Using NCCR*Explorer and NCCR Data in SEER*Stat

Childhood Cancer Data Initiative Webinar Series

NIH NATIONAL CANCER INSTITUTE

October 24, 2023

1. *NCCR*Explorer Overview*
2. *NCCR*Explorer Demo*
3. *NCCR Data in SEER*Stat*
4. *NCCR in SEER*Stat Demo*
5. *Q&A*

NCCR*Explorer Overview

Session Goals

- Provide a comprehensive overview of the CCDI National Childhood Cancer Registry Explorer (NCCR*Explorer), highlighting its features and functionalities
- Provide a brief overview of the NCCR data in SEER*Stat and the SEER*Stat tool, made available by NCI's Surveillance, Epidemiology, and End Results Program (SEER)
- Demonstrate practical applications of these tools, showcasing their effectiveness in addressing childhood cancer statistics questions
- Describe the differences, limitations, and capabilities inherent in both tools
- Share insights into the plans for the releases of both tools

Note: *This webinar does not provide detailed SEER*Stat training

SEER*Stat Resources

- YouTube video (youtube.com/watch?v=q6ywNxoJa7s)
- Recorded SEER*Stat past presentations (seer.cancer.gov/news/seerstat-webinars.html)
  - SEER and SEER*Stat Overview (October 14, 2021)
  - Survival (October 28, 2021)
  - SEER Variables (November 2, 2021)
- SEER*Stat tutorials (seer.cancer.gov/seerstat/tutorials/)

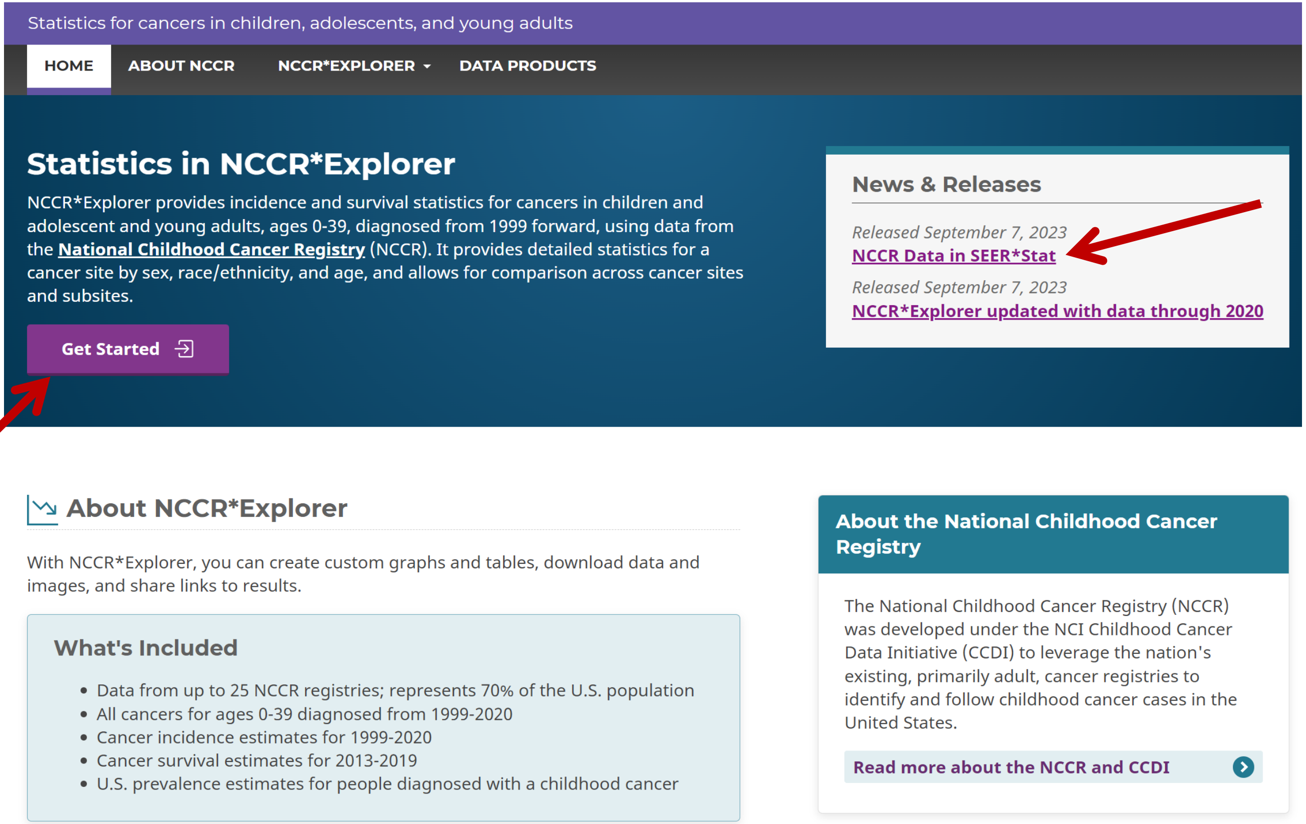
NCCR*Explorer Overview

- Includes **pre-calculated statistics** by cancer site, age, sex, and race
- Includes confidence intervals for all statistics, except for prevalence
- Adaptable, user-centric, and adept at delivering rapid responses through dynamic graphical representations and intuitive tables
- Figures and tables can be downloaded to be incorporated in documents

NATIONAL CANCER INSTITUTE
Childhood Cancer Data Initiative
National Childhood Cancer Registry Explorer

nccrexplorer.ccdi.cancer.gov

Google
NCCR*Explorer



Statistics in NCCR*Explorer

NCCR*Explorer provides incidence and survival statistics for cancers in children and adolescent and young adults, ages 0-39, diagnosed from 1999 forward, using data from the **National Childhood Cancer Registry** (NCCR). It provides detailed statistics for a cancer site by sex, race/ethnicity, and age, and allows for comparison across cancer sites and subsites.

News & Releases

Released September 7, 2023
NCCR Data in SEER*Stat

Released September 7, 2023
NCCR*Explorer updated with data through 2020

About NCCR*Explorer

With NCCR*Explorer, you can create custom graphs and tables, download data and images, and share links to results.

What's Included

- Data from up to 25 NCCR registries; represents 70% of the U.S. population
- All cancers for ages 0-39 diagnosed from 1999-2020
- Cancer incidence estimates for 1999-2020
- Cancer survival estimates for 2013-2019
- U.S. prevalence estimates for people diagnosed with a childhood cancer

About the National Childhood Cancer Registry

The National Childhood Cancer Registry (NCCR) was developed under the NCI Childhood Cancer Data Initiative (CCDI) to leverage the nation's existing, primarily adult, cancer registries to identify and follow childhood cancer cases in the United States.

Read more about the NCCR and CCDI

Which Data Are Included in NCCR*Explorer?

- All cancers for ages 0-39 diagnosed from 1999-2020 from up to 25 NCCR registries representing 70% of the U.S. population
- Cancer sites are defined using the International Classification of Childhood Cancer (ICCC), 3rd edition, 2017
- Categorizes childhood cancer based on the type of tissue affected, body location, as well as similar clinical management and outcomes

Statistics Included

Incidence rates (number of new cases per 1,000,000 of people at risk per year)

- Trends over time, 1999-2020
- Recent rates, 2016-2020
- Rates by single age at diagnosis

5-year survival: Measured as relative survival (but for this age group, very similar to all causes survival)

Complete prevalence: Number of people alive at 1/1/2020 ever diagnosed with childhood (0-19) cancer

NCCR*Explorer Demo

NCCR*Explorer Demo: Acute Lymphocytic Leukemia Diagnosed at Ages 0-19

- Q1. Is the incidence of lymphoid leukemias, aka acute lymphocytic leukemia (ALL), increasing or decreasing?
- Q2. Is ALL more common in boys or girls?
- Q3. Which race/ethnicity has the higher incidence rates?
- Q4. At what ages is the diagnosis of ALL more common?
- Q5. What is the 5-year survival for patients diagnosed with ALL? Does it vary by age at diagnosis?
- Q6. Which is the most common type of leukemia?
- Q7. How many people are living with a diagnosis of ALL at 1/1/2020?

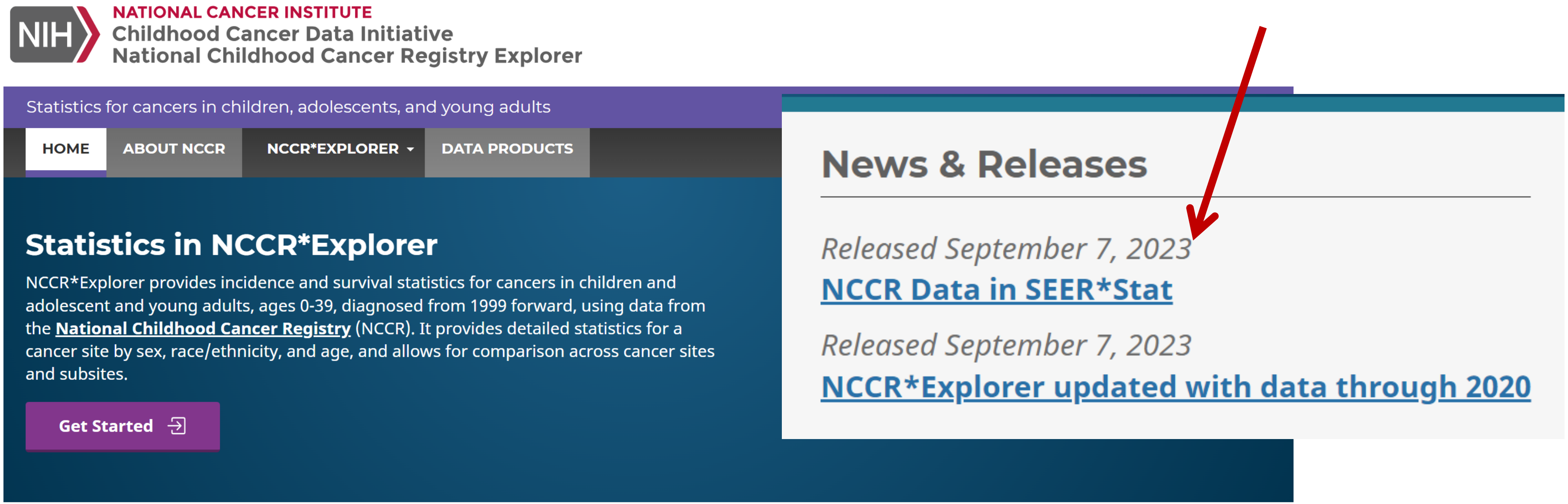
Limitations

- Stage, other tumor characteristics are not included
- Treatment information is not included
- Only available for a pre-defined list of ~70 cancer types and pre-calculated statistics
- Number of cases is not available for confidentiality issues

NCCR Data in SEER*Stat

How to Access NCCR Databases in SEER*Stat Software

nccrexplorer.ccdi.cancer.gov



NATIONAL CANCER INSTITUTE
Childhood Cancer Data Initiative
National Childhood Cancer Registry Explorer

Statistics for cancers in children, adolescents, and young adults

Statistics in NCCR*Explorer

NCCR*Explorer provides incidence and survival statistics for cancers in children and adolescent and young adults, ages 0-39, diagnosed from 1999 forward, using data from the **National Childhood Cancer Registry** (NCCR). It provides detailed statistics for a cancer site by sex, race/ethnicity, and age, and allows for comparison across cancer sites and subsites.

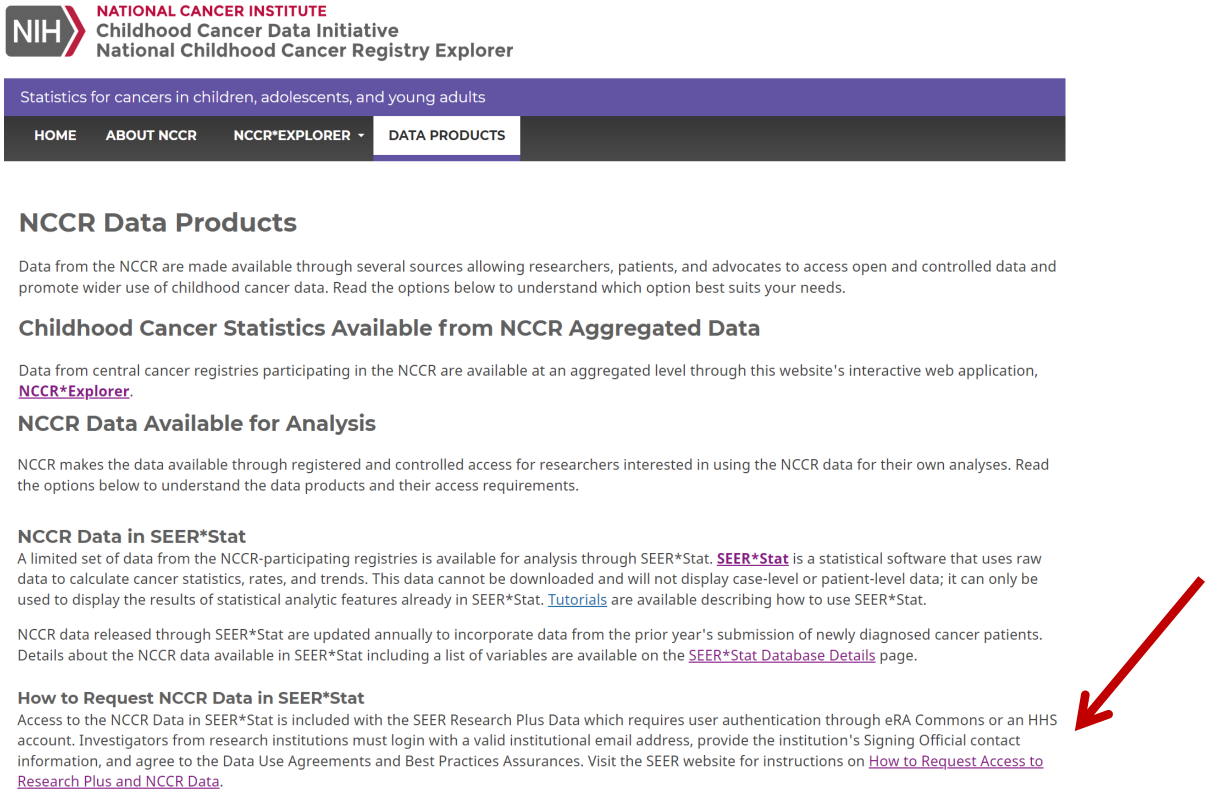
News & Releases

Released September 7, 2023
NCCR Data in SEER*Stat

Released September 7, 2023
NCCR*Explorer updated with data through 2020

Data Products

nccrexplorer.ccdi.cancer.gov/data-products.html



NCCR Data Products

Data from the NCCR are made available through several sources allowing researchers, patients, and advocates to access open and controlled data and promote wider use of childhood cancer data. Read the options below to understand which option best suits your needs.

Childhood Cancer Statistics Available from NCCR Aggregated Data

Data from central cancer registries participating in the NCCR are available at an aggregated level through this website's interactive web application, **NCCR*Explorer**.

NCCR Data Available for Analysis

NCCR makes the data available through registered and controlled access for researchers interested in using the NCCR data for their own analyses. Read the options below to understand the data products and their access requirements.

NCCR Data in SEER*Stat

A limited set of data from the NCCR-participating registries is available for analysis through SEER*Stat. **SEER*Stat** is a statistical software that uses raw data to calculate cancer statistics, rates, and trends. This data cannot be downloaded and will not display case-level or patient-level data; it can only be used to display the results of statistical analytic features already in SEER*Stat. Tutorials are available describing how to use SEER*Stat.

NCCR data released through SEER*Stat are updated annually to incorporate data from the prior year's submission of newly diagnosed cancer patients. Details about the NCCR data available in SEER*Stat including a list of variables are available on the SEER*Stat Database Details page.

How to Request NCCR Data in SEER*Stat

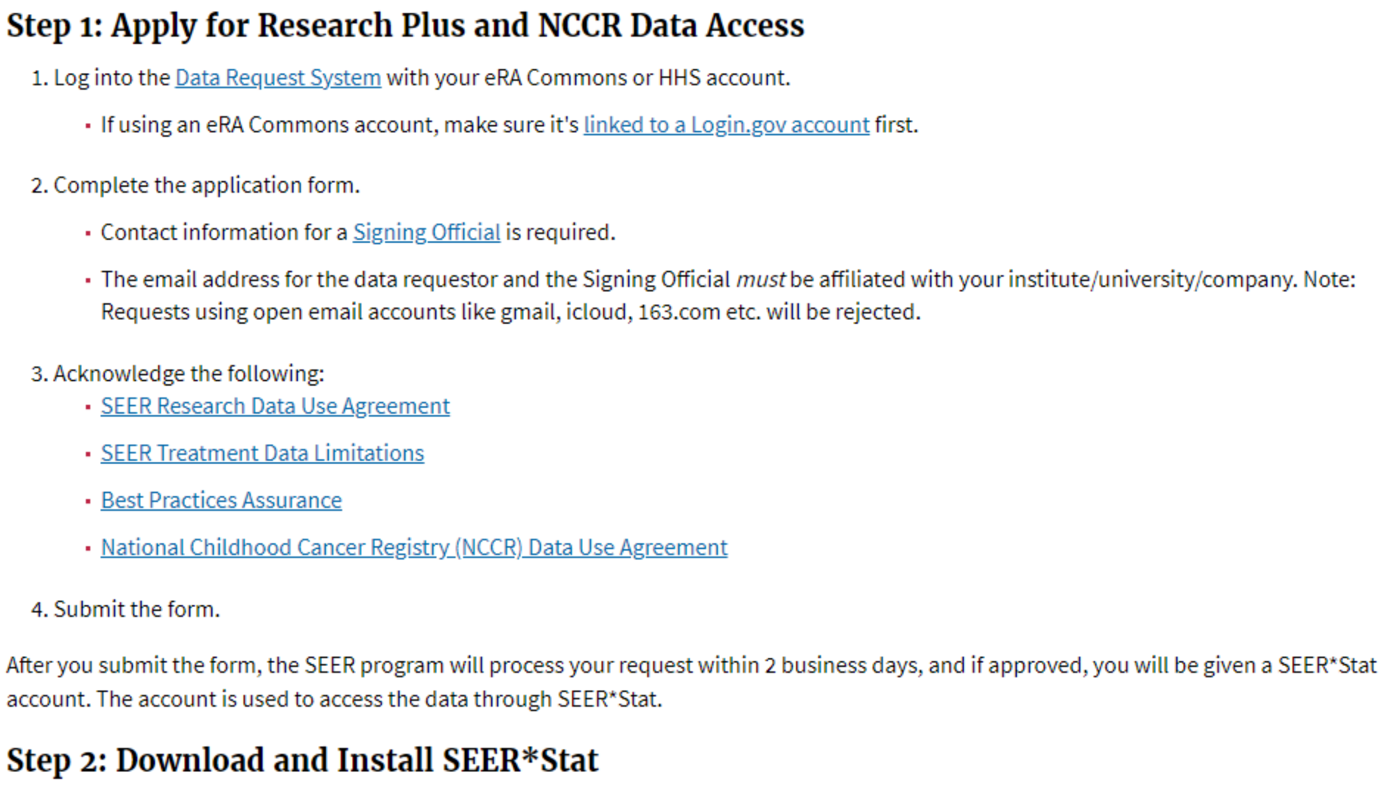
Access to the NCCR Data in SEER*Stat is included with the SEER Research Plus Data which requires user authentication through eRA Commons or an HHS account. Investigators from research institutions must login with a valid institutional email address, provide the institution's Signing Official contact information, and agree to the Data Use Agreements and Best Practices Assurances. Visit the SEER website for instructions on How to Request Access to Research Plus and NCCR Data.

How to Request Access to SEER and NCCR Data

- The NCCR databases are already available to SEER*Stat users with access to Research Plus SEER data
- NCCR data require user authentication with an eRA Commons or an HHS account
- Refer to the FAQs for instructions if you do not yet have an eRA Commons account. seer.cancer.gov/data/faqs.html

How to Request Access to SEER Data

seer.cancer.gov/data/access.html



Step 1: Apply for Research Plus and NCCR Data Access

1. Log into the Data Request System with your eRA Commons or HHS account.
    - If using an eRA Commons account, make sure it's linked to a Login.gov account first.
2. Complete the application form.
    - Contact information for a Signing Official is required.
    - The email address for the data requestor and the Signing Official *must* be affiliated with your institute/university/company. Note: Requests using open email accounts like gmail, icloud, 163.com etc. will be rejected.
3. Acknowledge the following:
    - SEER Research Data Use Agreement
    - SEER Treatment Data Limitations
    - Best Practices Assurance
    - National Childhood Cancer Registry (NCCR) Data Use Agreement
4. Submit the form.

After you submit the form, the SEER program will process your request within 2 business days, and if approved, you will be given a SEER*Stat account. The account is used to access the data through SEER*Stat.

Step 2: Download and Install SEER*Stat

SEER*Stat Overview

- SEER*Stat is an analytical software that allow users to perform a wide range of analyses with some flexibility in defining parameters
- Statistics for the NCCR data include:
  - Frequencies → counts
  - Incidence rates → counts/populations (default is cases per 100,000, but user can change the default to cases per 1,000,000)
  - Survival → 5 types of survival shown as percent or proportions
- SEER*Stat results are in a table format with spreadsheet features
  - View
  - Print
  - Export to text file
  - Copy and paste to other software (e.g. Excel for graphics)

NCCR Data in SEER*Stat

- In this “demonstration release,” there is a limited number of registries, including only those that provided permission (up to 15 states + Seattle)
- **The next release in Spring 2024 will include a larger number of registries**

NCCR Data in SEER*Stat

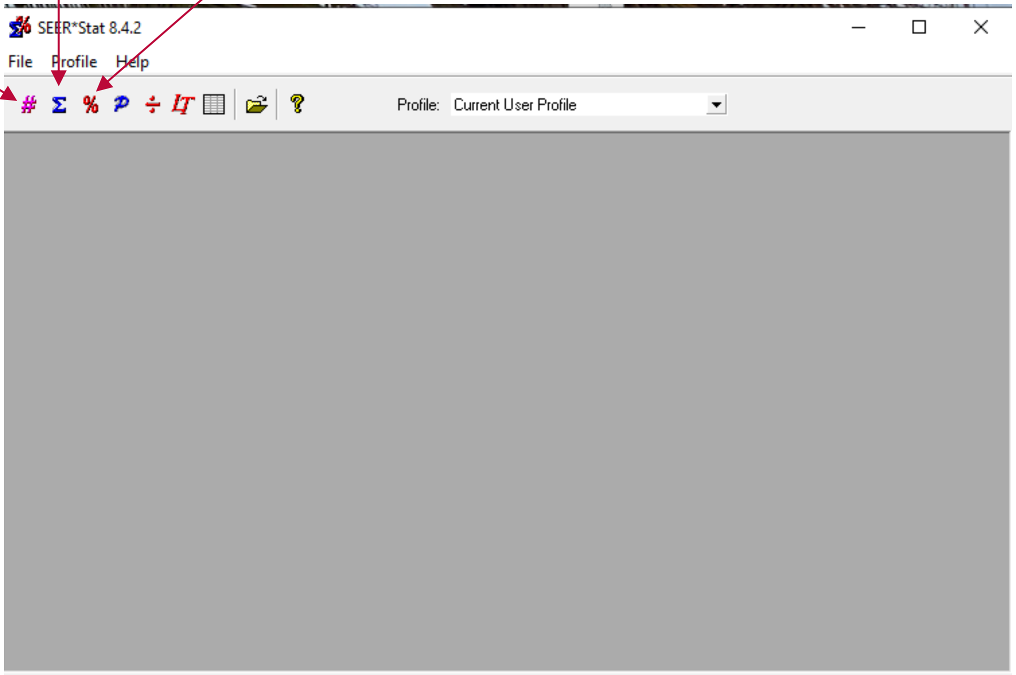
- NCCR data includes over 70 variables on demography, tumor characteristics, initial treatment, and survival information
- Only high-quality variables that were consistently collected during the same years across all the registries are included in the data set
- Not all the variables are available for all the years:
  - Variable name rule: Summary stage (2004+)
  - Years within parentheses indicated calendar years the variable is available
  - A year not in parentheses indicates the version of the variable: ICCC site recode 3rd edition/IARC 2017

NCCR in SEER*Stat Demo

NCCR in SEER*Stat Demo

- What is the number of cases diagnosed between 2016-2020 at ages 0-19 by detailed leukemia subtypes?
- Q2. What is the long-term survival (10 years) for patients diagnosed with leukemia at ages 0-19 between 2009-2019 by subtype

- Create a new session
- Select a database
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results

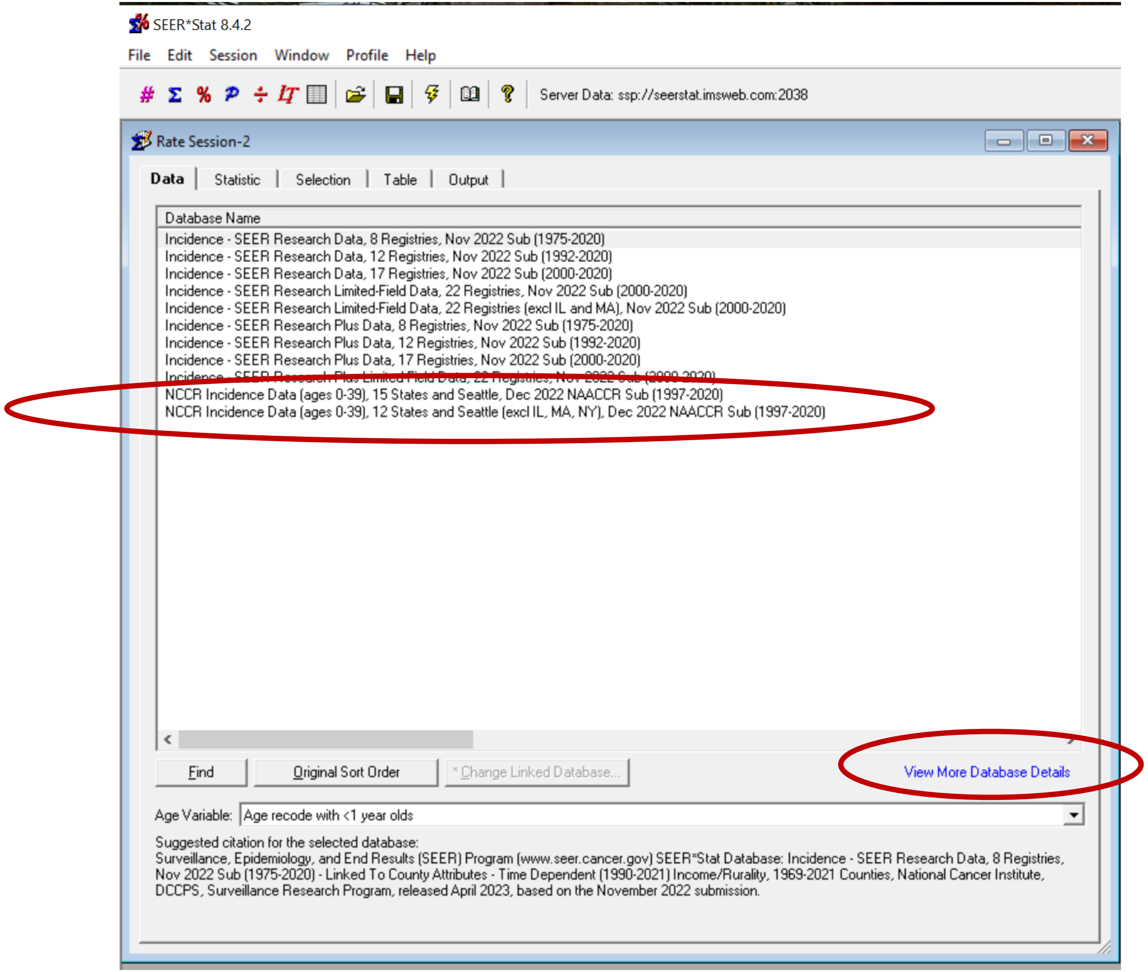


Frequencies (no. of cases)

Rates (Cases/pop)

Survival

- Create a new session
- **Select a database**
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results

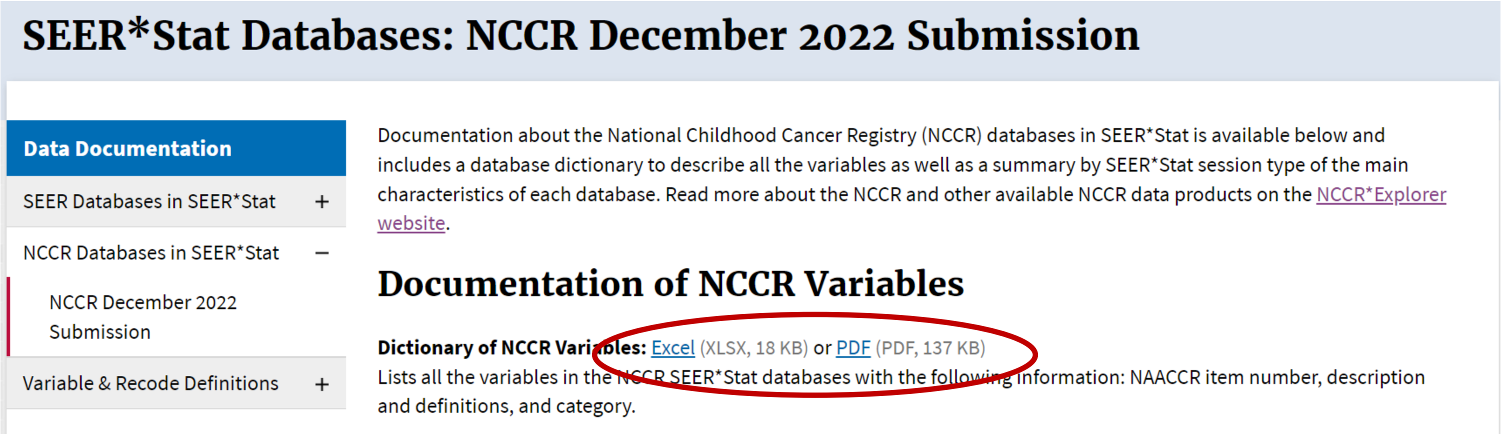


Database Name

Incidence - SEER Research Data, 8 Registries, Nov 2022 Sub (1975-2020)
Incidence - SEER Research Data, 12 Registries, Nov 2022 Sub (1992-2020)
Incidence - SEER Research Data, 17 Registries, Nov 2022 Sub (2000-2020)
Incidence - SEER Research Limited-Field Data, 22 Registries, Nov 2022 Sub (2000-2020)
Incidence - SEER Research Limited-Field Data, 22 Registries (excl IL and MA), Nov 2022 Sub (2000-2020)
Incidence - SEER Research Plus Data, 8 Registries, Nov 2022 Sub (1975-2020)
Incidence - SEER Research Plus Data, 12 Registries, Nov 2022 Sub (1992-2020)
Incidence - SEER Research Plus Data, 17 Registries, Nov 2022 Sub (2000-2020)
Incidence - SEER Research Plus Limited-Field Data, 22 Registries, Nov 2022 Sub (2000-2020)
NCCR Incidence Data (ages 0-39), 15 States and Seattle, Dec 2022 NAACCR Sub (1997-2020)
NCCR Incidence Data (ages 0-39), 12 States and Seattle (excl IL, MA, NY), Dec 2022 NAACCR Sub (1997-2020)

Age Variable: Age recode with <1 year olds

Suggested citation for the selected database:
Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER Research Data, 8 Registries, Nov 2022 Sub (1975-2020) - Linked To County Attributes - Time Dependent (1990-2021) Income/Rurality, 1969-2021 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2023, based on the November 2022 submission.

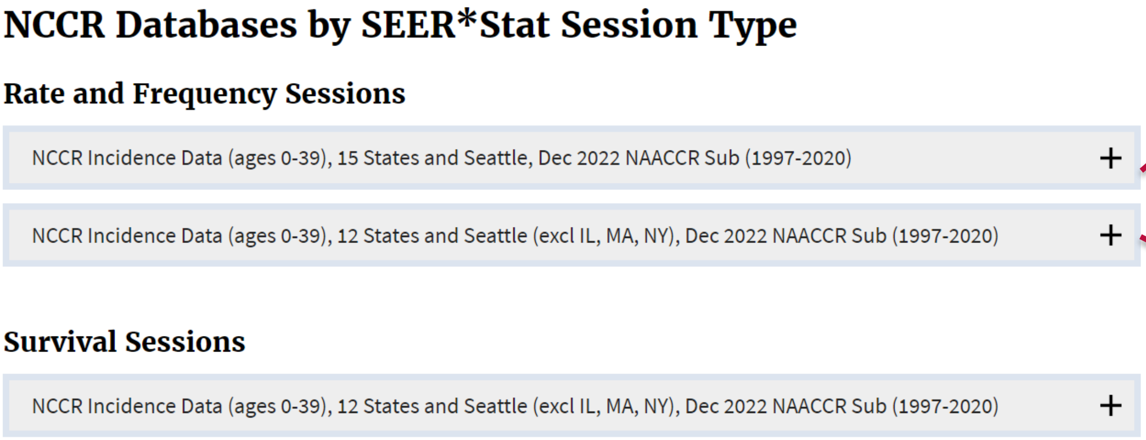


SEER*Stat Databases: NCCR December 2022 Submission

Documentation about the National Childhood Cancer Registry (NCCR) databases in SEER*Stat is available below and includes a database dictionary to describe all the variables as well as a summary by SEER*Stat session type of the main characteristics of each database. Read more about the NCCR and other available NCCR data products on the NCCR*Explorer website.

Documentation of NCCR Variables

Dictionary of NCCR Variables: Excel (XLSX, 18 KB) or PDF (PDF, 137 KB)
Lists all the variables in the NCCR SEER*Stat databases with the following information: NAACCR item number, description and definitions, and category.



NCCR Databases by SEER*Stat Session Type

Rate and Frequency Sessions

- NCCR Incidence Data (ages 0-39), 15 States and Seattle, Dec 2022 NAACCR Sub (1997-2020) +
- NCCR Incidence Data (ages 0-39), 12 States and Seattle (excl IL, MA, NY), Dec 2022 NAACCR Sub (1997-2020) +

Survival Sessions

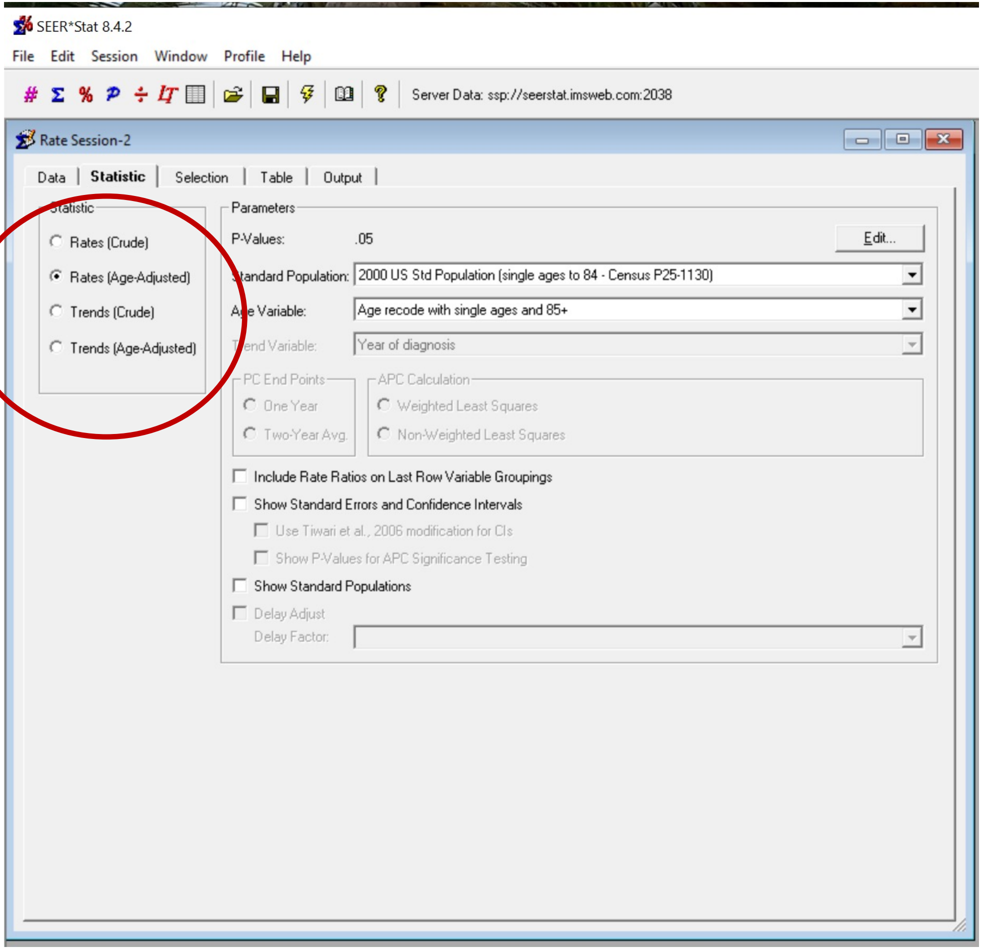
- NCCR Incidence Data (ages 0-39), 12 States and Seattle (excl IL, MA, NY), Dec 2022 NAACCR Sub (1997-2020) +

48% of the US pop ages 0-39

37% of the US pop ages 0-39

seer.cancer.gov/data-software/documentation/seerstat/nccr/dec2022/https://seer.cancer.gov/data-software/documentation/seerstat/nccr/dec2022/

- Create a new session
- Select a database
- **Choose a statistic**
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results



Server Data: ssp://seerstat.imsweb.com:2038

Statistic
○ Rates (Crude)
● Rates (Age-Adjusted)
○ Trends (Crude)
○ Trends (Age-Adjusted)

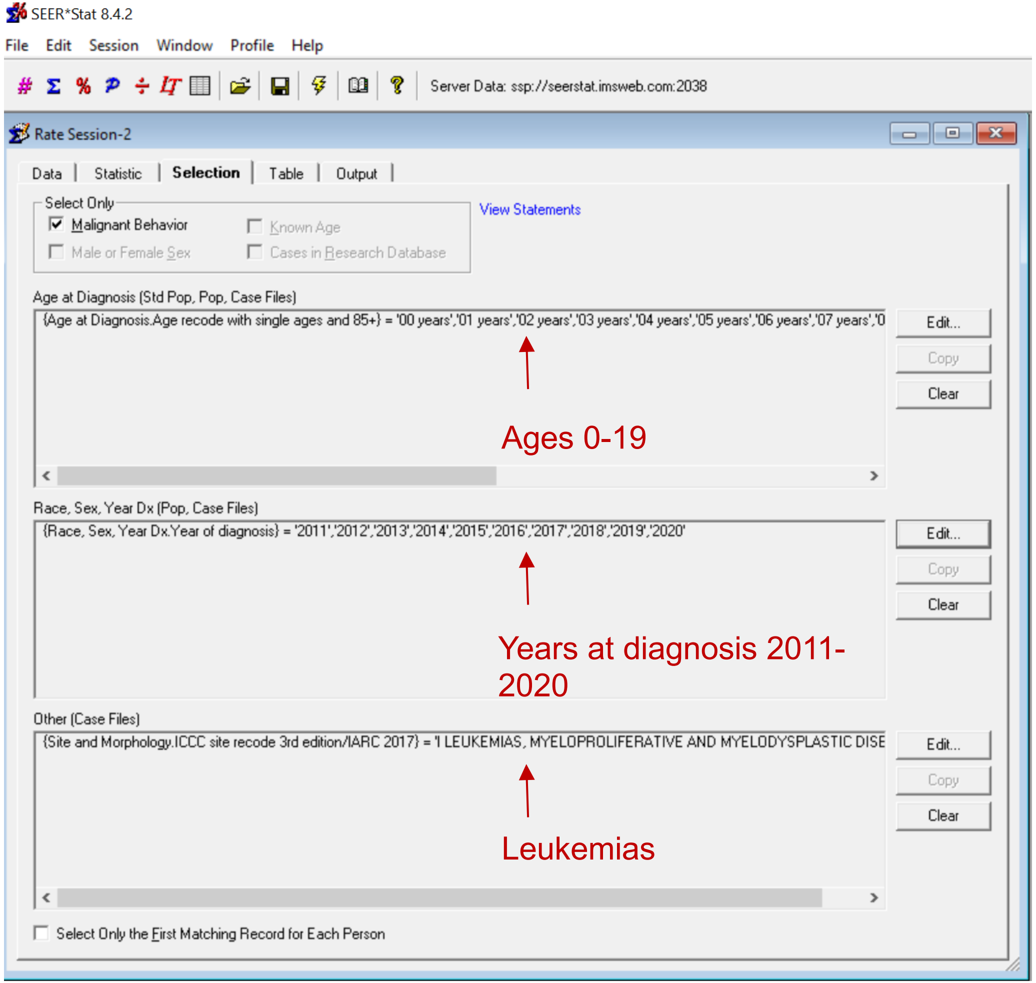
Parameters
P-Values: .05
Standard Population: 2000 US Std Population (single ages to 84 - Census P25-1130)
Age Variable: Age recode with single ages and 85+
Trend Variable: Year of diagnosis

PC End Points
○ One Year
○ Two-Year Avg.

APC Calculation
○ Weighted Least Squares
○ Non-Weighted Least Squares

☐ Include Rate Ratios on Last Row Variable Groupings
☐ Show Standard Errors and Confidence Intervals
  ☐ Use Tiwari et al., 2006 modification for CIs
  ☐ Show P-Values for APC Significance Testing
☐ Show Standard Populations
☐ Delay Adjust
Delay Factor:

- Create a new session
- Select a database
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results



Age at Diagnosis (Std Pop, Pop, Case Files)

{Age at Diagnosis.Age recode with single ages and 85+} = '00 years','01 years','02 years','03 years','04 years','05 years','06 years','07 years','0

Ages 0-19

Race, Sex, Year Dx (Pop, Case Files)

{Race, Sex, Year Dx.Year of diagnosis} = '2011','2012','2013','2014','2015','2016','2017','2018','2019','2020'

Years at diagnosis 2011-2020

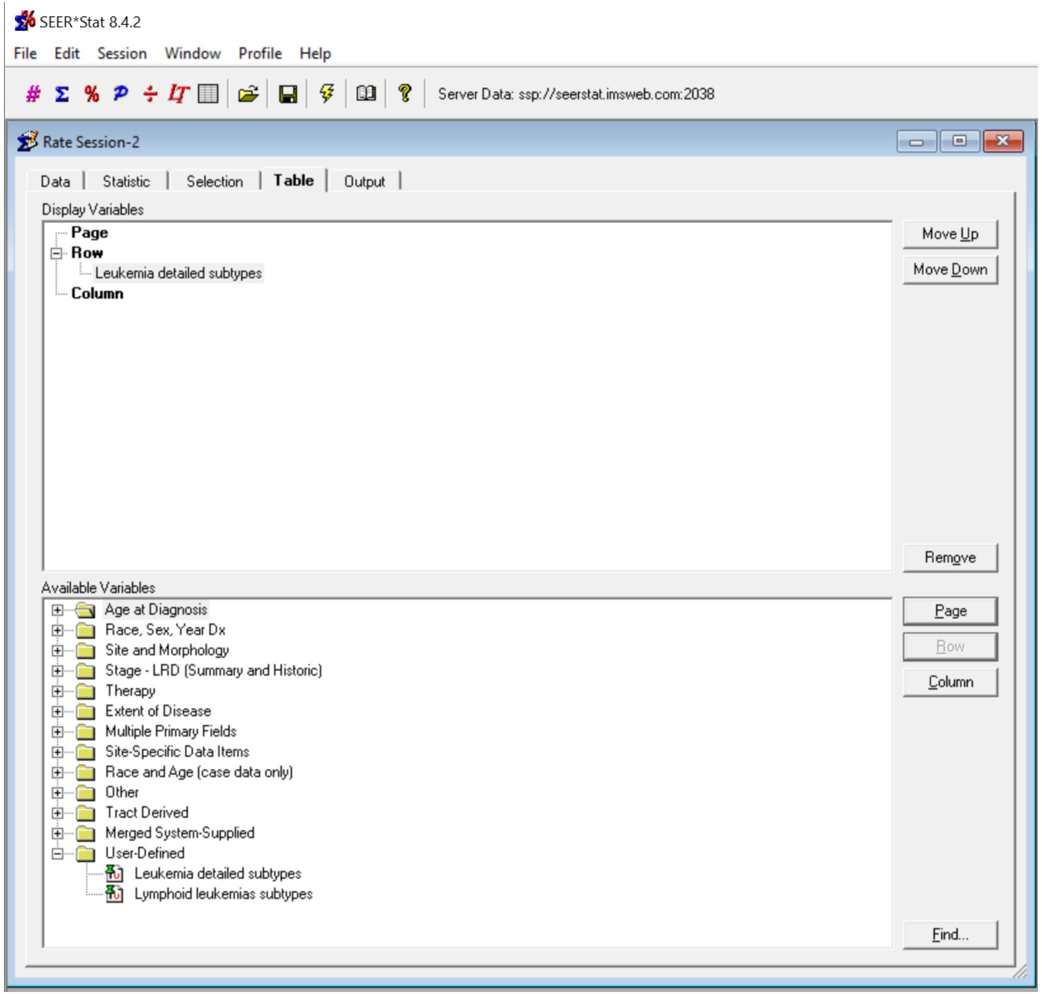
Other (Case Files)

{Site and Morphology.ICCC site recode 3rd edition/IARC 2017} = 'I LEUKEMIAS, MYELOPROLIFERATIVE AND MYELODYSPLASTIC DISE

Leukemias

☐ Select Only the First Matching Record for Each Person

- Create a new session
- Select a database
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results



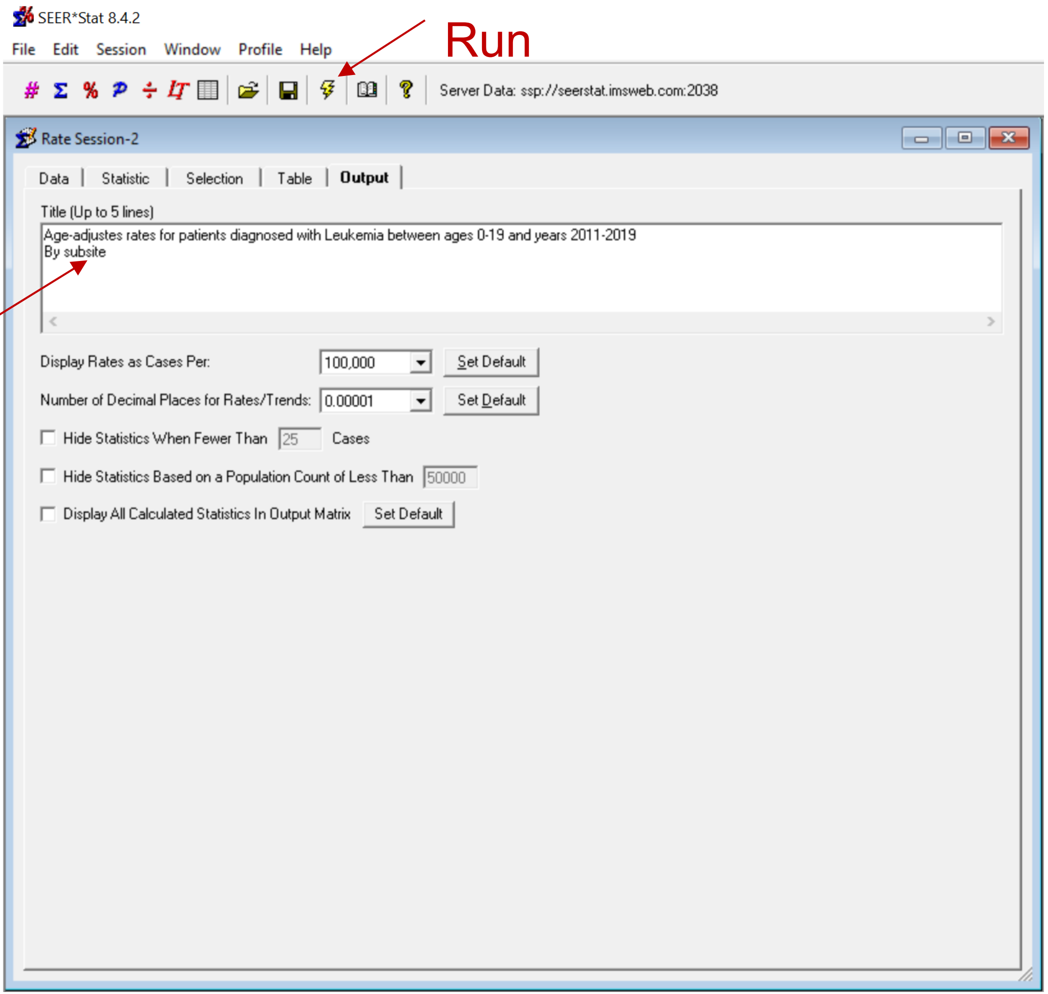
Display Variables

- Page
- Row
  - Leukemia detailed subtypes
- Column

Available Variables

- Age at Diagnosis
- Race, Sex, Year Dx
- Site and Morphology
- Stage - LRD (Summary and Historic)
- Therapy
- Extent of Disease
- Multiple Primary Fields
- Site-Specific Data Items
- Race and Age (case data only)
- Other
- Tract Derived
- Merged System-Supplied
- User-Defined
  - Leukemia detailed subtypes
  - Lymphoid leukemias subtypes

- Create a new session
- Select a database
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- **Run the software**
- Save/copy the results



Server Data: ssp://seerstat.imsweb.com:2038

Title (Up to 5 lines)

Age-adjustes rates for patients diagnosed with Leukemia between ages 0-19 and years 2011-2019
By subsite

Display Rates as Cases Per: 100,000

Number of Decimal Places for Rates/Trends: 0.00001

☐ Hide Statistics When Fewer Than 25 Cases

☐ Hide Statistics Based on a Population Count of Less Than 50000

☐ Display All Calculated Statistics In Output Matrix

Run

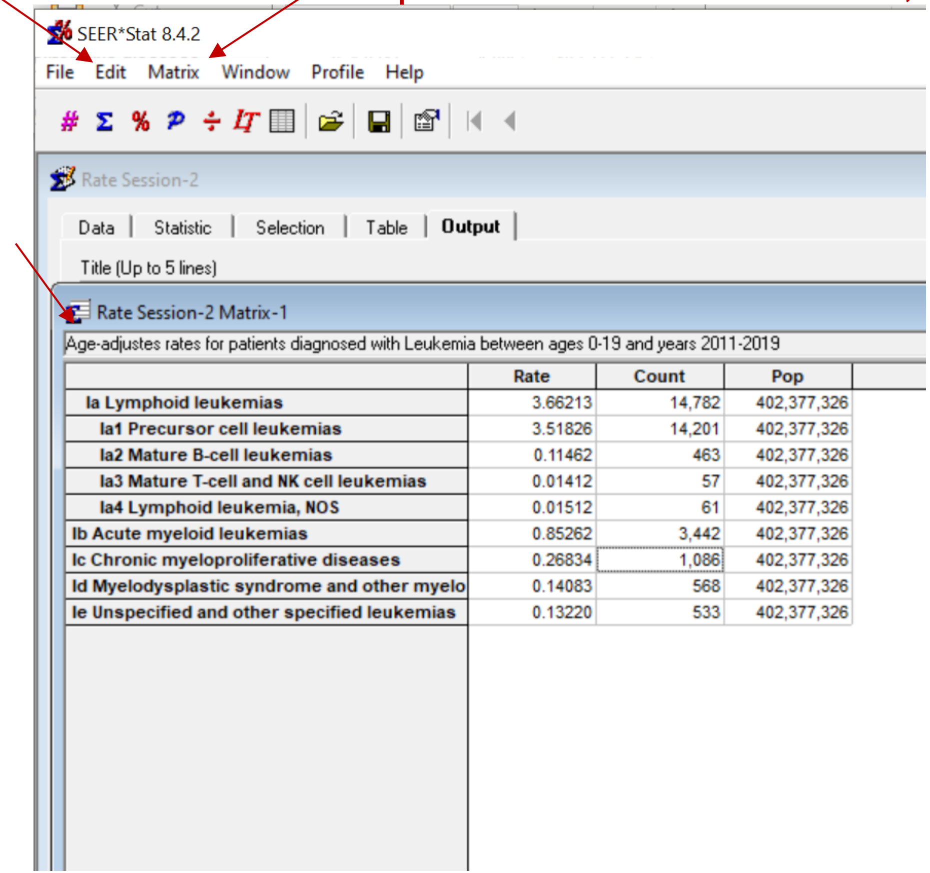
Title

- Create a new session
- Select a database
- Choose a statistic
- Define the cohort (subset)
- Create table (stratification)
- Run the software
- Save/copy the results

Copy function

Export to different format: csv, txt

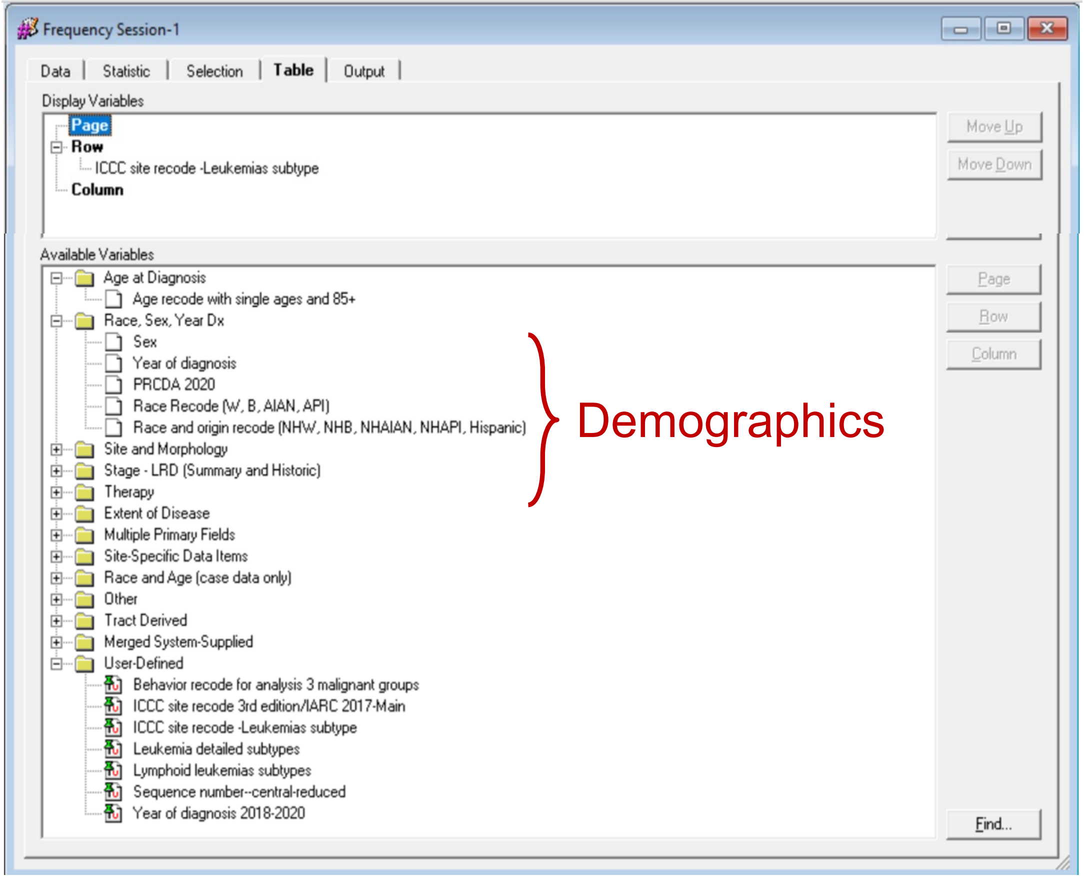
Results



Age-adjustes rates for patients diagnosed with Leukemia between ages 0-19 and years 2011-2019

	Rate	Count	Pop
Ia Lymphoid leukemias	3.66213	14,782	402,377,326
Ia1 Precursor cell leukemias	3.51826	14,201	402,377,326
Ia2 Mature B-cell leukemias	0.11462	463	402,377,326
Ia3 Mature T-cell and NK cell leukemias	0.01412	57	402,377,326
Ia4 Lymphoid leukemia, NOS	0.01512	61	402,377,326
Ib Acute myeloid leukemias	0.85262	3,442	402,377,326
Ic Chronic myeloproliferative diseases	0.26834	1,086	402,377,326
Id Myelodysplastic syndrome and other myelo	0.14083	568	402,377,326
Ie Unspecified and other specified leukemias	0.13220	533	402,377,326

Variables in SEER*Stat Are Organized in Folders



Display Variables

- **Page**
- **Row**
  - ICCC site recode -Leukemias subtype
- **Column**

Available Variables

- Age at Diagnosis
  - Age recode with single ages and 85+
- Race, Sex, Year Dx
  - Sex
  - Year of diagnosis
  - PRCDA 2020
  - Race Recode (W, B, AIAN, API)
  - Race and origin recode (NHW, NHB, NHAIAN, NHAPI, Hispanic)
- Site and Morphology
- Stage - LRD (Summary and Historic)
- Therapy
- Extent of Disease
- Multiple Primary Fields
- Site-Specific Data Items
- Race and Age (case data only)
- Other
- Tract Derived
- Merged System-Supplied
- User-Defined
  - Behavior recode for analysis 3 malignant groups
  - ICCC site recode 3rd edition/IARC 2017-Main
  - ICCC site recode -Leukemias subtype
  - Leukemia detailed subtypes
  - Lymphoid leukemias subtypes
  - Sequence number--central-reduced
  - Year of diagnosis 2018-2020

Demographics

Variables: Site and Morphology

- Extended International Classification of Childhood Cancer (ICCC) recodes
- Adolescents and young adults (AYA) site recode classifications
- ICD-O-3 topography, histology, behavior, and recodes
    - Behavior code ICD-O-3: (Benign, Borderline malignancy, In situ, & Malignant)
    - Behavior recode for analysis: created so that data analyses could eliminate major groups of histology/behaviors that weren't collected consistently over time
    - For more information, see seer.cancer.gov/behavrecode
- Grade

Behavior Recode for Analysis

	Benign	Borderline malignancy	In situ	Malignant	Only malignant in ICD-O-3 2001+	Only malignant 2010+
I LEUKEMIAS	0	0	0	44,056	1,832	125
II LYMPHOMAS	0	0	0	24,137	0	2,314
III CNS	8,259	5,311	0	28,999	50	0
IV NEUROBLASTOMA	0	0	0	7,907	0	0
V RETINOBLASTOMA	0	0	0	3,172	0	0
VI RENAL TUMORS	0	0	0	6,266	0	0
VII HEPATIC TUMORS	0	0	0	2,510	0	0
VIII MALIGNANT BONE TUMORS	0	0	0	8,693	0	0
IX SOFT TISSUE	0	0	0	11,394	1	0
X GERM CELL TUMORS	337	35	0	11,487	0	0
XI OTHER MALIGNANT EPITHELIAL NEOPLASMS …	0	0	0	18,296	0	0
XII OTHER AND UNSPECIFIED MALIGNANT NEOPLASMS	0	0	0	832	0	0
Not classified by ICCC or in situ	3,200	1,004	1,379	190	0	0

Extent of Disease and Stage

- Extent of disease (EOD) provide information about tumor size/thickness, lymph node involvement, and extension (metastasis), and have changed throughout the years
- We do not have a good stage system for childhood cancers
- The best stage to use is the Summary Stage (2004+), which classify tumors into local, regional, and distant

Therapy: First Course of Treatment

All methods of treatment recorded in the treatment plan and administered to the patient before disease progression or recurrence

- Surgery
- Radiation therapy
- Systemic treatment (chemotherapy)
    - **“Not received” and “Unknown if received” grouped together**(1)

seer.cancer.gov/tools/codingmanuals/index.html

seer.cancer.gov/seerstat/variables/seer/surgery

(1) Noone AM, Lund JL, Mariotto, A, et al. Comparison of SEER treatment data with Medicare claims. *Medical Care* 2016 Aug 22. doi: 10.1097/MLR.0000000000000073.

Multiple Primary Fields

- Date of diagnosis (month and year)
- Tumor sequence number
  - People can be diagnosed with more than one tumor
  - 00 = the only tumor of a person, 01 = first tumor, 02 = second tumor …
- 2% of all new tumors from patients age 0-19 who were diagnosed between 2016 and 2020 are second or later cancers (in contrast to 16% for tumors of all ages)

Other: Reporting Source and CoC Flag

- **Type of Reporting Source**
  - E.g. Hospital, physician's office, nursing home, autopsy/death certificate, other hospital outpatient
  - Over 95% of the tumors diagnosed at ages 0-19 are diagnosed in hospitals

- **CoC Accredited Flag (2018+)**
  - The CoC Accredited Flag identifies if the data abstraction was prepared at a Commission on Cancer (CoC) accredited facility
  - Only available for cases diagnosed 2018+

	2018-2020	
Abstract prepared at a facility:	No.	%
WITH CoC accreditation	16,752	63.3%
WIHOUT CoC accreditation	9,645	36.5%
Unknown	55	0.2%
Total	26,452	100.0%

Follow-Up and Survival Information

- Survival information was not available or had missing information for 3 registries → survival data available for 12 states and Seattle
- Death and cause of death collected through linkages with CDC's National Death Index database
- Cause of death information is not available because of confidentiality issues
- However, a recode identifying if the cause is likely associated with cancer versus not is included (seer.cancer.gov/causespecific)(1)

(1) Howlader N, Ries LA, Mariotto, A, et al. Improved estimates of cancer-specific survival rates from population-based data. *Journal of the National Cancer Institute* 2010 Oct 20. doi: 10.1093/jnci/djq366

Limitations

- Geographic information is not available. Statistics may only be calculated for all registries combined for confidentiality reasons.
- NCCR databases are not available in case listing sessions, meaning it is not possible to view individual records in the data (confidentiality reasons)
- Only SEER registries were included in this release
- This release does not offer graphical visualizations—the next version of SEER*Stat will have graphs

NCCR*Explorer vs NCCR Data in SEER*Stat

NCCR*Explorer

- Easier to use and get quick answers to questions
- Answers are limited to the statistics available
- Provides graphical output

NCCR Data in SEER*Stat

- More difficult to use; requires software and data knowledge
- More flexible and customizable; can provide answers for a much wider range of questions and analysis
- Includes more detailed information on tumor and treatment

Future Plans

- NCCR*Explorer
  - Will continue expanding on statistics, e.g. including long-term survival up to 10 years after diagnosis
- NCCR data in SEER*Stat
  - The current NCCR database includes 15 states and Seattle and covers 48% of the 0-39 US population
  - Next release (spring 2024) will include all participating NCCR registries and will cover over 70% of the 0-39 US population
  - It may include a few more variables

Future Plans

- Future SEER*Stat webinar in spring 2024
- Re-development of SEER*Stat → New SEER*Stat interface
- NCCR data platform:
  - Will include data beyond those collected by registries that will potentially have more detailed treatment, genomic characterization, residential history, social determinants of health, etc. …
  - Will require Institutional Review Board (IRB) for access

Acknowledgements

- Childhood Cancer Data Initiative (CCDI)
- NCCR Cancer Registries
- North American Association of Central Cancer Registries (NAACCR)
- Information Management System (IMS)
- National Cancer Institute (NCI)

Questions?

Thank you!

- Q&A
- Contact: mariotta@mail.nih.gov

Find Out More About CCDI

Visit the CCDI Hub and access the CCDI Data Ecosystem.
ccdi.cancer.gov

Subscribe to our monthly newsletter.
cancer.gov/CCDI

Questions? Email us.
NCIChildhoodCancerDataInitiative@mail.nih.gov

DEPARTMENT OF HEALTH & HUMAN SERVICES • USA

NIH

NATIONAL
CANCER
INSTITUTE

www.cancer.gov

www.cancer.gov/espanol