Title: Investigating new subsequent neoplasm ascertainment methods in the Childhood Cancer Survivor Study with population-based cancer registry data

CCSS Working Groups: Biostatistics/Epidemiology, Subsequent Neoplasm

Investigators:

Cindy Im, UMN
Kayla Stratton, Fred Hutch
Vikki Nolan, SJCRH
Shani Alston, SJCRH
Kumar Srivastava, SJCRH
Jennifer M. Yeh, Boston Children's Hospital
Chaya Moskowitz, MSKCC
Tara O. Henderson, Lurie Children's Hospital
Lindsay M. Morton, NCI
Yutaka Yasui, SJCRH
Joseph P. Neglia, UMN
Lucie M. Turcotte*, UMN
Wendy M. Leisenring*, Fred Hutch
Gregory T. Armstrong*, SJCRH

*Equal contribution

BACKGROUND

In the Childhood Cancer Survivor Study (CCSS), the current subsequent neoplasm (SN) ascertainment process involves confirming self-/proxy-reported SNs from periodic surveys by pathologist/oncologist review of pathology/medical records or death certificates. Although this CCSS SN confirmation protocol has generated robust data, it is labor intensive and may result in an under-ascertainment of SNs due its reliance on initial self-/proxy-report (e.g., false negatives due to loss to follow-up or death) and pursuit of medical documentation (e.g., false negatives when medical records cannot be reviewed). Furthermore, candidate SNs are also excluded from SN-specific analyses (though still included in mortality analyses) when SN-related death from the National Death Index (NDI) is the sole source of SN information and detailed information about SN diagnosis based on pathology or medical record report is unavailable. Cohort linkage to population-based cancer registry data potentially offers a cost-effective and potentially more comprehensive (as it can include CCSS non-participants) alternative to current CCSS SN ascertainment procedures.

Recent efforts by the US National Cancer Institute (NCI) have led to the development of the Virtual Pooled Registry-Cancer Linkage System (VPR-CLS), facilitating linkages between cohort studies and the majority of population-based state and regional cancer registries in the US with largely complete data from the mid-to-late 1990s. Preliminary work by NCI (Liu *et al.*[1]) assessed the feasibility and benefit of large-scale cohort linkage with 43 cancer registries participating in VPR-CLS with the US Radiologic Technologist (USRT) Study, where cancer-related follow-up is captured only by self-report in questionnaires. In their sample (N=98,944), their current method of cancer ascertainment had an overall sensitivity of 58%, i.e., proportion of USRT-reported first primary cancers also reported in registry data, matching by cancer type and diagnosis dates within a two-year window. Nearly 28% of false-negative reports were due to loss to follow-up. Agreement between USRT self-reported cancers and registry-reported cancers differed by cancer type, where cancers with higher survival were more likely to be

captured by self-report. In subsequent work in the same cohort,[2] registry data identified 4-times more incident cancer cases. Consequently, cancer incidence rates when using self-reported data only were underestimated. Combining registry data with self-reported cancers from USRT surveys provided the greatest cohort coverage. While there was limited evidence of bias in cancer risk association estimates with sociodemographic and lifestyle factors using self-reported cases, data combining the larger number of registry-based cases with more extended follow-up yielded risk association estimates with narrower confidence intervals. Similar conclusions were drawn in another SN ascertainment analysis comparing active reports versus cancer registry data in a linkage study between the Center for International Blood and Marrow Transplant Research (CIBMTR) and California Cancer Registry.[3]

In CCSS, a preliminary feasibility study testing VPR-CLS linkage was conducted with 23 state registries covering <74% of the US population with data through year 2016 (unpublished). Ultimately, 1,652 VPR-CLS SN cases that were likely matches for SNs among CCSS participants (full cohort) were evaluated. Despite the lower US coverage by VPR-CLS, 24% (N=400) were SNs reported to registries after CCSS last contact and were also not recoverable by NDI linkage, i.e., potential false negatives due to loss to follow-up. Among the 989 VPR-CLS cancer cases that were reported before CCSS last contact, 26% (N=256) were not reported or confirmed by CCSS, i.e., potential false negatives due to missing self/proxy reports or medical documentation. However, the current CCSS SN ascertainment protocol showed good sensitivity, i.e., 74% of VPR-CLS SNs were also confirmed in CCSS. Importantly, due to gaps in VPR-CLS coverage, 29% of SNs in CCSS were missing in VPR-CLS registry data.

Currently, VPR-CLS includes 45 registries covering approximately 95% of the US state population and Puerto Rico,[4] with most registries initiating data collection after 1995. The most recent CCSS-VPR-CLS linkage was initiated in the last 18 months. With this linkage, CCSS will obtain cancer incidence data from 42 states (non-participating states: AR, DE, IL, KS, NV, SD, WV, VT). Linkages have been feasible for 39 states, with 3 states pending data transfers (FL, MN, MS). In these data, the earliest year of registry data transferred to CCSS ranges from 1973-1997 and the latest year of data transferred to CCSS ranges from 2020-2021. In this study, we propose to comprehensively evaluate hypothetical approaches that supplement or replace aspects of current CCSS SN ascertainment procedures by integrating VPR-CLS population-based cancer registry data.

SPECIFIC AIMS

The main **objective** of this project is to evaluate a range of hypothetical SN ascertainment approaches (see Table 1) that supplement OR replace current CCSS procedures using VPR-CLS data. Hypotheses are as follows:

- Differing degrees of VPR-CLS data integration with existing CCSS SN data will yield different estimates of SN incidence.
- Differing degrees of VPR-CLS data integration will affect SN characteristic information capture (e.g., site, diagnosis date, histology/stage/grade, tumor size).
- Precision of association measures for SN risk factors will be improved with the incorporation of VPR-CLS data. Bias and precision may be affected for SN types with underreporting patterns, e.g., more lethal cancers.

This proposal has the following specific aims:

Aim 1: Assess whether alternative SN ascertainment approaches utilizing VPR-CLS: (a) can improve or maintain the validity of CCSS SN data, or (b) change SN incidence or treatment risk

association estimates, when compared to the standard CCSS approach of self-report with validation.

- Aim 1a: Compare the distributions of SN types/characteristics and SN information missingness across different approaches.
- Aim 1b: Compare survivor characteristics of identified cases (e.g., demographics, childhood cancer diagnosis, cancer treatment) across different approaches.
- Aim 1c: Identify bias and/or changes in precision for: (a) SN descriptive epidemiologic measures (incidence and SIR/AER), and (b) measures of association for major treatment risk factors across different approaches.

Aim 2: Conduct comparable analyses with registry-only data (VPR-CLS, NDI) in the all-eligible population and compare with results using confirmed CCSS SNs.

- Aim 2a: Characterize changes in the incidence of SN types during the analysis time period by attained age, treatment decade, and childhood cancer diagnosis.

Aim 3 (Exploratory): Compare time/effort and cost estimates between SN ascertainment approaches.

METHODS

Study population

For all analyses, the follow-up time period of interest is based on coverage for which all participating cancer registries had comprehensive cancer data, i.e., 1997 - 2020 (or 2021, considering NDI). We note that this time frame may change depending on the final assessment of available data from VPR-CLS and NDI linkages. For Aim 1, participants who completed CCSS baseline surveys and survived to the start of analysis follow-up will be included. For Aim 2, all CCSS-eligible five-year survivors who survived to the start of analysis follow-up with complete sociodemographic and primary cancer-related data at diagnosis (captured from each institution in the tracking database at the time of determination of cohort eligiblilty; i.e., not self-reported) will be included.

Outcomes

In this project, all SNs during the analysis period will be counted; cancer types will be limited to those reported to NAACCR (and also consistent with CCSS). Separate analyses will be conducted for SNs overall, SMNs overall, and specific SN types including breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors.

A harmonized classification system for SNs will be applied as needed to assure comparisons between CCSS-confirmed SNs and SNs identified with linkage with VPR-CLS. Data that will be assessed for each SN includes: cancer type; cancer site; diagnosis date; histology/stage. SNs will be ascertained using different approaches that adopt varying levels of VPR-CLS registry-based cancer incidence data (Table 1) and compared to the benchmark approach, i.e., the current procedure for confirming CCSS self/proxy-reported SNs.

Table 1: SN ascertainment in CCSS versus other SN ascertainment approaches adopting VPR-CLS registry-based cancer incidence data

SN ascertainment approaches Orange=CCSS practice Green=CCSS reference Blue=Registry reference	**Method**	**Follow-up window**	**Sample size**
Confirmed SN (benchmark)	Current practice in CCSS: • Pathology/oncology clinician review and confirmation of self-reported from CCSS surveys • Death certificate information used to capture SNs after last CCSS FU[a]	From FU start to event, death, or max[NDI search date, last CCSS questionnaire with SNs], whichever comes first	CCSS enrolled
A. Registry-Enhanced	Supplement confirmed self-reported CCSS SNs: • CCSS confirmed SNs = starting point • Supplement NDI SNs or unconfirmed SNs with VPR-CLS data ○ *All SNs in CCSS or NDI-identified, but NDI SNs must have a VPR-CLS match*	Same as above	CCSS enrolled
B. Registry-Enhanced + Added	Supplement confirmed self-reported CCSS SNs: • CCSS confirmed SNs = starting point • Supplement NDI SNs or unconfirmed SNs with VPR-CLS data (same as above) • Add VPR SNs after last CCSS FU	Same as above	CCSS enrolled
C. Registry + Survey	VPR-CLS, NDI, and CCSS surveys used. • Registry data = starting point; use registry-reported date of diagnosis • SNs reported in surveys (may or may not be pathology/oncology-validated in CCSS) but not registries: Use imputed diagnosis dates from self/proxy report • NDI SNs without a VPR-CLS match: included if self-report data is consistent	From FU start to event, death, or max[VPR-CLS linkage date, date of NDI search date, last CCSS questionnaire with SNs], whichever comes first	CCSS enrolled
D. Registry + Confirmed SN	VPR-CLS, NDI, and confirmed self-reports used. • Registry data = starting point; use registry-reported date of diagnosis • SNs reported in surveys but not registries: Use diagnosis dates from confirmed SNs • NDI SNs without a VPR-CLS match: included if confirmed CCSS SN data is consistent	Same as above	CCSS enrolled
E. Registry Only	Only VPR-CLS and NDI used. • Use registry-reported date of diagnosis (no CCSS SN data used) • NDI SNs included, only ruled out if cause of death may reflect recurrent or metastatic disease	From FU start to event, death, or max[VPR-CLS linkage date, date of NDI search date], whichever comes first	All eligible for CCSS

a. Included if cause of death from death certificate is clearly related to SN; diagnosis date is imputed with a date immediately before the death date.

Given that SN diagnosis dates may be incomplete (e.g., in CCSS surveys, only age at diagnosis is provided and some registries only provide diagnosis year), missing months/days will be imputed with the midpoint of year (June 15) if only year of diagnosis is available. For this reason, a window of time (e.g. ±2 years) around an SN occurrence will be used to assess concordance between SNs from different sources.

Exposures and other covariates

- Age/date at primary cancer diagnosis
- Primary cancer diagnosis
- Sex
- Race and ethnicity (Hispanic, Non-Hispanic Black, Non-Hispanic White, Other)
- Attained age
- Vital status, date of death, and cause of death (SN-related death)
- Last follow-up date
- Cancer treatment
    - Any radiation therapy (RT)
    - Radiation therapy exposure and dose, each of 7 major body regions (cranial, neck, chest, abdomen, pelvis, arm, leg)
    - Any chemotherapy
    - Alkylating agents, exposure and dose (cyclophosphamide-equivalent dose)
    - Anthracyclines, exposure and dose (doxorubicin-equivalent dose)
    - Epipodophyllotoxins, exposure and dose
    - Platinum, exposure and dose (cisplatin-equivalent dose)

Statistical analysis

To compare confirmed CCSS SNs and VPR-CLS SNs, we will assess agreement using an hierarchical approach adopted in previous work,[1,2] considering: (a) same cancer type and site, (b) diagnosis within ±2 years, and (c) both same cancer type and site and diagnosis within ±2 years. As a sensitivity analysis, we will consider less granular criteria for concordance, e.g., considering matches by broader organ or SN types. The referent cancer type and diagnosis date information depend on SN ascertainment approach: specifically, if CCSS is the starting point (approaches A-B), the confirmed CCSS cancer type and diagnosis date are treated as the reference; if VPR-CLS is the starting point (approaches C-E), the cancer type and diagnosis date/year from registries are treated as the reference. For approaches C-D, we will use DatStat free text responses from surveys available in CCSS data freezes and follow existing CCSS protocols used to preliminarily review self-reported SN free text responses to identify self-reported SNs that may overlap with VPR-CLS cancer entries. All self-report and registry-based SN discrepancies, i.e., self-reported SNs in CCSS surveys that are cannot be reconciled with existing VPR or NDI data, will be corroborated by CCSS pathologists/oncologists involved in current CCSS SN confirmation procedures. Given the relatively high VPR-CLS US population coverage, we note that agreement will not rely on aligning participants' residential information over time with cancer reporting by specific state registries.

Potential SNs identified via NDI for approaches A-B, i.e., without CCSS SN confirmation, will be reviewed and counted as a valid SN if corroborated by VPR-CLS data. Potential SNs identified via NDI for approaches C-D, i.e., without matching a VPR-CLS SN, may be included if the SN is consistent with CCSS follow-up data, i.e., reconcilable with self-reported SN information (approach C) or confirmed SN data (approach D). For approaches A-D, agreement will require the same general criteria above, i.e., same cancer type and diagnosis age ±2 years between information sources. For approach E, NDI-only SNs will be included without a secondary confirmation source. For all approaches, NDI-based SNs will be considered only if cause of death information clearly excludes non-unique SNs and recurrent disease or metastatic disease only. All SNs where NDI information is the primary ascertainment source (especially with respect to approach E) will be reviewed by CCSS pathologists/oncologists.

All SN ascertainment approaches will use existing confirmed CCSS SN data as the benchmark, identifying participants as:

- True positive (TP) = Confirmed CCSS SN and other-approach SN, or
- False negative (FN) = Other-approach SN, but not a confirmed CCSS SN.

Proposed analyses are summarized in Table 2 (below).

Table 2: Summary of measures by SN ascertainment approaches and proposal aims

	Summary of measures by approach and aim		
Other SN ascertainment approaches	**Aim 1a**	**Aim 1b**	**Aim 1c and 2**
A. Registry-Enhanced	• Missingness: site/histology and diagnosis date/year • # of TPs, FNs • FN rate (FN/[FN+TP]), i.e., % other-approach SNs missed by CCSS • Separately assess loss due to protocol incompletion (approach A) or loss to follow-up (approach B)	Calculate the following for each of the other approaches separately: • Break out Aim 1a statistics by sociodemographic and clinical characteristics • Associations between FN probability and characteristics (ORs), from logistic regression models for FN status (yes/no)	Calculate the following for confirmed CCSS SNs (benchmark approach) and each set of SNs ascertained with other approaches: • Cumulative incidence • Incidence rate • SIR, AER • Treatment associations (RRs), from multivariable piecewise exponential models
B. Registry-Enhanced + Added			
C. Registry + Survey	• Missingness: site/histology and diagnosis date/year • # of TPs, FNs, ***and CCSS-only SNs (approaches C, D)*** • FN rate (FN/[FN+TP]), i.e., % other-approach SNs missed by CCSS • ***Predictive value = TP/(# CCSS SNs)***		
D. Registry + Confirmed SN			
E. Registry Only	*Approach E: CCSS participants only*	*Approach E: CCSS eligibles*	*Approach E: CCSS eligibles*

For Aim 1a, analyses will characterize the following across SN ascertainment approaches: SN data missingness (e.g., missing site/histology information or diagnosis date/year); the number of true positives and false negatives (i.e., changes in the total number of SNs per approach); and the false negative rate (i.e., proportion of other-approach SNs that were not captured by the benchmark approach). For approaches that will leverage VPR-CLS to supplement the confirmed CCSS data (A-B), we will assess the false negative rate related to lack of SN confirmation (e.g., no self-report, no medical validation) and loss to follow-up. For approaches that replace aspects of CCSS SN confirmation, we will also evaluate the predictive value of each approach, i.e., the proportion of other-approach SNs among all confirmed CCSS SNs. These statistics will be evaluated by SNs and SMNs overall and by specific SN types; the latter will provide information as to whether the accuracy of SN ascertainment approaches differ by cancer type.

For Aim 1b, we will assess Aim 1a statistics by sociodemographic and clinical characteristics of interest, which will provide insights into factors that may underlie underreporting in CCSS. We will also evaluate adjusted associations between the false negative probability and these characteristics with odds ratios from logistic regression models considering each participant's false negative status (yes/no) by approach.

For Aim 1c, we will separately evaluate datasets with SNs generated from each of the different SN ascertainment approaches while considering confirmed CCSS SNs as the benchmark. To compare SN ascertainment approaches, we estimate the cumulative incidence of SNs/SMNs (overall and by major types) treating death as a competing risk, and the excess risk relative to the general population using standardized incidence ratios (SIRs) and absolute excess risks (AERs) using age-/sex-/year-matched SEER cancer incidence data. Time at risk will be defined by the analysis follow-up period (Table 1). Changes in these measures over time across the different SN ascertainment approaches will be described. Relative rates (RRs) from multivariable piecewise exponential models will assess previously reported associations between SNs (overall and by type) and RT dose (Turcotte et al.,[5] 2018 JAMA) and chemotherapy dose (Turcotte et al.,[6] 2019 JCO), adjusting for covariates as previously described.

Aim 2 analyses with the CCSS-eligible study population with registry-only SN ascertainment will reflect the principles and methods described above for Aim 1. Multiple imputation will be performed following previously described methods[7] to address treatment data missingness in the CCSS-eligible population. In analyses to understand patterns of SN incidence over time considering attained age, treatment decade, treatment exposures, and childhood cancer diagnosis, sensitivity analyses will be conducted considering survivors free of SNs at the start of at-risk time period.

Aim 3 analyses are exploratory and secondary to the primary aims described above. We will identify content experts familiar with CCSS to pragmatically quantify cost and effort across approaches.

DRAFT TABLES AND FIGURES

Tables and figures for analyses with CCSS participants (Aim 1)

Table 1: Demographic and clinical characteristics of childhood cancer survivors included in analyses

	CCSS (N=)		All eligible for CCSS (N=)	
Characteristic	N	Percent or median (IQR)	N	Percent or median (IQR)
Sex				
Female				
Male				
Race and ethnicity				
Hispanic/Latine				
Non-Hispanic Black				
Non-Hispanic White				
Other				
Age at last follow up (years)				
Age at childhood cancer diagnosis (years)				
Primary childhood cancer diagnosis				
Leukemia				
Acute lymphoblastic leukemia				
Acute myeloid leukemia				
Other leukemia				
Hodgkin lymphoma				
Non-Hodgkin lymphoma				
Sarcoma				
Soft tissue sarcoma				
Osteosarcoma				
Ewings sarcoma				
Other bone tumors				
Kidney tumors				
Neuroblastoma				
Central nervous system tumors				
Astrocytomas				
Medulloblastoma				
Other				
Childhood cancer treatment decade				
1970-1979				
1980-1989				
1990-1999				
Treatments for childhood cancer				
Any radiation therapy (RT)				
Hematopoietic cell transplantation (HCT)				
Any chemotherapy				
Any alkylating agents				
Any anthracyclines				
Any epipodophyllotoxins				
Any platinums				
Dosages: Median (IQR)[a]				
Maximum RT tumor dose, any field (cGy)				
Alkylating agents (mg/m^2)				
Anthracyclines (mg/m^2)				
Epipodophyllotoxins (mg/m^2)				
Platinum (mg/m^2)				
Vital status				
Deceased				

Table 2: Total counts of subsequent neoplasms reported between 1997-2021 across different ascertainment approaches in CCSS participants with VPR-CLS linkage

	CCSS only (N= individuals)				
Cancer type	**CCSS-confirmed (N)**	**Approach A (N or %)**	**Approach B (N or %)**	**Approach C (N or %)**	**Approach D (N or %)**
SNs (all types) True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
SMNs (all types) True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Breast True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Colorectal True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Thyroid True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
CNS tumors True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Soft tissue sarcomas True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Bone tumors True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					
Leukemia True positive (TP) Eliminated neoplasms[a] False negative (FN) FN rate[b] (%) Predictive value[c] (%)					

Cancer type	CCSS only (N= individuals)				
	CCSS-confirmed (N)	**Approach A (N or %)**	**Approach B (N or %)**	**Approach C (N or %)**	**Approach D (N or %)**
Lymphoma					
True positive (TP)					
Eliminated neoplasms[a]					
False negative (FN)					
FN rate[b] (%)					
Predictive value[c] (%)					
Melanoma					
True positive (TP)					
Eliminated neoplasms[a]					
False negative (FN)					
FN rate[b] (%)					
Predictive value[c] (%)					
Other solid tumor					
True positive (TP)					
Eliminated neoplasms[a]					
False negative (FN)					
FN rate[b] (%)					
Predictive value[c] (%)					

a. For the benchmark approach (CCSS-confirmed), the number of eliminated neoplasms reflects those not counted due to lack of confirmation for cancers identified with NDI information only or when medical record documentation could not be obtained. For other approaches, the number of eliminated neoplasms reflect potential FNs that were not counted due to data missingness (e.g., no diagnosis date or year, cancer type information lacks specificity).
b. False negative rate of approach= FN/(TP+FN)
c. Predictive value of approach = TP/(total CCSS-confirmed SNs)

Table 3a: Comparison of subsequent neoplasm counts by sociodemographic and clinical characteristics across ascertainment approaches

	CCSS (N=)				
	Confirmed in CCSS	Approach A	Approach B	Approach C	Approach D
Total count					
Characteristics	N (%)	N (%)	N (%)	N (%)	N (%)
Sex					
Female					
Male					
Race and ethnicity					
Hispanic/Latine					
Non-Hispanic Black					
Non-Hispanic White					
Other					
Age at last follow up (years)					
<30 years					
31-45 years					
>45 years					
Age at primary cancer diagnosis					
<10 years					
10 years or older					
Primary childhood cancer diagnosis					
Leukemia					
Hodgkin lymphoma					
Non-Hodgkin lymphoma					
Sarcoma					
Kidney tumors					
Neuroblastoma					
Central nervous system tumors					
Other					
Childhood cancer treatment decade					
1970-1979					
1980-1989					
1990-1999					
Treatment					
Surgery only					
Any radiation therapy ± surgery, no chemotherapy					

	CCSS (N=)				
	Confirmed in CCSS	Approach A	Approach B	Approach C	Approach D
Any chemotherapy ± surgery, no radiation therapy					
Any radiation therapy or chemotherapy					
HCT					

Table 3b-f: Repeat Table 3a for SMNs and specific SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors). Additional SMN types to be considered if counts are sufficient.

Table 4a: Associations between sociodemographic and clinical characteristics and the false negative probability for ascertainment approaches for SNs overall (CCSS only, N=)

Characteristics	**Approach A** OR (95% CI)	**Approach B** OR (95% CI)	**Approach C** OR (95% CI)	**Approach D** OR (95% CI)
Sex				
Female	Ref.	Ref.	Ref.	Ref.
Male				
Race and ethnicity				
Hispanic/Latine				
Non-Hispanic Black				
Non-Hispanic White	Ref.	Ref.	Ref.	Ref.
Other				
Age at last follow up (years)				
<30 years	Ref.	Ref.	Ref.	Ref.
31-45 years				
>45 years				
Age at primary cancer diagnosis				
<10 years	Ref.	Ref.	Ref.	Ref.
10 years or older				
Primary childhood cancer diagnosis				
Leukemia	Ref.	Ref.	Ref.	Ref.
Hodgkin lymphoma				
Non-Hodgkin lymphoma				
Sarcoma				
Kidney tumors				
Neuroblastoma				
Central nervous system tumors				
Other				
Childhood cancer treatment decade				
1970-1979	Ref.	Ref.	Ref.	Ref.
1980-1989				
1990-1999				
Treatment				
Surgery only	Ref.	Ref.	Ref.	Ref.
Any radiation therapy ± surgery, no chemotherapy				
Any chemotherapy ± surgery, no radiation therapy				
Any radiation therapy or chemotherapy				
HCT				

Table 4b-f: Repeat Table 4a for SMNs and specific SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors). Additional SMN types to be considered if counts are sufficient.

Table 5: Comparison of standardized incidence ratios (SIRs) and absolute excess risks (AERs) of SNs by ascertainment approaches

SN type and ascertainment pproach	Obs	Exp	Person-years of follow-up	SIR (95% CI)	SIR P_{trend}	AER (95% CI)	AER P_{trend}
SN, overall							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
SMN, overall							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Breast							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Colorectal							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Thyroid							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
CNS tumor							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Soft tissue sarcoma							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Bone tumor							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Leukemia							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Lymphoma							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Melanoma							

SN type and ascertainment pproach	Obs	Exp	Person-years of follow-up	SIR (95% CI)	SIR P_{trend}	AER (95% CI)	AER P_{trend}
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							
Other solid tumor							
CCSS-confirmed							
Approach A							
Approach B							
Approach C							
Approach D							

Figure 1: CONSORT-style flow describing exclusion criteria and final analytic sample

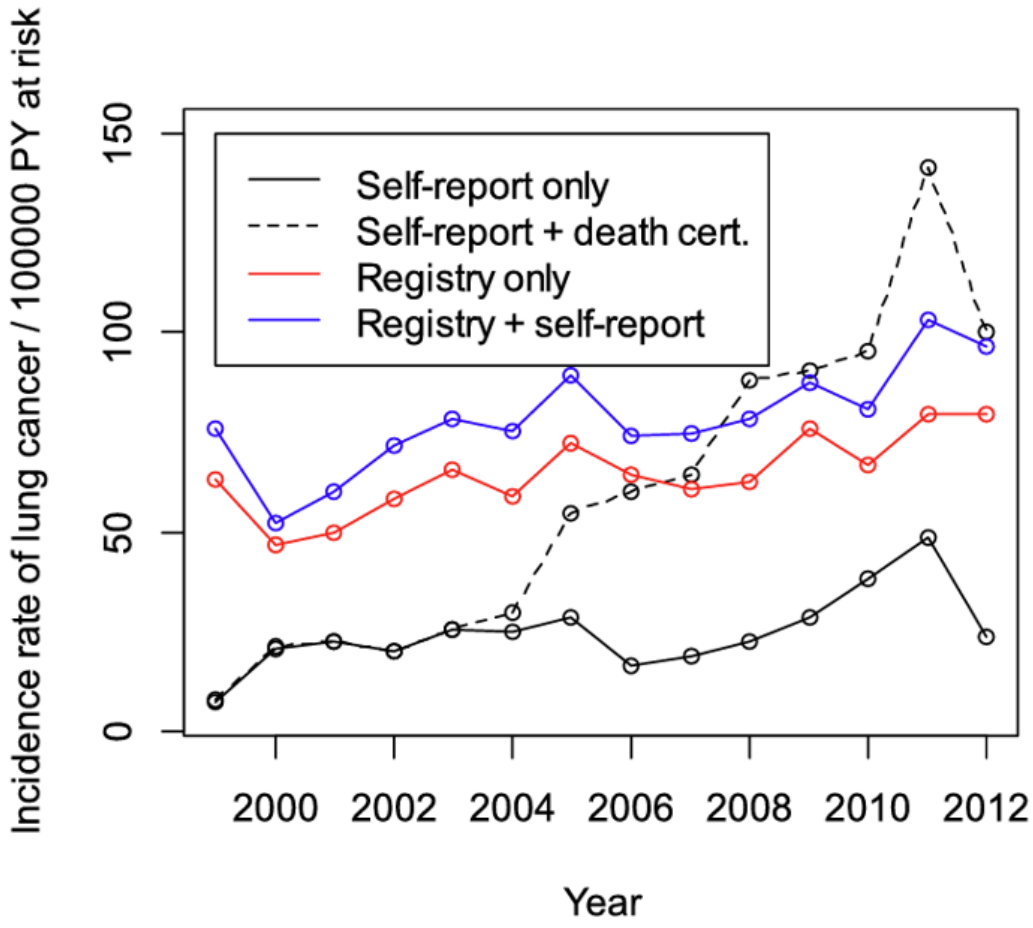

Figure 2a: Comparison of SN incidence rates (per 10,000 person-years) across ascertainment approaches

Figure 2b: Comparison of SMN incidence rates (per 10,000 person-years) across ascertainment approaches

Figure 2c-f: Comparison of incidence rates (per 10,000 person-years) across ascertainment approaches for SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors)

Example figure above; corresponding tables for each figure with incidence rates per year by approach will be included.

Figure 3a: Comparison of SN cumulative incidence across ascertainment approaches

Figure 3b: Comparison of SMN cumulative incidence across ascertainment approaches

Figure 3c-f: Comparison of cumulative incidence across ascertainment approaches for SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors)

Corresponding tables for each figure with cumulative incidence in 5-year follow-up increments with 95% CIs by approach will be included.

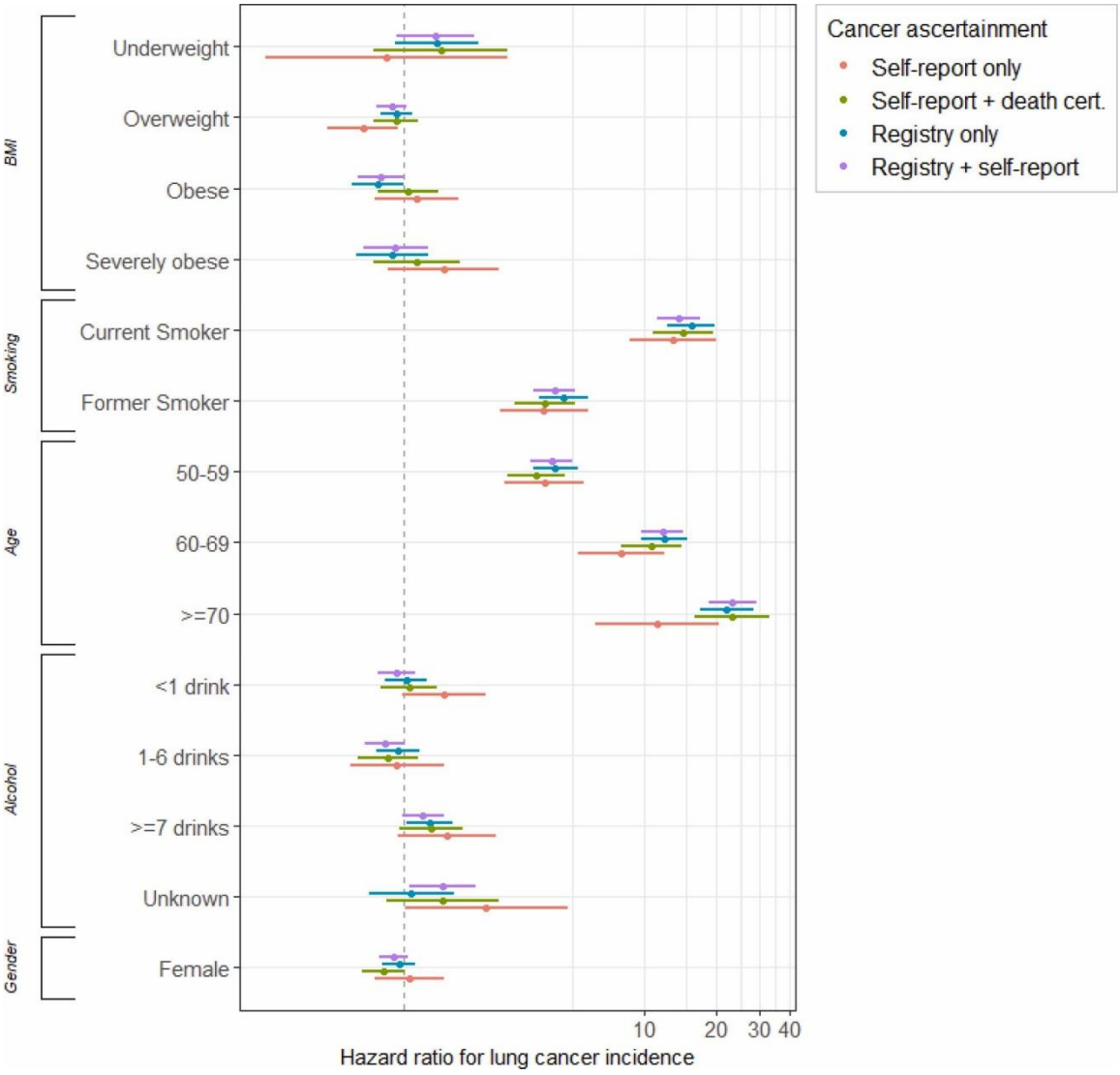

Figure 4a: Comparison of adjusted SN relative rates for radiotherapy and chemotherapy exposures across ascertainment approaches

Figure 4b: Comparison of adjusted SMN relative rates for radiotherapy and chemotherapy exposures across ascertainment approaches

Figure 4c-f: Comparison of adjusted relative rates across ascertainment approaches for SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors)

Example figure above. Corresponding tables for each figure by approach will be included.

Tables and figures for the all-eligible analyses (Aim 2):

Table 1: Total counts of subsequent neoplasms overall and by type reported between 1997-2021 using an SN ascertainment approach that utilizes registry data only (approach E)

	All eligible for CCSS (N= individuals)		Enrolled in CCSS among all eligible (N= individuals)				
Cancer type	Total (N)	# eliminated[a]	a. CCSS-confirmed only (N)	b. CCSS-confirmed and registry-identified (N)	c. Registry-identified only (N)	d. False-negative rate (%, c/[b+c])	e. Predictive value (%, b/[a+b])
SNs (all types)							
SMNs (all types)							
Breast							
Colorectal							
Thyroid							
CNS tumors							
Soft tissue sarcomas							
Bone tumors							
Leukemia							
Lymphoma							
Melanoma							
Other solid tumor							

a. Number of eliminated neoplasms reflect those that were not counted due to data missingness (e.g., no diagnosis date or year, cancer type information lacks specificity).

Table 2a: Comparison of subsequent neoplasm counts by sociodemographic and clinical characteristics across ascertainment approaches

	CCSS only (N=)	All eligible (N=)
	Confirmed in CCSS	Approach E
Total count		
Characteristics	N (%)	N (%)
Sex		
Female		
Male		
Race and ethnicity		
Hispanic/Latine		
Non-Hispanic Black		
Non-Hispanic White		
Other		
Age at last follow up (years)		
<30 years		
31-45 years		
>45 years		
Age at primary cancer diagnosis		
<10 years		
10 years or older		
Primary childhood cancer diagnosis		
Leukemia		
Hodgkin lymphoma		
Non-Hodgkin lymphoma		
Sarcoma		
Kidney tumors		
Neuroblastoma		
Central nervous system tumors		
Other		
Childhood cancer treatment decade		
1970-1979		
1980-1989		
1990-1999		
Treatment		
Surgery only		
Any radiation therapy ± surgery, no chemotherapy		
Any chemotherapy ± surgery, no radiation therapy		
Any radiation therapy or chemotherapy		
HCT		

Table 2b-f: Repeat Table 2a for SMNs and specific SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors). Additional SMN types to be considered if counts are sufficient.

Table 3a: Associations between sociodemographic and clinical characteristics and the false negative probability for ascertainment approaches for SNs overall (CCSS only, N=)

Characteristics	**Approach E** OR (95% CI)
Sex	
Female	Ref.
Male	
Race and ethnicity	
Hispanic/Latine	
Non-Hispanic Black	
Non-Hispanic White	Ref.
Other	
Age at last follow up (years)	
<30 years	Ref.
31-45 years	
>45 years	
Age at primary cancer diagnosis	
<10 years	Ref.
10 years or older	
Primary childhood cancer diagnosis	
Leukemia	Ref.
Hodgkin lymphoma	
Non-Hodgkin lymphoma	
Sarcoma	
Kidney tumors	
Neuroblastoma	
Central nervous system tumors	
Other	
Childhood cancer treatment decade	
1970-1979	Ref.
1980-1989	
1990-1999	
Treatment	
Surgery only	Ref.
Any radiation therapy ± surgery, no chemotherapy	
Any chemotherapy ± surgery, no radiation therapy	
Any radiation therapy or chemotherapy	
HCT	

Table 3b-f: Repeat Table 3a for SMNs and specific SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors). Additional SMN types to be considered if counts are sufficient.

Table 4: Comparison of standardized incidence ratios (SIRs) and absolute excess risks (AERs) of SNs by ascertainment approaches

SN type and ascertainment pproach	Obs	Exp	Person-years of follow-up	SIR (95% CI)	SIR P_{trend}	AER (95% CI)	AER P_{trend}
SN, overall							
CCSS-confirmed							
Approach E							
SMN, overall							
CCSS-confirmed							
Approach E							
Breast							
CCSS-confirmed							
Approach E							
Colorectal							
CCSS-confirmed							
Approach E							
Thyroid							
CCSS-confirmed							
Approach E							
CNS tumor							
CCSS-confirmed							
Approach E							
Soft tissue sarcoma							
CCSS-confirmed							
Approach E							
Bone tumor							
CCSS-confirmed							
Approach E							
Leukemia							
CCSS-confirmed							
Approach E							
Lymphoma							
CCSS-confirmed							
Approach E							
Melanoma							
CCSS-confirmed							
Approach E							
Other solid tumor							
CCSS-confirmed							
Approach E							

Table 4a: Same as table above, but broken out by attained age and treatment decade.

Table 4b: Same as table above, but broken out by childhood cancer diagnosis.

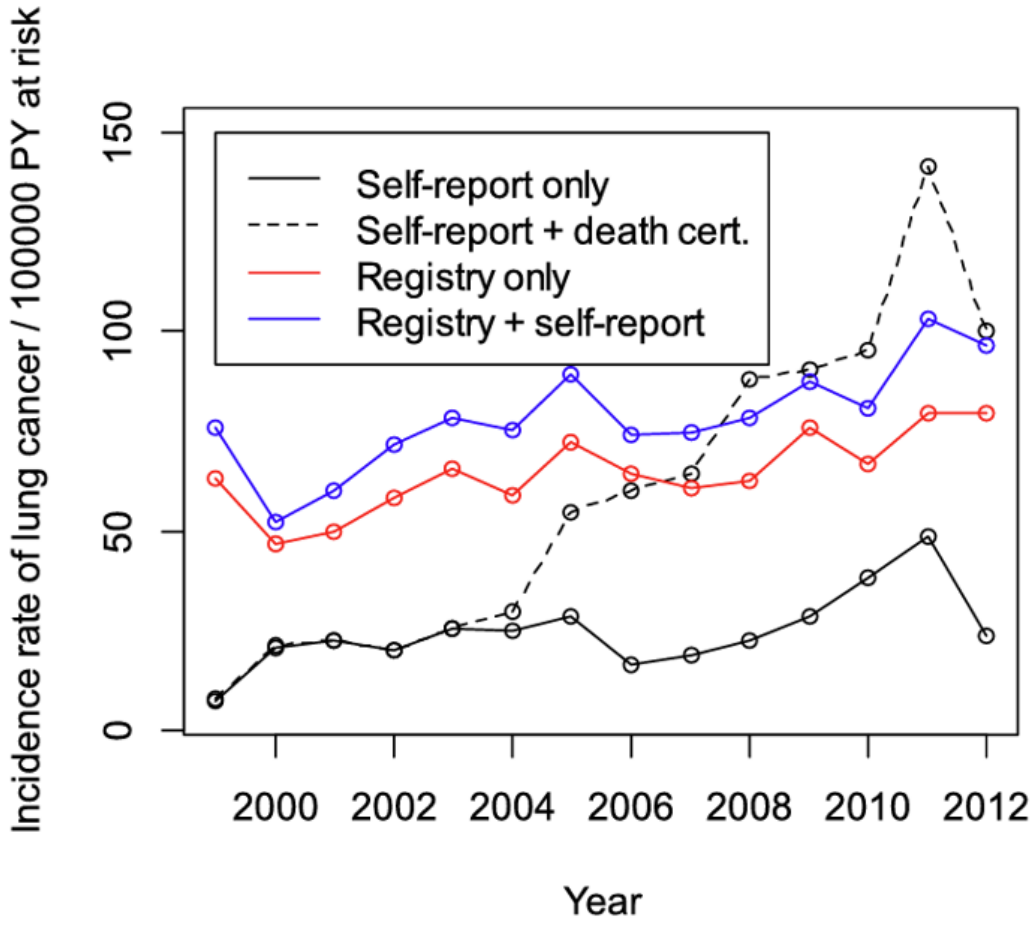

Figure 1: SN incidence rates (per 10,000 person-years) in the CCSS-eligible population

- Comparison of incidence rates (per 10,000 person-years) across ascertainment approaches for SNs overall, SMNs overall, and SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors) in a single figure or multi-panel figure.
- *Example figure above; corresponding tables will be included.*

Figure 1a: Same as F1 above, but broken out by attained age and treatment decade.

Figure 1b: Same as F1 above, but broken out by childhood cancer diagnosis.

Figure 2: SN cumulative incidence in the CCSS-eligible population

- Cumulative incidence curves for SNs overall, SMNs overall, and for SMN types (breast, colorectal, thyroid, CNS, soft tissue sarcoma, bone tumors, leukemia, lymphoma, other solid tumors) in a single figure or multi-panel figure.
- *Corresponding tables with cumulative incidence in 5-year follow-up increments with 95% CIs by approach will be included.*

Figure 2a: Same as F2 above, but broken out by attained age and treatment decade.

Figure 2b: Same as F2 above, but broken out by childhood cancer diagnosis.

References

1. Liu D, Linet MS, Albert PS, et al. Ascertainment of incident cancer by US population-based cancer registries versus self-reports and death certificates in a nationwide cohort study, the US Radiologic Technologists Study. *American journal of epidemiology*. 2022;191(12):2075-2083.
2. Liu D, Linet MS, Albert PS, et al. Examining bias due to method of follow-up for cancer incidence in a large US cohort: Self-report versus registry linkage. *Annals of epidemiology*. 2025;107:44-50.
3. Schonfeld SJ, Valcarcel B, Meyer CL, et al. Cumulative incidence estimates for solid tumors after HCT in the CIBMTR and California Cancer Registry. *Blood Advances*. 2024;8(15):4102-4112.
4. Deapen D, Clerkin C, Howe W, et al. Virtual Pooled Registry-Cancer Linkage System: an improved method for ascertaining cancer diagnoses. *JNCI Monographs*. 2024;2024(65):191-197.
5. Turcotte LM, Liu Q, Yasui Y, et al. Temporal trends in treatment and subsequent neoplasm risk among 5-year survivors of childhood cancer, 1970-2015. *Jama*. 2017;317(8):814-824.
6. Turcotte LM, Liu Q, Yasui Y, et al. Chemotherapy and risk of subsequent malignant neoplasms in the childhood cancer survivor study cohort. *Journal of Clinical Oncology*. 2019;37(34):3310.
7. Armstrong GT, Chen Y, Yasui Y, et al. Reduction in late mortality among 5-year survivors of childhood cancer. *New England Journal of Medicine*. 2016;374(9):833-842.