**Statement of Work**

Childhood Cancer Data Initiative: Multi-Omics Research Characterization Network

**Center for Biomedical Informatics and Information Technology**

# BACKGROUND

# The National Cancer Institute (NCI)’s Childhood Cancer Data Initiative (CCDI) focuses on the critical need to collect, analyze, and share data to address the burden of cancer in children, adolescents, and young adults (AYAs). The data initiative supports childhood cancer research and aims to facilitate cancer researchers to learn from each of the approximately 16,000 children and adolescents diagnosed with cancer in the United States each year.

The CCDI foundational goals are to:

* Gather data (existing and new) 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 research 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.

The CCDI aims to strengthen data generation and collection for childhood cancers, foster data utilization and sharing to advance progress in combating these cancers by ensuring the Findability, Accessibility, Interoperability, and Reusability (FAIR) Principles are upheld, and consequently incentivize the cancer research community to develop new treatments for children with cancer. In that regard, the generation of molecular data on children and AYA patients enrolled in cancer clinical trials are extremely valuable to realizing the promise of precision oncology. Through these clinical trials and protocols, pediatric and AYA cancer patients across the U.S. and around the world can access state-of-the-art therapies available through NCI-funded Pediatric Early Phase Clinical Trials Network (PEP-CTN), NCI Community Oncology Research Program (NCORP), My Pediatric and Adult Rare Tumor (MyPart) Network, Children’s Oncology Group (COG), and other studies. Specimens are routinely collected from patients enrolled across these clinical trials, presenting opportunities for molecular subtyping to aid in clinical decision-making, as well as for complementary research efforts to further the understanding of disease biology. Patient samples from children and AYAs with cancer on clinical trials and studies, however, are critically limited, posing challenges for research studies. Furthermore, data generated from these specimens in the past are often fragmented and not broadly available to researchers and oncology teams.

CCDI has a unique opportunity to generate and aggregate many sources of data to enhance data use for the broad research community. Since its launch, [The CCDI Molecular Characterization Initiative (MCI)](https://www.cancer.gov/research/areas/childhood/childhood-cancer-data-initiative/data-ecosystem/molecular-characterization) has contributed to and will continue to further the CCDI’s goals by generating and sharing clinical-grade next-generation sequencing (NGS) data through Clinical Laboratory Improvement Amendments (CLIA)-certified laboratories using specimens from pediatric and AYA patients enrolled to pediatric cancer focused trials and studies. These include but are not limited to NCI-supported clinical trials and studies conducted by the NCI’s Center for Cancer Research (CCR), PEP-CTN and COG for use in the conduct of the trials and related correlative studies. Other sources of specimens from clinical trials supported by pharmaceutical industry and foundations could also be leveraged for this endeavor to augment the cohort size. These studies have increasingly been utilized to identify gene mutations, genetic alterations, and changes in global gene expression that are associated with different tumor types and stages of disease.

Importantly, remaining specimens from clinical trials and protocols, whenever available and of sufficient quantity and quality, present an unprecedented opportunity to generate multi-omics characterization data including NGS, proteomics, and metabolomics on patients to complement clinical sequencing data (i.e., performed by clinical laboratories properly certified by Clinical Laboratory Improvement Amendments (CLIA) to receive Medicare or Medicaid payments). Molecular characterization of the residual samples using multi-omics approaches at the research level (non-CLIA-certified) can provide a more unified, comprehensive view of the underlying disease biology to facilitate researchers to learn from every child. Prior studies have demonstrated the new insights gleaned from cancer specimens that went through multidimensional characterization, such as identification of proteomic-centric subtypes, prioritization of driver mutations by correlative analysis of copy number alterations and protein abundance, understanding cancer-relevant pathways through posttranslational modifications (PTMs), and identification of characteristic metabolic phenotypes as biomarkers for treatment response. Thus, maximizing the use of childhood and AYA cancer specimens from clinical trials with additional characterization platforms with the resulting data made broadly available for research will have a profound impact on our ability to prevent, diagnose, and treat this patient population. This effort can serve as a model for improving the use of data across cancer types and other disease continuums.

## SCOPE

To accomplish the goals of this research characterization program complementary to clinical sequencing efforts of CCDI, the envisioned data generation network will be comprised of research laboratories that possess one or more molecular characterization expertise and are sufficiently agile to accommodate the characterization needs of CCDI amongst several network members (see diagram below).

The base contract will support the generation of childhood cancer multi-omics research data that will be delivered to the [CCDI data ecosystem](https://www.cancer.gov/research/areas/childhood/childhood-cancer-data-initiative/data-ecosystem) and/or NIH/NCI repositories. In performance of this SOW, the Contractor shall serve as Multi-omics Research Characterization Network (MRCN) members capable of providing high-throughput and quality-controlled analytical pipelines for comprehensive omics characterizations of different childhood cancer types using validated methods and platforms. The Contractor shall receive analytes such as DNA, RNA and proteins extracted or enriched from original specimens (e.g., tumor tissue, blood, saliva), and/or original specimens as well as their associated data from investigators/ institutions with access to these materials for additional molecular profiling that will advance the understanding of childhood cancer.

**Molecular Research Characterization Network Diagram**

Molecular Research Characterization Network Diagram

Process flow diagram that shows Specimens and Data (Clinical/Pathology), Analytes, Data Generation.

Under the process flow is an outer circle diagram that shows Genomics, Transcriptomics and Epigenetics, Metabolomics, and Proteomics. The inner circle shows Clinical Trials and Studies.

The molecular characterization for consideration may include the following activities at bulk sample, subpopulation of cells, and/or single-cell/single nuclei level:

* Whole Exome Sequencing (WES)
* Whole Genome Sequencing (WGS)
* Whole Genome Bisulfate Sequencing (WGBS)
* Methylation Arrays
* Transcriptome Sequencing (e.g., total RNA-seq and Nanostring)
* Additional molecular assays (e.g., miRNA analysis, circulating tumor DNA analysis, targeted DNA/RNA sequencing panels)
* Global Proteomics and Phosphoproteomics, or additional PTM profiling
* Metabolomics Profiling
* Single-Cell DNA/RNA/protein analysis

Specimens may include fresh frozen tumor biopsy or resected tumor tissue (e.g., curled or cryopulverized powder), formalin-fixed paraffin-embedded (FFPE) tissue blocks, laser-capture microdissection (LCM) tissue cells, bone marrow aspirate, bone marrow biopsy, whole blood, plasma, and serum, as well as extracted analytes from processed specimens such as DNA, RNA, protein and metabolite. In addition to characterization of tumor samples, CCDI may also need to support germline analysis from normal tissue and/or blood in selected settings.

These data and results will be deposited by the Contractor(s) into CCDI data ecosystem or NIH/NCI repositories, including multi-omics molecular characterization data, specimen data, clinical data, or other summary reports, depending on the specific needs of the clinical trials and studies.

NOTE: Although the exact number of specimens that will require characterization is not known, it is anticipated that this research network will be able to complete molecular characterization of up to 50 cases/participants per week in the base contract, with options for increased quantities in out years.

**SCIENTIFIC OBJECTIVES**

The overarching scientific objective of the parent contract is to obtain research-grade multi-omics characterization data on specimens from pediatric and AYA patients enrolled in applicable clinical trials and studies, including but not limited to NCI-supported clinical trials conducted by the NCI’s CCR, PEP-CTN, and COG for use in the conduct of the trials and related correlative science studies.

# REQUIREMENTS

The Contractor(s) shall have sequencing/characterization laboratories with the capacity and expertise to perform activities listed above. NCI will provide specifications for the characterization laboratories identified for each activity. Characterization laboratories shall be in place and ready to receive analytes or process specimens to analytes (DNA, RNA, proteins, metabolites) for sequencing using validated platforms within 6-8 months of NCI identifying a new molecular characterization need under this project. A robust platform is considered analytically validated if it has successfully been tested in at least one other laboratory, is capable of generating reproducible results within and across laboratories using standards and metrics in previous research studies and has been previously published in a peer-reviewed journal. The NCI-required research characterization specifications may change based on changing scientific needs, incremental increase needs in sample throughput in the out years, available characterization options, and validated novel assays coming online. As a result, agreements with the characterization laboratories should allow for potential changes to be expeditiously made when NCI specifications change.

## TASK AREAS

In order to meet the project objectives, the following task areas shall be performed:

**TASK AREA 1**: **High-throughput DNA and RNA sequencing of pediatric and AYA clinical trial specimens**

The Contractor shall:

1. Build up characterization pipelines of a minimum of 1,000 cases/participants using validated, high-throughput DNA and RNA sequencing platforms with prior data demonstrating equivalency to large-scale genomic research programs and widely adopted commercial off-the-shelf products.
2. Implement and report quality control/quality assurance (QA/QC) procedures to ensure and monitor the data quality from characterization pipelines.
3. Monitor, identify and take steps to mitigate sources of systematic error and bias to improve experimental reproducibility.
4. Deliver primary genomic, transcriptomic and epigenomic data, and derived data (e.g., fusion or somatic mutation calls) using the newest human reference genome as well as associated metadata to the CCDI data ecosystem or NIH/NCI repositories.

Protocol modifications shall be reviewed and pre-approved by the CCDI COR.

Sequencing platforms applicable to this project include:

*High-throughput DNA Sequencing*

DNA sequencing results can be used to identify novel cancer mutations, familial cancer mutation carriers, and provide clues for effective and targeted treatments.

Analyze DNA aliquots extracted from specimens using validated platforms and technologies for high-throughput DNA sequencing or DNA microarrays. The scale of DNA sequencing or microarray can range from targeted genes to whole exome or whole genome (depth and coverage requirements dependent on the specific needs of the project).

*High-throughput RNA Sequencing*

High-throughput RNA sequencing (RNA-seq) is a sequencing method for studying transcriptome and non-coding RNAs such as microRNAs (miRNAs). RNA-seq data can be used in studies such as whole transcriptome analysis of differential gene expression, differential splicing of mRNAs, and detection of gene fusions.

Analyze RNA aliquots extracted from specimens using validated platforms and technologies for high-throughput sequencing. Generate RNA sequencing data for cancer transcriptome and non-coding RNA expression profiles (with depth and coverage dependent on the specific needs of the project).

*Epigenomics with Array- or Sequencing-based Assays*

Epigenomics involves the profiling and analysis of epigenetic marks across the genome. The whole epigenome analysis can be carried out with sequencing-based assays.

Analyze specimens using validated platforms and technologies for high-throughput methylation sequencing or methylation arrays and generate molecular data for methylation profiles.

*Single-Cell/Single-Nuclei DNA/RNA/Methylation Sequencing*

Single-Cell/Single-Nuclei sequencing (sc/sn)-based assays have advanced significantly. Different types of sequencing technologies such as DNA-seq, RNA-seq, Methylation-seq, and ATAC-seq have been developed for single cell/single nuclei studies. These applications help researchers tackle problems such as tumor heterogeneity and cell population characterization with a high resolution unavailable with bulk sequencing methods.

Analyze specimens of tumor tissues or cells using approved technologies for high-throughput sc/sn sequencing. Generate molecular data for profiling the genomes, the epigenomes, or transcriptomes on the level of individual cells.

**TASK AREA 2**: **Comprehensive proteomic characterization of pediatric and AYA clinical trial specimens**

The Contractor shall:

1. Build up characterization pipelines to meet a minimum operational capacity of 300 samples per year using multiplexed, validated liquid chromatography mass spectrometry (LC-MS) platforms to characterize cancer specimens. At a minimum, the Contractor shall perform untargeted characterizations of global proteomes (unmodified) and phosphoproteomes. Additional post-translational modifications (PTMs) such as glycosylation, lysyl acetylation and ubiquitinylation, or single-cell proteomics will be considered with scientific justifications.
2. Analyze specimens or extracted proteins/peptides using validated technologies to identify and quantify unique proteins and phosphorylation sites from tumor samples in the forms of fresh frozen tissues, FFPE tissue blocks, or purified/isolated cells from blood or bone marrow samples using additional processing and enrichment procedures.
3. Implement a QA/QC procedure to monitor the performance of work and follow SOPs for each protocol.
4. Deliver primary and derived proteomic data using the newest human reference genome as well as associated metadata to the CCDI data ecosystem or NIH/NCI repositories.

Protocol modifications shall be reviewed and pre-approved by the CCDI COR.

*Comprehensive Global Proteome Profiling*

Comprehensively identify and quantify the unmodified protein composition (i.e., global proteomics) resulting from proteins or peptides extracted from specimens. The depth and coverage of proteomes will depend on the sample input and type of starting materials.

*Comprehensive Phosphoproteome Profiling*

Enrich for and comprehensively identify and quantify site-specific phosphoserine-, phosphothreonine and phosphotyrosine-containing proteins and peptides extracted from cancer specimens with relative quantification (i.e., phosphoproteomics). The depth and coverage of phosphoproteomes will vary dependent on the sample input and type of starting materials.

**Task AREA 3**: **Metabolomic profiling for pediatric and AYA clinical trial specimens**

The Contractor shall:

1. Comprehensively identify and measure (relative concentrations) of metabolites and/or metabolite features using validated LC-MS or GC-MS platforms in specimens such as tissue, serum, plasma, urine.
2. Build up metabolomic profiling capacity to analyze >300 samples per year.
3. Curate metabolomic data and identify a subset of metabolites using the newest reference spectral library for any given sample set.
4. Implement a QA/QC process to monitor the performance of work and shall follow SOPs for each protocol.
5. Deliver primary mass spectrometry data, metadata and derived data to the CCDI data ecosystem or NIH/NCI repositories.

## 4.0 PROJECT MANAGEMENT

## Relevant NCI programs and the Office of Data Sharing (ODS) will coordinate and provide oversight for overall MRCN activities including but are not limited to timeliness and quality of specimens, coordination of sample shipment, receipt, processing and molecular characterization, data management and sharing progress, data delivery to the CCDI data ecosystem supported by this program during the period of performance. The Contractor must include a data management and sharing (DMS) plan (Plan) or a statement that the Contractor will comply with CCDI data management and sharing expectations.

## 5.0 SCHEDULE

## The period of performance for the base contract is one (1) year, with options for increased quantities of samples for research characterization up to a total of five (5) years.