

CCSS Data and Statistics Center Report

CCSS Investigator Meeting, June 18, 2025

Deokumar Srivastava
Director, Data and Statistics Center

CCSS
Childhood Cancer
Survivor Study



An NCI-funded Resource

Outline

CCSS

- Team Members and Transitions
- Activities Overview
- Core Activity Updates
- Challenges and Priorities

Lead Biostatistician of Data & Statistics Center

ccss



2023-Present



2004-2023
20-years



1995-2004
10-years

Childhood Cancer
Survivor Study
An NCI-funded
resource

Thank you, Wendy! Her new role in CCSS

CCSS



Wendy Leisenring, ScD
Co-Chair Epidemiology/Biostat WG

2004-2023
20-years

Childhood Cancer
Survivor Study
An NCI-funded
resource

Lead Biostatistician as of May 1, 2023

ccss



**Deo Kumar Srivastava,
PhD
Lead CCSS Biostatistician**



Mingjuan Wang



Zhuo Qu



Grace Zhou



Nivya George

Analysts

**Childhood Cancer
Survivor Study**
An NCI-funded
resource

Staffing

Faculty

Analysts

CCSS



Sedigheh Mirzaei



Mengqi Xing



Shalini Bhatia



Kendrick Li



Lu Xie

Childhood Cancer
Survivor Study
An NCI-funded
resource

Staffing

Manager/Faculty

Analysts

CCSS



Yan Chen
Manager



Lei Wang



Fang Wang



Kerry Ye



Heather Xiang



Rik
Dhaduk



Yutaka Yasui



Qi Liu
Univ. of Alberta



Jessica Baedke
Univ. of Alberta

Staffing

Faculty

Analyst

CCSS



Wendy Leisenring
Co-chair Epidemiology/Biostat WG



Kayla Stratton

Statistical Center Core Activities

CCSS

- **Data Management**
 - Close collaboration with Coordinating Center and MD Anderson's RT Dosimetry Center (weekly meetings)
 - Cleaning data
 - Generating analytic data sets
 - Linkage with external data source
- **Concept Proposal and Ancillary Study Grant Development**
 - Design and write in collaboration with investigators ([involve us early!](#))
- **Statistical Analyses**
 - Collaborate on analyses via regular virtual meetings with investigators

Manuscript and Analyses Volume

ccss

January 1, 2020-Present*:	
Published Manuscripts:	141
With CCSS Biostatistician:	130
Currently Submitted Manuscripts:	22
With CCSS Biostatistician:	17
Manuscripts in Preparation:	27
Analyses Ongoing/Data sent:	26

Concept Proposals

CCSS

Status	N
Approved Concepts in Queue Waiting for Analyst:	0
Approved Concepts in Queue Waiting for Linkage/Data/DUA:	0

Frozen CCSS Analytic Database

ccss

Treatment for Primary Cancer (to 5 years post dx)

- Chemotherapy abstracted from medical records, cumulative dose
- Radiation Dosimetry to body regions and select organs (MD Anderson)
- Surgery Operative notes coded as ICD9

Survey Data

- Survey Response data through FU7
- Subsequent Neoplasms reported and confirmed through FU7
- Chronic Health Conditions coded through FU7

Mortality

- Vital status and cause of death through 2021 (National Death Index linkage)

Cloud Data Repositories

CCSS

dbGaP Data Repository:

- GWAS data for n=5,912 original cohort survivors
- Whole Exome Sequencing for n=5,451 original cohort survivors
- CCSS phenotype data for n=25,665 survivors; N=5,051 siblings

Promoting the CCSS Resource
CCSS on St. Jude Survivor Portal
Dr. Yasui 5.00 – 5.30 pm

St Jude Cloud

- CCSS Expansion cohort subjects with Whole Genome Sequencing data (N=2,641)
- Online tools at St Jude Cloud available (<http://survivorship.stjude.cloud/>)
- CCSS Phenotype data for these 2,641 survivors

Phenotype Data on clouds includes:

Characteristics	Treatment Exposures	Outcome Measures
Demographics	Chemotherapy (Y/N and Doses), HCT	Mortality
Primary Cancer DX	Radiation (body region Y/N and Doses)	Subsequent Neoplasms /recurrences
	Surgeries (summary, icd9code/label)	Chronic conditions overall and by organ system (max grade, age)

External Data Linkages with CCSS cohort

ccss

Completed

- Society for Assisted Reproductive Technology (SART)
 - Identified 137 survivors / 33 siblings who used ART
- Medicaid Data
 - CCSS participants alive and aged 18-64 years as of Jan 1, 2010
 - Matched to 2009-2016 Medicaid claims data
- Center for International Blood and Marrow Transplant Research (CIBMTR)
 - Identified 1018 matched stem cell transplant recipients
 - Provides transplant characteristics (donor, conditioning, etc.) and graft vs host disease

All pose challenges
Update these resources

Ongoing

- Virtual Pooled Registry (VPR) for SMN identification
 - Piloted with 26 states' linkage of cancer registries to CCSS
 - 43 states total available and in process

Progress Made Since Data Center Move

May 1, 2023

CCSS

Completed the FU7 Freeze

- Subsequent Neoplasms reported and confirmed
- Updated Chronic Health Conditions
- Moved RT data to St. Jude
- Uploaded CCSS data to St. Jude Cloud

Mortality

- Vital status and cause of death through 2021 (National Death Index linkage)

Converted Questionnaire Data from PDF format to EXCEL Files

- From Baseline – FU8 (Searchable, accessible)

CCSS Longitudinal Questionnaire Dictionary

ccss

SECTION	SECTION NA	QUESTION TOPIC	QUESTION	VARIABLE in SAS	DATASET in SAS	NOTE FOR VARIABLES	FORMAT in SAS	CATEGORIES in QUESTIONNAIRES	NOTE FOR CATEGORIES in SAS	LABELS in SAS
C	Medical Tests	c1.g	An ultrasound of the thyroid gland?	thyroid_us	fu7c		MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1g Last thyroid ultrasoun
C	Medical Tests	c1.h	An ultrasound of the carotid arteries (blood vessels in the neck)?	carotid_us	fu7c		MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1h Last carotid artery ult
C	Medical Tests	c1.i	A skin exam for skin cancer by a healthcare provider?	skinexam	fu7c		MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1i Last skin exam for skin
C	Medical Tests	c1.j	A mammogram?	mammo	fu7c	For females	MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1j Last mammogram
C	Medical Tests	c1.k	A breast ultrasound?	breast_us	fu7c	For females	MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1k Last breast ultrasound
C	Medical Tests	c1.l	A breast MRI?	breast_mri	fu7c	For females	MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1l Last breast MRI
C	Medical Tests	c1.m	A pap smear?	lpapsmr	fu7c	For females	MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1m Last pap smear
C	Medical Tests	c1.n	A PSA or blood test to detect prostate cancer?	psa	fu7c	For males	MEDTESTF.	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	1 - Never, 2 - Less than 1 year ago, 3 - 1-2 ye	C1n Last PSA test
C	Medical Tests	c2	Please indicate all medicines/drugs you took regularly during the PAST 2 YEARS. We are only asking about medicines/drugs that you took consistently for more than one month, or for 30 days or more in a year. Please list only drugs prescribed by a doctor and							
C	Medical Tests	c2.1	Please specify the name of each prescription medicine/drug1	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.1	Please specify the name of each prescription medicine/drug1_age	fu7c		Age at first use	6			C2 Prescription drug 1 age
C	Medical Tests	c2.1	Please specify the name of each prescription medicine/drug1_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.2	Please specify the name of each prescription medicine/drug2	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.2	Please specify the name of each prescription medicine/drug2_age	fu7c		Age at first use	6			C2 Prescription drug 2 age
C	Medical Tests	c2.2	Please specify the name of each prescription medicine/drug2_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.3	Please specify the name of each prescription medicine/drug3	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.3	Please specify the name of each prescription medicine/drug3_age	fu7c		Age at first use	6			C2 Prescription drug 3 age
C	Medical Tests	c2.3	Please specify the name of each prescription medicine/drug3_yr	fu7c		Are you currently taking?	6	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.4	Please specify the name of each prescription medicine/drug4	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.4	Please specify the name of each prescription medicine/drug4_age	fu7c		Age at first use	6			C2 Prescription drug 4 age
C	Medical Tests	c2.4	Please specify the name of each prescription medicine/drug4_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.5	Please specify the name of each prescription medicine/drug5	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.5	Please specify the name of each prescription medicine/drug5_age	fu7c		Age at first use	6			C2 Prescription drug 5 age
C	Medical Tests	c2.5	Please specify the name of each prescription medicine/drug5_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.6	Please specify the name of each prescription medicine/drug6	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.6	Please specify the name of each prescription medicine/drug6_age	fu7c		Age at first use	6			C2 Prescription drug 6 age
C	Medical Tests	c2.6	Please specify the name of each prescription medicine/drug6_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.7	Please specify the name of each prescription medicine/drug7	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.7	Please specify the name of each prescription medicine/drug7_age	fu7c		Age at first use	6			C2 Prescription drug 7 age
D	Medical Tests	c2.7	Please specify the name of each prescription medicine/drug7_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
D	Medical Tests	c2.8	Please specify the name of each prescription medicine/drug8	fu7c			\$4.00			C2 Prescription drug name
D	Medical Tests	c2.8	Please specify the name of each prescription medicine/drug8_age	fu7c		Age at first use	6			C2 Prescription drug 8 age
D	Medical Tests	c2.8	Please specify the name of each prescription medicine/drug8_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
D	Medical Tests	c2.9	Please specify the name of each prescription medicine/drug9	fu7c			\$4.00			C2 Prescription drug name
D	Medical Tests	c2.9	Please specify the name of each prescription medicine/drug9_age	fu7c		Age at first use	6			C2 Prescription drug 9 age
D	Medical Tests	c2.9	Please specify the name of each prescription medicine/drug9_yr	fu7c		Are you currently taking?	YESNOF.	1: Yes, 2: No	1: Yes, 2: No; 3: Not sure	C2 Currently taking prescr
C	Medical Tests	c2.10	Please specify the name of each prescription medicine/drug10	fu7c			\$4.00			C2 Prescription drug name
C	Medical Tests	c2.10	Please specify the name of each prescription medicine/drug10_age	fu7c		Age at first use	YESNOF.			C2 Prescription drug 10 age

CCSS Longitudinal Questionnaire Dictionary

Benefits

CCSS

- Eliminates Multi-Tool hunting
 - Single Interface replaces scattered PDFs + SAS datasets
- Consistency across Questionnaires
 - Baseline → FU8 variables and formats in one view
- Instant Cross-reference
 - Question → SAS name → Labels without switching tools
- Universal Standardization
 - One source of information for all teams

Challenges

CCSS

Methodological:

- **Drop-out concerns:** Bias due to non-participation / Loss to follow-up
 - **Solution:** Continue to advocate using inverse probability weighting (IPW) to adjust all analyses for likelihood of dropout.
 - Conduct sensitivity analysis.
- **Challenge:** Have applied weights to recent analyses
 - Overly complicates some analyses and can be time consuming to implement.
 - Still advocating, but thus far, bias is minimal. Be flexible and vigilant.
- **Self-reported Outcomes:** Potential recall/reporting bias
 - **Validation data in subset:** Could improve analyses?

Priorities

CCSS

- **Analyses:** Maintain productivity, quality and consistency on analytic projects. Continue to evaluate and account for drop out.
- **Data:** Many incoming data sources for us to clean, incorporate and analyze. Moved this process to St Jude and getting ready to freeze FU8 data.
- **Study Designs:** Support core and ancillary studies with innovative and rigorous study designs and analyses.
- **Grant Renewal:** Working with others, we are developing plans for cohort expansion, aimed at studying late effects of modern therapies (from 2000-2023) and accelerated aging.
 - Innovations in data collection, validation.