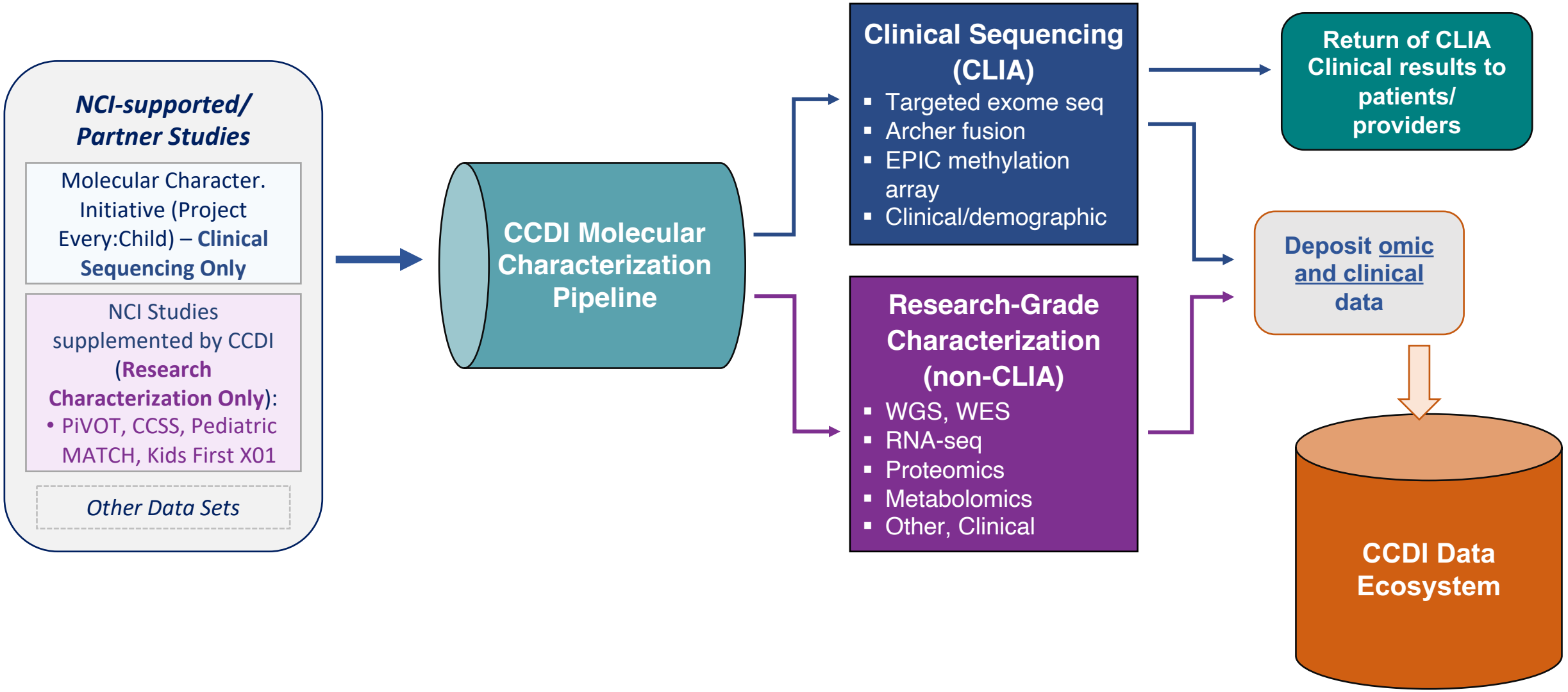
Foundational Phase of CCDI (2020 – 2022) – Develop a framework of critical activities that will fill major areas of need in the pediatric research community and support future efforts

Discovery and Expansion phases of CCDI (2023 – 2026….2029) – Establish opportunities to expand foundational efforts to make them work well together and create feasibility studies in the wider community

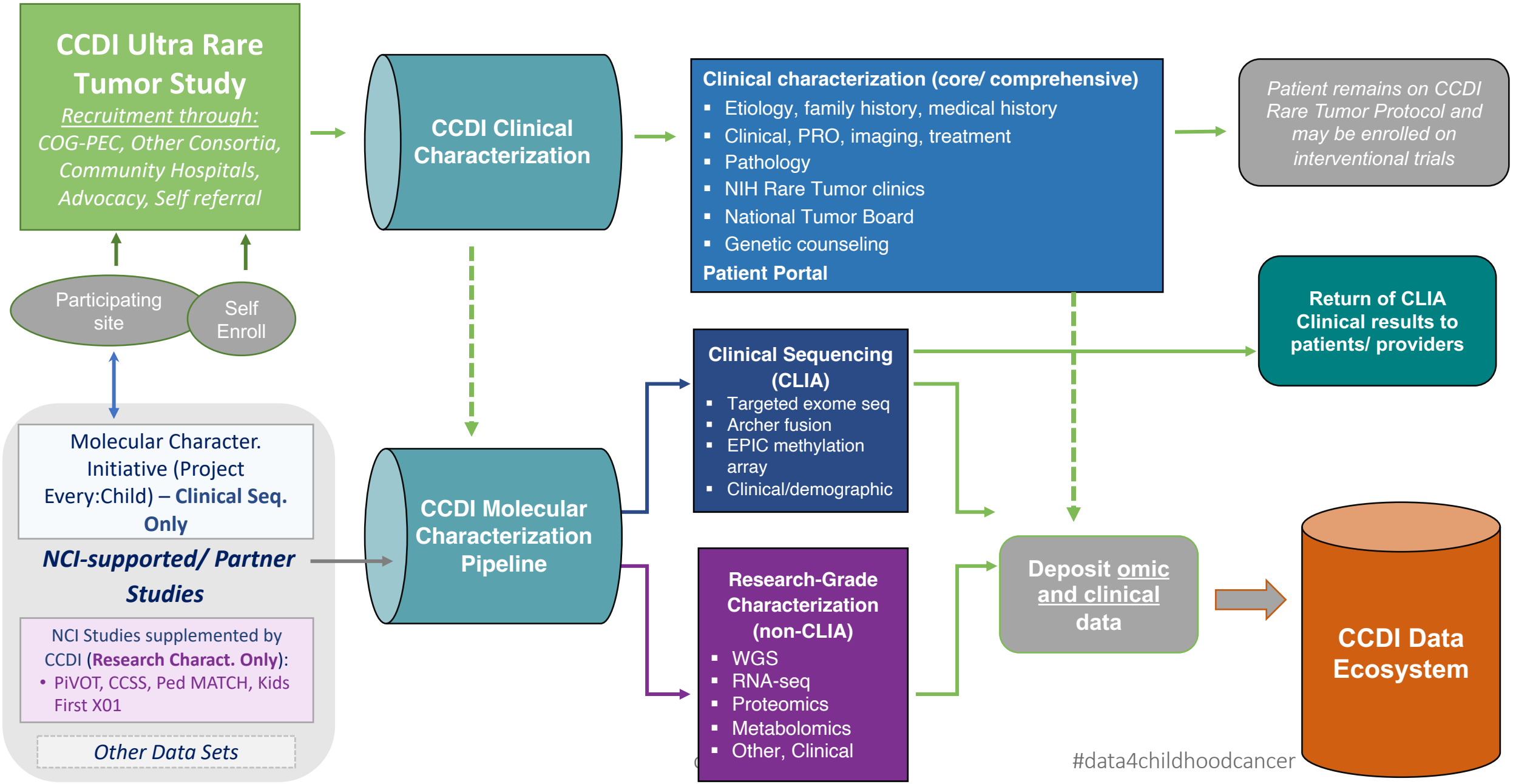
Workshop - Advancing a National Initiative for Rare Cancers in Children, Adolescents, and Young Adults - November 18, 2022

- BSA remit: Generate new data; focus on **populations of particular need** – rare disease, AYA patients, relapse/refractory disease
- Rare cancers – account for approx.15% of childhood cancer
- As a group, generally associated with poor outcomes
- Absent SOC; sparse clinical trials despite isolated institutional focused programs
- Broad support for an NCI-sponsored national strategy/trials – observational/interventional
- Need for patient support to navigate all aspects of the system – partnership with foundations and advocacy groups required
- Need for sustainable infrastructure to support collaborative efforts

CCDI Molecular Characterization Workflow



CCDI-Coordinated Rare Pediatric/AYA Tumor Study



#data4childhoodcancer

Next Steps (3)

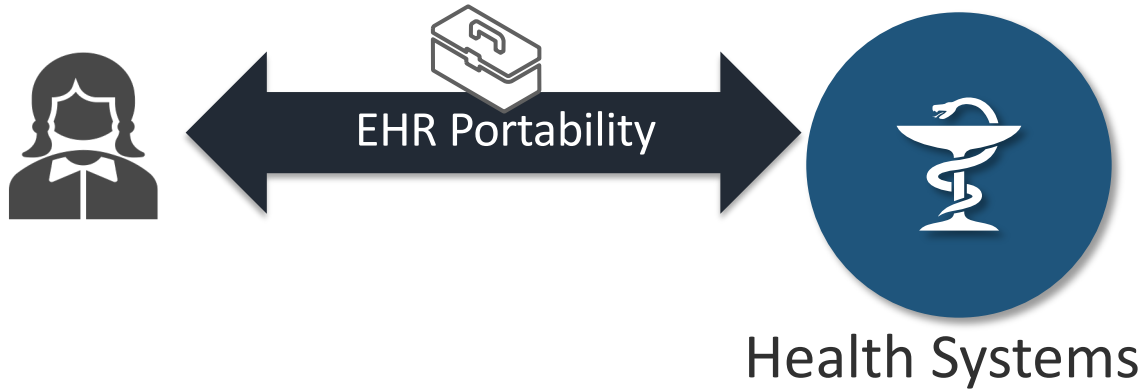
Building Data Infrastructure	Aggregate and Generate Data	Learn from and Use the Data
Continue to expand data ecosystem capabilities with focus on tools and portal for access and broad use	Establish CCDI Ultra Rare Tumor Protocol that includes comprehensive clinical and molecular characterization, collected over time	Develop consortia/network to oversee and explore a series of EHR extraction feasibility studies to support all types of research

Foundational Phase of CCDI (2020 – 2022) – Develop a framework of critical activities that will fill major areas of need in the pediatric research community and support future efforts

Discovery and Expansion phases of CCDI (2023 – 2026….2029) – Establish opportunities to expand foundational efforts to make them work well together and create feasibility studies in the wider community

Workshop: The Importance of Electronic Health Record Data in Clinical Care and Research - November 2, 2022

- Patients should be empowered to provide their consent for data transfer and research use
- Address challenges around lack of consistent data structures and standards, particularly within EHR systems (Epic, Cerner, etc.)
- Record needs to be widely useful for research, including for EHR-directed clinical trials, although data quality and consistency remain challenges to address
- Patients should have ways to enter, manage, and share their data, until improved technology supports automated data sharing



Health Systems

CCDI EHR Pilots
(*for Research Use*)

Structured EHR Data
& Translation

Surveillance
Public Health

Patient Participation
Clinical Trials

Discover
Basic Research

Next Steps

Building Data Infrastructure	Aggregate and Generate Data	Learn from and Use the Data
Continue to expand data ecosystem capabilities with focus on tools and portal for access and broad use	Establish CCDI Ultra Rare Tumor Protocol that includes comprehensive clinical and molecular characterization, collected over time	Develop consortia/network to oversee and explore a series of EHR extraction feasibility studies to support all types of research

Foundational Phase of CCDI (2020 – 2022) – Develop a framework of critical activities that will fill major areas of need in the pediatric research community and support future efforts

Discovery and Expansion phases of CCDI (2023 – 2026….2029) – Establish opportunities to expand foundational efforts to make them work well together and create feasibility studies in the wider community

Many Thanks to our CCDI Working Group & Support Members

Steering Committee
- Chairs: Dr. James Doroshow, Dr. Warren Kibbe*
- CCDI NCI Leadership Team
- CCDI WG Co-Chairs
- Mr. David Arons
- Dr. Doug Hawkins*
- Dr. Andrea Hayes-Jordan*
- Ms. Ellyn Miller
- Dr. Charlie Roberts

Cross-Cutting Issues
- Chairs: Drs. Jaime Guidry Auvil*, Sam Volchenboum*, Jinghui Zhang
- Dr. Ethan Cerami
- Dr. Sharon Diskin
- Dr. Casey Greene
- Ms. Julie Guillot*
- Dr. Melissa Haendel
- Ms. Amanda Haddock*
- Dr. Andrew Kung*
- Dr. Jill Mesirov
- Dr. Greg Reaman*
- Dr. Karlyne Reilly
- Dr. Adam Resnick
- Dr. David Siegel

Engagement Committee
- Chairs: Drs. Greg Aune, Lee Helman, Brigitte Widemann*
- Dr. Vickie Beunger
- Dr. Daphne Haas-Kogan
- Dr. D. Ashley Hill
- Dr. Beth Lawlor
- Dr. Troy McEachron
- Dr. Corrie Painter
- Dr. Abby Sandler
- Dr. Philip Lupo
- Ms. Sarah Milberg
- Dr. Alejandro Sweet-Cordero

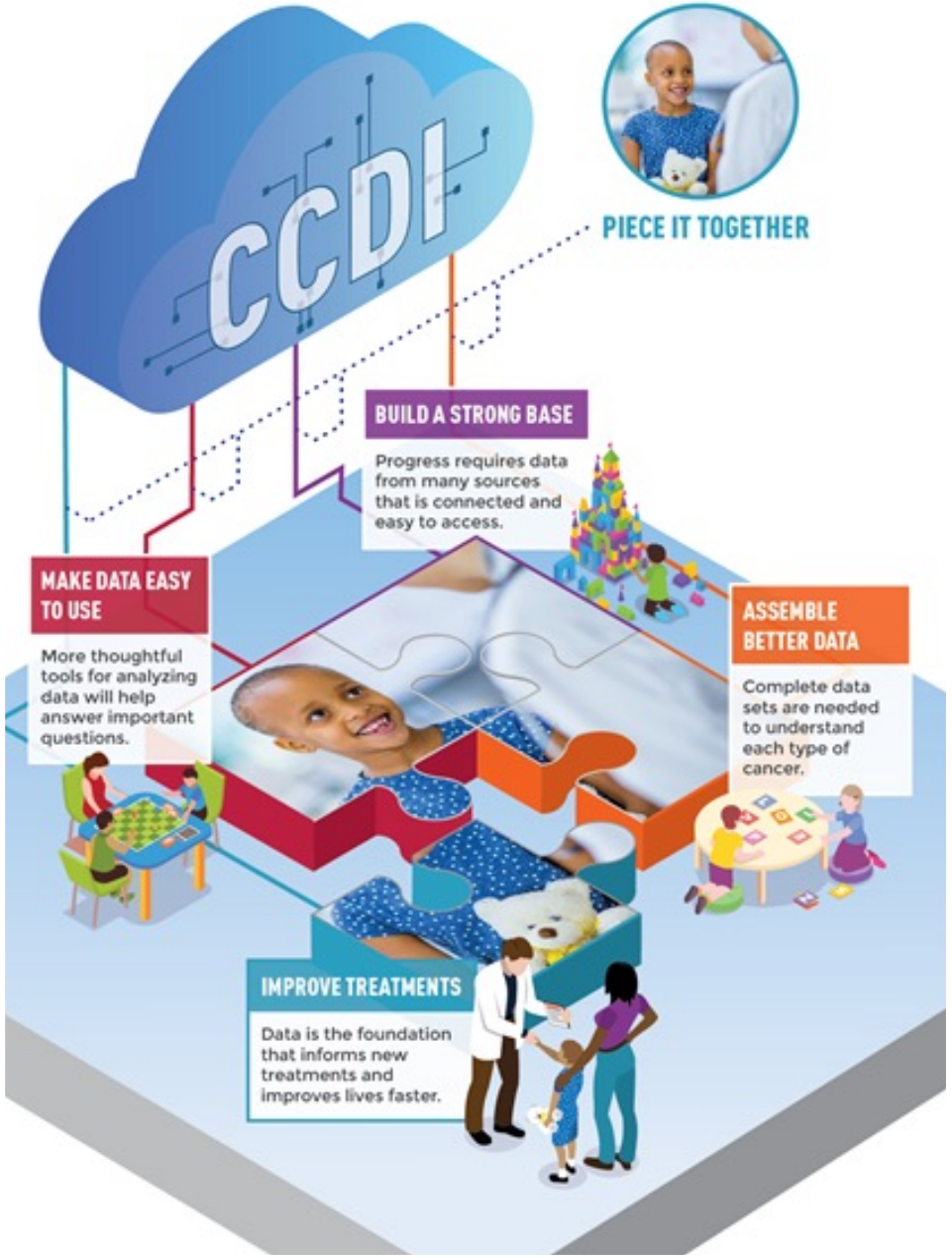
Data Platform
- Chairs: Drs. Allison Heath, Tony Kerlavage*, Jason Levine
- Ms. Naomi Bartley
- Dr. Tanja Davidsen
- Dr. Suzanne Forrest
- Ms. Johanna Goderre-Jones
- Dr. Subhashini Jagu
- Dr. Salvatore La Rosa
- Dr. Clay McLeod
- Dr. Shannon McWeeney
- Dr. Olena Vaske

Cancer Cohort
- Chairs: Drs. Greg Armstrong, Stephen Chanock*, Lynne Penberthy*
- Dr. Richard Aplenc
- Dr. Maria Monica Gramatges
- Dr. Amie Hwang
- Dr. Andy Olshan
- Ms. Aubrey Reirchard-Eline
- Dr. Logan Spector
- Dr. Nita Seibel
- Dr. Gail Tomlinson
- Dr. Emily Tonorezos
- Dr. Joe Flores-Toro
- Dr. Brent Weill
- Dr. Torunn Yock

BSA Ad Hoc Working Group
- CCDI WG members (*)
- Dr. Otis Brawley
- Dr. Kevin Shannon
- Dr. Peter Adamson
- Dr. Tom Curran
- Dr. Jim Downing
- Dr. John Maris

Molecular Characterization Protocol
- Chairs: Drs. Katie Janeway*, Jack Shern, Malcolm Smith*
- Dr. Caitlin Barrett
- Dr. Sue Cohn
- Dr. Carrye Cost
- Dr. Brian Crompton
- Dr. Maryam Fouladi
- Dr. Amar Gajjar
- Dr. Julie Glade-Bender
- Ms. Nancy Goodman
- Dr. Andy Kolbs
- Dr. Ted Laetsch
- Dr. Nilay Shah
- Dr. Jim Tricoli
- Dr. Mary Frances Wedekind

NCI Staff Supporting CCDI
- Ms. Anne Lubenow
- Mr. Richard Takamoto
- Ms. Eve Shalley
- Ms. Bonny Sheppard
- NCI Office of Advocacy Relations, Office of Communication & Public Liaison, and Office of Data Sharing



Questions?

Engage with CCDI

- Visit the CCDI website to learn more. on the NCI website
- Review the BSA CCDI Working Group final report
- Receive email updates from NCI on the CCDI
- Contact the CCDI with questions about engagement, ongoing activities, or funding opportunities