

## Pediatric Cancer Survivorship Research: Experience of the Childhood Cancer Survivor Study

Wendy M. Leisenring, Ann C. Mertens, Gregory T. Armstrong, Marilyn A. Stovall, Joseph P. Neglia, Jennifer Q. Lancot, John D. Boice Jr, John A. Whitton, and Yutaka Yasui

### ABSTRACT

The Childhood Cancer Survivor Study (CCSS) is a comprehensive multicenter study designed to quantify and better understand the effects of pediatric cancer and its treatment on later health, including behavioral and sociodemographic outcomes. The CCSS investigators have published more than 100 articles in the scientific literature related to the study. As with any large cohort study, high standards for methodologic approaches are imperative for valid and generalizable results. In this article we describe methodological issues of study design, exposure assessment, outcome validation, and statistical analysis. Methods for handling missing data, intrafamily correlation, and competing risks analysis are addressed; each with particular relevance to pediatric cancer survivorship research. Our goal in this article is to provide a resource and reference for other researchers working in the area of long-term cancer survivorship.

*J Clin Oncol* 27:2319-2327. © 2009 by American Society of Clinical Oncology

From the Cancer Prevention and Clinical Statistics Programs, Fred Hutchinson Cancer Research Center, Seattle, WA; Emory University, Atlanta, GA; St Jude Children's Research Hospital, Memphis; Vanderbilt University Medical Center, Nashville, TN; The University of Texas M. D. Anderson Cancer Center, Houston, TX; University of Minnesota Medical School, Minneapolis, MN; International Epidemiology Institute, Rockville, MD; and the University of Alberta, Edmonton, Alberta, Canada.

Submitted November 19, 2008; accepted November 20, 2008; published online ahead of print at www.jco.org on April 13, 2009.

Supported by National Institutes of Health-National Cancer Institute Grants No. U24CA55727 and 5R01CA104666.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Wendy M. Leisenring, ScD, Cancer Prevention and Clinical Statistics, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, D5-360, PO Box 19024, Seattle, WA 98109; e-mail: wleisenr@fhcrc.org.

© 2009 by American Society of Clinical Oncology

0732-183X/09/2714-2319/\$20.00

DOI: 10.1200/JCO.2008.21.1813

### INTRODUCTION

The remarkable successes in treating childhood cancers over the past 40 years have made it imperative to study the long-term outcomes after pediatric cancer and its associated intensive treatments. The Childhood Cancer Survivor Study (CCSS) is one of the first large cohorts of pediatric cancer survivors to be formed and followed successfully. The purpose of this article is to summarize many of the methodological lessons learned over the past 15 years of our experience carrying out survivorship research.<sup>1-3</sup> Our intent is to provide a comprehensive document that will prove to be a resource to other researchers in the field.

The topics covered are divided into four main areas. First, we describe the study design and issues involved in contacting and recruiting members of the cohort. This includes the recent changes required as a consequence of the Health Insurance Portability and Accountability Act Privacy Rule (HIPAA). Second, we describe methods for obtaining high quality treatment data that have proven to be vital to the success of CCSS. The third important topic deals with outcome data that are necessarily gathered via self-report mechanisms. For certain outcomes, we have undertaken a validation process to ensure data quality and this is described. Finally, we summarize a number of topics related to statistical analysis of the data, such as handling missing data and competing risks analysis.

### STUDY DESIGN AND COHORT RECRUITMENT

Much of the early research conducted on long-term survivors of childhood cancer was carried out as single-institution studies. These early studies documented the occurrence of a variety of late effects of therapy, but were limited in scope due to small number of participants and homogeneity in treatment. Also, because the concept of survivorship was still fairly new, many of these single institution studies only followed participants 5 to 10 years after diagnosis. Early research on survivorship was also performed within the cooperative clinical trials groups. However, because of the therapeutic intent of these protocols, they were not designed for long-term follow-up and often suffered from incomplete participant ascertainment. Thus, much of the early research was restricted to events in the first 5 to 10 years from diagnosis on limited participant populations. Finally, lack of appropriate comparison populations often made interpretation of rates and effect sizes difficult.

The CCSS was designed as a multicenter hospital-based retrospective cohort study with longitudinal follow-up.<sup>1</sup> The CCSS has been estimated to have captured approximately 40% to 45% of 5-year survivors diagnosed between 1970 and 1986 in the United States and, in doing so, has established a cohort of sufficient size and heterogeneity to overcome many of these previous limitations.

The CCSS identified and recruited all survivors meeting eligibility criteria (Table 1) at 26 institutions in the United States (n = 25) and Canada (n = 1). Ascertainment and registration of participants occurred at each center using a comprehensive unified protocol to achieve complete ascertainment of eligible participants. Of the 22,124 participants initially registered with the CCSS Coordinating Center (Memphis, TN), 20,691 were confirmed to be eligible.

Contact began in 1992, and after an initial letter from the treating institution, a letter from the CCSS Coordinating Center containing the baseline survey, informed consent, and a request for medical record release was sent to each eligible patient (or parent if the patient was younger than 18 years at the time of contact). If no response was received, a postal reminder was sent, ultimately followed by a telephone call from the Coordinating Center by a trained telephone interviewer who provided the option of completing the baseline survey by telephone. For eligible participants who were known to have died after achieving 5-year survivorship, their next of kin were contacted and asked to complete the baseline questionnaire. For 7,030 participants unable to be located using the address obtained from the treating institution, a tracing protocol was completed by a national survey research firm (Westat Inc, Rockville, MD). Tracing was successful for 4,188 persons (60%).<sup>2</sup>

Overall, 3,058 participants (15%) could not be located and were lost to follow-up, 3,205 (15%) declined participation, and 65 participants were unable to participate due to language difficulties. Ultimately, 14,357 eligible participants completed the baseline questionnaire, representing 69% of the total eligible population (approximately six other patients agreed to participate and did so only for subsequent questionnaires). Since the time of the baseline questionnaire, the CCSS has completed three additional follow-up surveys of this cohort, achieving participation rates between 77% and 81% among those participants who were eligible and successfully contacted (Fig 1). In addition, this cohort has been contacted for participation in several other survey-based investigations regarding barriers to health care, sleep and fatigue, use of mammography, health information, and men's and women's specific health issues, in addition to questionnaires specifically targeting quality of life in survivors of bone tumors, and health behaviors in survivors during adolescence.

Comparisons of available demographic and cancer-related characteristics between participants and nonparticipants at the initial base-

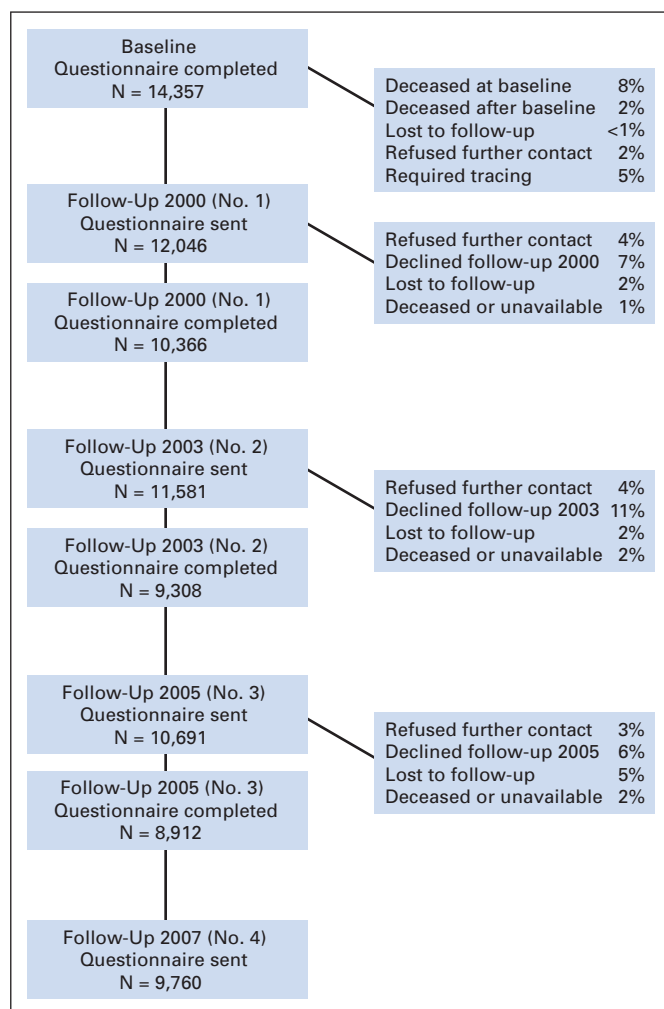


Fig 1. Participation and contact in the Childhood Cancer Survivor Study.

line questionnaire showed that the only significant difference between these groups was vital status. That is, the next-of-kin relatives of patients who died more than 5 years after diagnosis were less likely to agree to participate than patients (or parents of patients) who were still alive.<sup>1</sup> Among participants who agreed to participate at the baseline questionnaire, we have also evaluated whether demographic and cancer-related characteristics differed between participants and nonparticipants at subsequent questionnaires and have determined that, while differences are moderate in size (< 10% increase), the study retains more female, White race, college-educated, higher-income, and older participants (data not shown).

For certain outcomes, siblings provide a readily available comparison population. Among a random sample of participating survivors, the sibling closest in age to the survivor was contacted requesting their participation in this cohort study. Identical to survivors, participation included informed consent, a request for medical record release, and completion of a 24P-page baseline questionnaire. Of the 5,857 siblings randomly selected for participation, 3,899 (67%) completed the baseline questionnaire. Siblings who participated in CCSS were more likely to be older, female, and white race than participating survivors, though the differences are

Table 1. Childhood Cancer Survivor Study Eligibility Criteria

Criteria
1. Diagnosis and initial treatment of leukemia, CNS malignancy (excludes meningioma and craniopharyngioma), Hodgkin's lymphoma, non-Hodgkin's lymphoma, neuroblastoma, soft tissue sarcoma, kidney cancer, or bone cancer (specific ICD-O codes defining eligible cases within each diagnostic category are provided at <a href="http://www.stjude.org/ccss">www.stjude.org/ccss</a> )
2. Diagnosis date between January 1, 1970, and December 31, 1986
3. Age younger than 21 years at diagnosis
4. Alive 5 years from the date of diagnosis, regardless of disease or treatment status
5. English or Spanish speaking, because of logistics of questionnaires and interviews
6. Resident of United States or Canada at the time of initial follow up contact

Abbreviation: ICD-O, International Classification of Diseases-Oncology.

relatively small (Table 2). Thus, comparative analyses are always adjusted for these factors.

The CCSS is currently expanding its cohort to include patients diagnosed between 1987 and 1999. This expansion will provide important information regarding late effects of more modern therapeutic protocols, and will employ similar methods of data ascertainment to assure comparability of data with the original cohort. However, a number of challenges to recruitment of this cohort in the current era have been identified. Most importantly, modern privacy laws, including the HIPAA, place limits on contact with eligible participants until their consent for study participation is obtained. In addition, survivors from this era who are age 20 to 39 years are a highly mobile group and not as available or as responsive to contact by traditional mail mechanisms or traditional land-line telephone. Successful recruitment of this population will require innovative use of electronic methods of contact including e-mail and Web-based modalities.

### EXPOSURE ASSESSMENT

Assessment of therapeutic exposures has been critical to the correct attribution of late outcomes. The CCSS used a methodology of case-by-case chart abstraction for each member of the cohort. Individual abstracters for each center were centrally trained to carry out abstraction of chemotherapy, surgery, and radiotherapy for those consenting cohort members using a standardized medical records abstraction form (MRAF) and treatment data were abstracted from the medical record for each case. In the MRAF the abstracter was asked to specify the dates of therapy covered by that abstraction form, the protocol the patient was treated on (if applicable), and then provide specific data for the treatments of interest (ie, chemotherapy, radiation therapy, and surgery). An individual MRAF form was completed for each treatment plan, but it was recognized that the medical record may not be complete or that some treatments would have been given outside of the participating CCSS centers. In those instances, abstracters were asked to infer doses given for patients and remark on the incompleteness of the dose information so that the incompleteness could be accounted for in the subsequent analyses.

As many pediatric patients were treated on cooperative group studies (Children's Cancer Group and Pediatric Oncology Group during the treatment era of this cohort) expected doses were calculated for the most common protocols used by both groups. As a quality control measure, abstracted treatment information for each case was compared with the calculated expected doses within each protocol. Outliers were returned to the abstracter to double check the medical records and verify data.

### Chemotherapeutic Agents

A yes/no evaluation of exposure was asked for each of 42 common chemotherapeutic agents used during this time. For 22 specific agents of the 42, the quantitative dose was abstracted as outlined above. These agents included anthracyclines, alkylating agents, epipodophyllotoxins, and platinum compounds.

For many drugs, the cumulative dose can be used as a measure of total exposure. However, when a number of agents fall into a single class, such as anthracycline or alkylating agents, to enable a succinct assessment of exposure effects for the class, several methods were used. For anthracyclines, the cumulative dose of doxorubicin, daunomycin,

**Table 2.** Characteristics of Survivor and Sibling Participants at Baseline

Characteristic	CCSS Participants (n = 14,357)		Siblings (n = 3,899)	
	No.	%	No.	%
<b>Age at baseline, years</b>				
0-4	NA	NA	6	0.2
5-9	175	1.2	70	1.8
10-14	1,933	13.5	355	9.1
15-19	3,126	21.8	656	16.8
20-24	3,118	21.7	671	17.2
25-29	2,770	19.3	709	18.2
30-34	1,977	13.8	655	16.8
35-39	926	6.4	461	11.8
≥ 40	332	2.3	316	8.1
<b>Sex</b>				
Male	7,713	53.7	1,875	48.1
Female	6,644	46.3	2,024	51.9
<b>Race/ethnicity</b>				
White not Hispanic	11,942	83.2	3,414	87.6
Black not Hispanic	668	4.7	102	2.6
American Indian/Alaska Native	88	0.6	15	0.4
Asian or Pacific Islander	167	1.2	41	1.1
Hispanic	397	2.8	60	1.5
Other	1,045	7.3	130	3.3
Unknown	50	0.3	137	3.5
<b>Age at diagnosis, years</b>				
0-4	5,754	40.1	NA	NA
5-9	3,200	22.3	NA	NA
10-14	2,912	20.3	NA	NA
15-20	2,491	17.4	NA	NA
<b>Diagnosis</b>				
Acute lymphoblastic leukemia	4,151	28.9	NA	NA
Acute myeloid leukemia	356	2.5	NA	NA
Other leukemia	320	2.2	NA	NA
Astrocytomas	1,184	8.2	NA	NA
Medulloblastoma, PNET	381	2.7	NA	NA
Other CNS tumors	315	2.2	NA	NA
Hodgkin's lymphoma	1,927	13.4	NA	NA
Non-Hodgkin's lymphoma	1,083	7.5	NA	NA
Kidney tumors	1,258	8.8	NA	NA
Neuroblastoma	951	6.6	NA	NA
Soft tissue sarcoma	1,243	8.7	NA	NA
Ewings sarcoma	403	2.8	NA	NA
Osteosarcoma	733	5.1	NA	NA
Other bone tumors	52	0.4	NA	NA
<b>Vital status at baseline</b>				
Alive	13,182	91.8	NA	NA
Dead	1,175	8.2	NA	NA
<b>Treatment</b>				
Chemo + RT + surgery	5,550	38.7	NA	NA
Chemo + RT	1,459	10.2	NA	NA
Chemo + surgery	2,274	15.8	NA	NA
Chemo only	816	5.7	NA	NA
RT + surgery	1,485	10.3	NA	NA
RT only	33	0.2	NA	NA
Surgery only	910	6.3	NA	NA
Surgery + RT	12	0.1	NA	NA
No medical records	1,818	12.7	NA	NA

Abbreviations: NA, not applicable; PNET, primitive neuroectodermal tumor; Chemo, chemotherapy; RT, radiation therapy.

and idarubicin (multiplied by three) were summed. The cumulative platinum compound exposure was calculated by summing the cisplatin and carboplatin (carboplatin divided by four) exposures.<sup>4</sup> Given the wide variety of alkylating agents used, a summary variable was created as follows: First, the dose of each agent was abstracted. Across all patients exposed in the cohort, the dose (standardized by body-surface area) was divided into tertiles of exposure for the individual agent. Each participant was assigned an exposure code of 0 (no exposure), 1, 2, or 3 for each alkylating agent he/she had received.<sup>5</sup> The cumulative score for each individual was summed, and then, across the cohort, these summed exposures were again assigned tertiles. This resulted in individual alkylating agent exposure scores ranging from 0 to 3 for each cohort member which can be utilized in analyses.

### Surgical Procedures

Surgical procedures were also abstracted and entered into the MRAF. Each procedure requiring general anesthesia was abstracted with the exception of procedures for the placement of vascular access devices. The date, name of procedure, and International Classification of Diseases (9th revision, clinical modification) code were requested for each surgery performed.

### Radiation Dosimetry

Radiation therapy was also indicated in the MRAF. Abstracters were, however, asked only whether the participant had received radiation, the dates of treatment, and the names of the radiation oncologist and facility where it was given. The abstracters then copied records from the radiation oncology department, including treatment plans, patient placement photographs, daily treatment logs, and radiation summaries; these records were sent to the Radiation Physics Center at

the M. D. Anderson Cancer Center (Houston, TX), where the records were scanned and stored in an image database.

The aim was to provide for each patient in a study the radiation absorbed dose to organs or anatomic sites appropriate to the outcome under investigation. Basic treatment information was abstracted for the entire cohort and entered into a database. This first-level abstracting included first and last date of treatment, body region treated, beam energy, treatment field size, configuration and laterality, and total treatment dose. The basic coding is useful for study planning and sufficient for many cohort analyses. Table 3 shows body regions treated for all patients in the database who had radiation therapy, stratified by disease.

Case/control studies and some cohort analyses with specific interest in radiation exposure effects require additional record review, with more detailed coding, in particular where the organ or anatomic site of interest was shielded during treatment (eg, ovaries, testes, breasts, kidneys, or eyes). Dose to the site of interest for each patient is estimated by applying out-of-beam data measured in a water phantom to an age-specific mathematical phantom.<sup>6,7</sup> Detailed dosimetry is provided for each of these studies, depending on the regions of interest and study population determined by the investigators.

## OUTCOMES

### Validation of Self-Reported Medical Outcomes

Validation of medical outcomes has been an important topic in the CCSS. Due to the increase in the personnel effort and cost required to conduct validation through medical records, however, careful consideration has to be made as to what major end points require this

**Table 3.** Childhood Cancer Survivor Study Radiotherapy: Anatomic Regions With One or More Radiation Fields

First Malignant Neoplasm	Sites Treated																					
	Brain		Head (not brain)		Cranio-Spinal*		Neck		Chest		Spine		Abdomen		Pelvis		Limb		Total-Body Irradiation		Region Unknown	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Leukemia (n = 2,927)	2,666	91.1	77	2.6	442	15.1	69	2.4	93	3.2	443	15.1	77	2.6	379	12.9	26	0.9	244	8.3	56	1.9
CNS (n = 1,166)	1,096	94.0	70	6.0	416	35.7	63	5.4	26	2.2	428	36.7	24	2.1	26	2.2	3	0.3	1	0.1	32	2.7
Hodgkin's lymphoma (n = 1,584)	15	0.9	190	12.0	1	0.1	1,484	93.7	1,435	90.6	1	0.1	989	62.4	609	38.4	34	2.1	4	0.3	30	1.9
Non-Hodgkin's lymphoma (n = 639)	178	27.9	150	23.5	39	6.1	239	37.4	264	41.3	41	6.4	152	23.8	160	25.0	48	7.5	17	2.7	13	2.0
Kidney tumor (n = 694)	4	0.6	1	0.1			16	2.3	263	37.9			673	97.0	342	49.3	3	0.4			9	1.3
Neuroblastoma (n = 414)	41	9.9	35	8.5	11	2.7	72	17.4	171	41.3	15	3.6	215	51.9	115	27.8	10	2.4	16	3.9	8	1.9
Soft-tissue tumor (n = 694)	64	9.2	272	39.2	6	0.9	120	17.3	124	17.9	7	1.0	84	12.1	191	27.5	131	18.9	1	0.1	15	2.2
Bone tumor (n = 388)	30	7.7	20	5.2	5	1.3	24	6.2	168	43.3	5	1.3	47	12.1	95	24.5	185	47.7	3	0.8	9	2.3
Total (N = 8,506)	4,094	48.1	815	9.6	920	10.8	2,087	24.5	2,544	29.9	940	11.1	2,261	26.6	1,917	22.5	440	5.2	286	3.4	172	2.0

NOTE. Includes radiotherapy at any time. Rows sum to more than 100% due to some patients receiving treatment to multiple regions.

\*Cranio-spinal added as a category for patients who were treated to the brain and spine. Brain and spine may not have been treated at the same time or to the same dose. Patients with cranio-spinal treatments also are counted in both the "brain" and "spine" categories.



additional effort. At the conception of the CCSS, six primary hypotheses were postulated to be addressed: excess risk of mortality; risk of a therapy-related subsequent cancer; risk of clinically apparent cardiopulmonary events; loss of fertility, adverse pregnancy outcomes, and abnormalities in offspring; distinct patterns of family history of cancer; and increased risk of adverse health events due to health behaviors. Of these hypotheses, the first four listed outcomes were selected for validation in this study.

Vital status and the cause of death were determined through the National Death Index (NDI). There is extensive documentation of the advantages and limitations of the use of the NDI, which is covered in another article in this issue.<sup>8</sup> The remaining three outcomes of interest within CCSS (subsequent cancers, cardiopulmonary outcomes, and adverse pregnancy outcomes) are described in detail herein.

The validation procedure used within CCSS, which has been successful in other large epidemiologic studies, is depicted in Figure 2. Medical outcome data were collected using a self-report survey sent to the home of the eligible participant. A HIPAA form requesting release of medical records was also requested from the participant. On return, a request for photocopies of relevant medical records is then made to the hospital/clinic where the participant was diagnosed for this condition; medical records are reviewed and data coded by trained abstracters/physicians.

Validation through medical records of self-reported medical outcomes from mailed surveys have significant limitations in our current medical care environment. One change to the procedures was brought about by the enactment of the HIPAA Privacy Rule during 2001 and 2002 (modified rule). During recruitment for CCSS, we were able to obtain medical release on 93% of the survivor participants. We subsequently needed to obtain signed HIPAA release for future medical record validation. Although we have ultimately been successful in obtaining these consents for 95% of our participating participants, accomplishing this required significant added resources. Secondly, as

these survivors age and become adults, their medical care has transitioned from the pediatric institutions where they were treated for their primary diagnosis to adult care facilities. Because of this transition, and the constant change in health care providers within the current medical system, collection of records from such facilities can be costly and inefficient.

Medical records are useful for identifying false-positive self-reported outcomes; however, it is difficult to identify false-negative outcome events that are not reported by the respondent. In our experience with validation of self-reported outcomes, concordance between self-report and medical records was good for well-known complications that have clear diagnostic criteria, such as the occurrence of subsequent cancers, and for records where the participant had good recollection of where they were seen for the condition, such as pregnancy records and place of delivery. Conditions, however, with nonestablished diagnostic criteria such as cardiac outcomes demonstrated a lower level of agreement and the ability to successfully collect records.

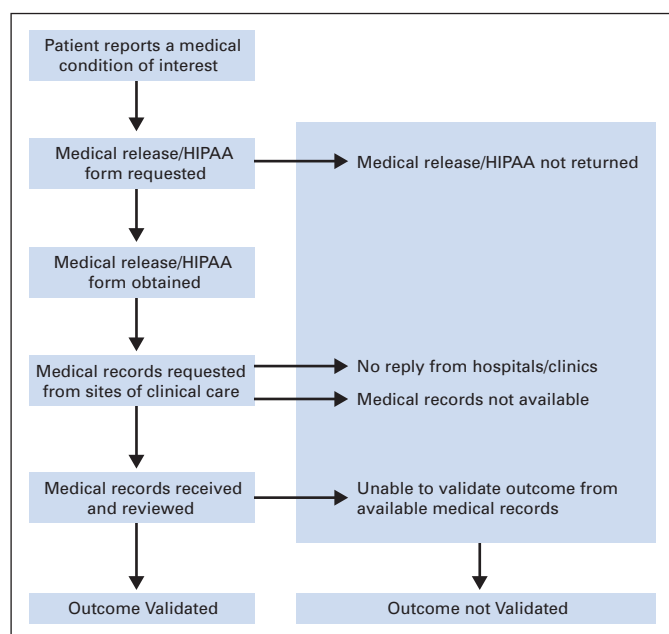
### Subsequent Neoplasm Validation

Subsequent neoplasms (SNs) were initially identified by self-report of any relapse or recurrence of their original cancer and/or the occurrence of a new cancer after treatment for their primary malignancy. The name of the hospital where the subsequent cancer had been diagnosed was also requested. All positive responses were screened by a CCSS investigator (J.P.N.), and those responses considered likely or possible SNs were forwarded to the CCSS Pathology Center (Columbus, OH) for verification. Reports of late recurrences of the original cancer (10 years or more after the original diagnosis) were also forwarded for verification. For all positive responses from individuals who signed a medical release, a copy of the pathology report was requested. Returned reports were reviewed by the CCSS pathologist for inclusion or exclusion in the study. Data collected included the specific type of SN, date of diagnosis, and location of tumor(s). If a pathology report could not be obtained, the patient and/or parent response or death certificate and/or other institutional records were reviewed to determine the presence of an SN.

At the time of this report, we had reviewed and verified a total of 2,508 SN events using the above methodology. Among these, 2,196 were verified from the pathology report, and an additional 17 were confirmed from death certificates. The remaining 295 were determined to be valid using participant or proxy responses or other sources as described above: 154 of these neoplasms were in participants who had not signed a medical release.

### Adverse Pregnancy Outcomes

To study adverse pregnancy outcomes and possible germline mutagenesis, we evaluated self-reported genetic and congenital diseases among the approximately 6,100 offspring of survivors and the 3,100 offspring of sibling controls. The self-administered questionnaire included questions on pregnancy histories, live births, stillbirths, miscarriages, abortions, cancers, birth defects, and hereditary conditions. Genetic disease included cytogenetic abnormalities, single-gene birth defects, and simple malformations. The approach to validate or confirm the self-reported conditions began with an initial review, including family history, by a cancer geneticist. A decision was made as to whether the self-reported condition could be accepted, rejected, or



**Fig 2.** Strategy for validation of self-reported outcomes. HIPAA, Health Insurance Portability and Accountability Act Privacy Rule.

that additional information was required. In instances where additional information was required, individualized scripts or questions were prepared for each participant, who was then contacted by CCSS staff to provide additional clarification of the self-report and/or to obtain a medical release for medical records.

All available information on the self-reported condition was then reviewed by a three-person panel to reach a consensus decision. The final decision could be: accept; accept but not count because the condition could be explained by family history or nongenetic factors; or reject. Validated genetic and congenital diseases occurred in 157 (2.6%) of the children of survivors, compared with 111 (3.6%) of the children of sibling controls. There were no apparent differences in the proportion of offspring with cytogenetic syndromes (seven in case offspring, six in sibling offspring), single-gene defects (14 and eight), or simple malformations (136 and 97). Analyses based only on the self-reported genetic diseases were reassuring<sup>9</sup> and were then confirmed through the validation procedure.<sup>10</sup>

### Cardiac Outcome Validation

For survivors who reported a specific cardiopulmonary outcome and were still alive at the time of contact, an additional stage of validation was incorporated which consisted of a series of telephone-based questions (telephone script) to further document the specifics of selected self-reported adverse cardiopulmonary events and to determine where the participant received treatment for the reported outcomes. Participants contacted by telephone were also asked to sign a HIPAA form and return it to the CCSS Coordinating Center. Once received, medical records were requested from the physicians listed and were returned to the CCSS Coordinating Center. The first 100 medical records that were returned were reviewed independently by two physicians. Consensus among the physician's validation was reached when the two differed and a standardized protocol was developed to determine validation of each condition. Subsequent records were reviewed and validated by one physician.

As an example, a flowchart summarizing the validation of 292 survivors who reported congestive heart failure (CHF) is detailed in Figure 3. Among participants for whom validation was successfully carried out, CHF was confirmed for 83% and 67%, and determined not to have occurred in 11% and 9%, for telephone script and medical record validation, respectively. Notably, among participants for whom medical records were received, 25 (24%) did not have enough information in the records to determine CHF status. A key difficulty identified with this validation process was that validation either by script or by medical records was only possible in a relatively small subset of participants who self-reported the outcome (65% and 35% for script and medical records, respectively). Of further concern is the fact that this subset of participants was by-and-large limited to those who were alive at the time of validation. In particular, 64% of those participants not validated by telephone script had died (as opposed to 14% of those validated). Since CHF is a potentially fatal condition, we can only assume that a relatively large proportion of the participants who died would have been confirmed had we a means of validating them. Confirmation of cause of death via death certificate does not provide a specific enough cardiac diagnosis, nor would it identify patients who had a cardiac condition and subsequently died of a different cause. Thus, for cardiac conditions, it has not been feasible to utilize only validated outcomes in an unbiased fashion for

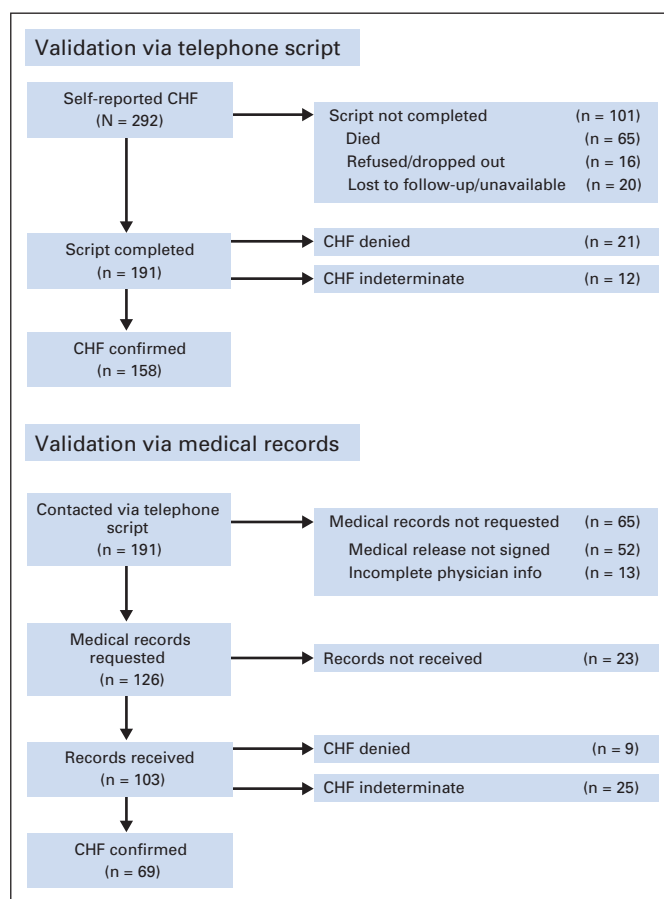


Fig 3. Validation experience for congestive heart failure (CHF) via telephone script and medical record review.

analyses. Instead we have relied on self-reported outcomes for current analyses.

## STATISTICAL METHODOLOGY

In the course of analyzing the data from the CCSS over the past 10 or more years, we have needed to carefully consider a number of key statistical issues. Many of these issues are generalizable to other settings, although they have the common theme of being specifically applicable to survivorship research and thus are useful tools for anyone else carrying out statistical analyses on similar data.

### Long-Term Survivor Cohort Definition: Impact on Analyses

The requirement that participants have attained 5-year survivorship for eligibility into the CCSS cohort has implications on late events that can be utilized in valid and generalizable analyses. Because our questionnaires are typically worded to ask the first age at which an event occurred, it is not unusual for the first event time to be before the cohort inception time point of 5 years after diagnosis. It is tempting for researchers to examine the rates or carry out time-to-event analyses that incorporate these events. However, caution must be taken since potential patients who died during those first 5 years are not part of the CCSS cohort and have been removed from the denominator: hence,

the full cohort of patients who were at risk for events in the first 5 years are not all included in a survivorship cohort. As such, any time-to-event analyses would not accurately represent true rates or relative risks in that time period and would not be generalizable to any existing prospective population. In statistical terms, this analysis would violate standard principles of time-to-event analyses by conditioning on the future event of survival to 5 years. The most appropriate way of handling time-to-event analyses in a 5-year survivor cohort is thus to begin analyses at the 5-year postdiagnosis time point, only prospectively considering events that occur after the inception of the cohort. One way that outcomes occurring before 5 years could be reported would be as the proportion (or prevalence) of participants who had experienced at least one event by the time the cohort was formed at 5 years. However, for the reasons stated above, one should avoid use of time-dependent rates, or time-to-event analyses. The key point must be emphasized; these results are only generalizable to the population of participants who have survived at least 5 years after their diagnosis of primary cancer.

### Accounting for Correlation Between Survivors and Siblings

A statistical issue that needs to be addressed in any analysis that incorporates both survivors and siblings is the intradependence of the outcomes from members of the same family. Since siblings would be expected to have more similar health outcomes than a randomly selected individual from the general population, standard assumptions of independence required for most statistical analyses are violated. In a correlated data setting such as this, unadjusted statistical methods typically lead to incorrect estimation of the variability of measures of association and thus, resulting naïve *P* values and CIs are also incorrect. To appropriately handle this issue in analyses, a generalized estimating equation approach with robust variance estimates can be used in analyses. These methods have been developed for use with generalized linear models (eg, logistic and Poisson regression)<sup>11</sup> as well as Cox proportional hazards models.<sup>12</sup> The idea behind the methodology is that it incorporates an appropriate adjustment that accounts for the intrafamily correlation and assures that inferences are valid. Other methods for handling the correlation between survivors and siblings that we have used are generalized linear mixed models<sup>13</sup> and bootstrapping approaches.<sup>14</sup>

### Impact of Attained Age on Risk of Disease: Appropriate Methods for Analyses

The risk of many key outcomes in long-term survivorship studies, especially those of childhood-disease survivors, can be highly dependent on the attained age of the participant and thus, attained age should be incorporated in a meaningful way into analyses. Indeed, if time since diagnosis is used as the time scale for relative risk analyses, rather than age of participant, for example, this can lead to flawed conclusions. In a cohort such as CCSS, participants who enter the cohort between the ages of 5 and 20 years, with 20 years of follow-up, will be age 25 to 40 years at last contact, an age range in which risks of some chronic diseases increase considerably with age. If time since diagnosis were used as the scale for analyses, then participants who were age 25 years would be treated on equal footing with participants who were age 40 years, two groups who might have markedly different risk of disease. As an example, Figure 4 illustrates the difference in

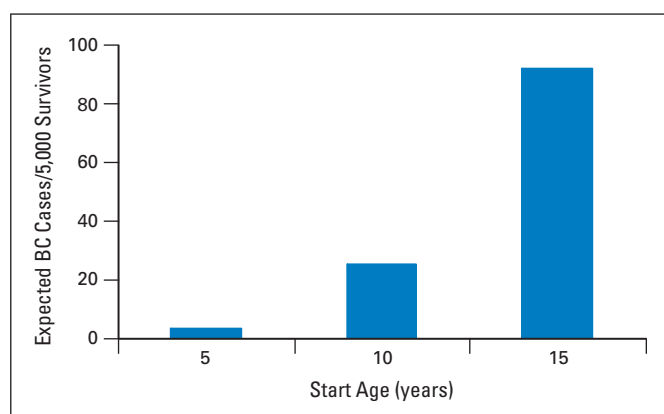


Fig 4. Expected numbers of breast cancers (BCs) in three age cohorts after 20 years of follow-up.

expected number of breast cancer cases between three groups of 5,000 participants diagnosed at ages 5, 10, and 15 years, respectively, based on Surveillance, Epidemiology, and End Results (SEER) incidence rates, assuming these participants had the same rates of breast cancer as the general population.<sup>3,15</sup> Without appropriately taking attained age into account, an analyst might erroneously conclude that girls treated at older ages were more likely to develop breast cancer. One can ameliorate the impact of attained age in descriptive analyses by utilizing standardized incidence rates, adjusted for age. Moreover, for multivariable regression analyses, the use of age as the time scale in a Cox proportional hazards model is an elegant way to directly adjust for changes in risk with age, without needing to incorporate age as a covariate or assuming a specific form for its effect. In this setting, participants enter the analysis at the age at which they enter the cohort and are followed until their attained age at end of follow-up. Another method for analysis is to use Poisson regression models to directly model standardized incidence ratios (SIR) in multivariable models. This method uses external reference age-specific rates such as the rates from SEER to adjust for the effects of attained age on risk of disease. Both these methods provide valid ways of adjusting for the effects of age on outcome and are useful tools for a long-term survivorship data analyst.

### Cumulative Incidence for Nondeath Outcomes

Most health outcomes of interest are reported using time-to-event analyses and results are often illustrated with figures displaying their cumulative probability over time. Because cohort participants could die any time before that outcome, an analysis of nondeath outcomes must appropriately consider death as a competing risk event when evaluating the probability of these outcomes. Readily available software provides Kaplan-Meier methodology that can be erroneously used in such situations. As described elsewhere, since Kaplan-Meier estimates treat time of death exactly the same way as a censored outcome, the estimates can become overly inflated when many deaths occur during the follow-up period.<sup>16,17</sup> The appropriate methodology in this setting is to utilize cumulative incidence estimates, which handle deaths differently from censored observations.<sup>18</sup> With the long follow-up period and high mortality rate present in the survivor population, this is an important issue for any analyst to address appropriately in order to obtain valid estimates of cumulative probability.

## Missing Data

In any epidemiological study, missing data raise concerns as a potential source of bias. In the CCSS, we have dealt with two types of missing data. The first type arises due to nonparticipation of eligible patients, where all data from the surveys and MRAF could not be obtained. The second type occurs among participants, where some of the survey items or MRAF elements were not answered or collected: this includes survivors who participated in the surveys but did not consent for medical record release.

The first type is difficult to deal with as we have no data on nonparticipants' outcomes and exposures, except their cancer type, age at diagnosis, sex, diagnosis year, and treating institution, which were collected for the initial eligibility establishment. We have compared and reported these characteristics between the participants and nonparticipants, and also made an aggregated-level comparison of MRAF data between the two groups.<sup>1</sup> The CCSS mortality analyses<sup>19,20</sup> have been an exception. For these analyses, all eligible participants' vital status was ascertained by the public NDI data, and mortality-risk analyses by cancer type, age at diagnosis, sex, and diagnosis year were conducted. When assessing treatment effects on mortality risk, however, we used multiple imputation under the assumption of missing at random of treatment data, given the known characteristics.<sup>21</sup> Note that the multiple-imputation approach imputes missing data multiple times to construct multiple complete datasets, runs an identical analysis with each of the complete datasets, and makes statistical inference using results from the multiple analyses. This is in contrast to the single-imputation approach, where the imputed and observed data are not distinguished in the single analysis of the complete data set: in multiple imputation, the variability across the multiple complete datasets appropriately reflects the uncertainty in the missing data.<sup>19,20</sup>

For the second type of missing data, the frequencies in the CCSS are mostly no more than 10%, often 0% to 5%. In many CCSS analyses, therefore, we confirm that the extent of missing data in key outcome and exposure variables is small and proceed to perform complete case analyses.<sup>21</sup> When the extent of missing data is large (ie, the number of incomplete cases is appreciable), or when an adverse event occurrence was indicated in the survey but the age at the occurrence was not reported (thus, a complete-case analysis would bias time-to-event analysis), we used multiple imputation under the assumption of "missing at random." Specifically, we extensively used the multiple-imputation method of Taylor et al<sup>22</sup> with slight modifications for the cases where there were an appreciable number of participants who reported an adverse event of interest, but did not report their age at its first occurrence. This method employs piecewise, exponential models to describe the rate of development of each adverse event by relevant demographic, clinical, and treatment variables with possible interactions. Its model fitting uses an expectation-maximization algorithm before proceeding to multiple imputation.<sup>23</sup>

Currently, the Statistical Center (Seattle, WA) of the CCSS is constructing ten complete datasets of CCSS survey participants through an extensive application of multiple imputation so that those who answered the surveys but did not consent medical record release can be entered into analyses of treatment effects. This work involves (1) elicitation of clinical knowledge from pediatric and radiation oncologists on the treatments used from 1970 to 1986 by diagnosis type, age, treating institution, and calendar period; (2) construction of imputation models based on the elicited knowledge as well as statisti-

cally driven model selection; (3) imputation of missing data ten times using the models; and (4) checking the imputed data by pediatric and radiation oncologists to see if they are sensible clinically. Such central multiple-imputation of missing data to construct multiple complete datasets has been successfully used in other large epidemiologic studies.<sup>24,25</sup>

## SUMMARY AND FUTURE DIRECTIONS

This article summarizes many of the procedures and methodologies implemented in the successful conduct of the CCSS over the past 15 years. These represent our efforts to ensure that conclusions drawn are unbiased and generalizable to the larger population of long-term pediatric cancer survivors. Challenges that require further development of methodology will continue to arise with continued follow-up of the current CCSS cohort and as the more recently treated expansion cohort is incorporated. We already confront issues related to recruiting and maintaining contact with a younger, more mobile cohort. There is an urgent need for recruitment strategies that utilize modern means of contact (eg, via cell phones and the internet). As we continue to collect data on an aging population, the patterns of missing data will be monitored and documented to assess the need for multiple imputation strategies for additional specific data elements. In addition, the participation rates and demographics of the participating CCSS population will be regularly evaluated. Applying appropriate methodology to the data to adjust for under- or overrepresentation of certain subpopulations will be important if disparities develop. Continued efforts at maintaining high levels of participation will be a priority to reduce any potential biases and to maximize statistical power.

CCSS has been and will continue to be successful at its goals of better understanding and quantifying risks of sequelae to cancer and its treatment. As new knowledge is developed there will be more opportunity for focused interventional studies aimed at reducing the morbidity due to these outcomes. Typically, these will be in the form of screening interventions studies, which can require large numbers of participants to see a significant impact on patient survival or morbidity. This will require the development of efficient study designs, with accurately characterized high risk populations and well defined and meaningful end points that best utilize the available resources and answer the cogent questions.

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Wendy Leisenring

**Data analysis and interpretation:** Jennifer Lanctot

**Manuscript writing:** Wendy Leisenring, Ann Mertens, Gregory Armstrong, Marilyn Stovall, Joseph Neglia, Jennifer Lanctot, John Boice, John Whitton, Yutaka Yasui



## REFERENCES

1. Robison LL, Mertens AC, Boice JD, et al: Study design and cohort characteristics of the Childhood Cancer Survivor Study: A multi-institutional collaborative project. *Med Pediatr Oncol* 38:229-239, 2002
2. Mertens AC, Walls RS, Taylor L, et al: Characteristics of childhood cancer survivors predicted their successful tracing. *J Clin Epidemiol* 57:933-944, 2004
3. Yasui Y, Liu Y, Neglia JP, et al: A methodological issue in the analysis of second primary cancer incidence in long-term survivors of childhood cancers. *Am J Epidemiol* 158:1108-1113, 2003
4. Travis LB, Holowaty EJ, Bergfeldt K, et al: Risk of leukemia after platinum-based chemotherapy for ovarian cancer. *N Engl J Med* 340:351-357, 1999
5. Tucker MA, D'Angio GJ, Boice JD Jr, et al: Bone sarcomas linked to radiotherapy and chemotherapy in children. *N Engl J Med* 317:588-593, 1987
6. Stovall M, Weathers R, Kasper C, et al: Dose reconstruction for therapeutic and diagnostic radiation exposures: Use in epidemiological studies. *Radiat Res* 166:141-157, 2006
7. Stovall M, Donaldson SS, Weathers RE, et al: Genetic effects of radiotherapy for childhood cancer: Gonadal dose reconstruction. *Int J Radiat Oncol Biol Phys* 60:542-552, 2004
8. Armstrong GT, Liu Q, Yasui Y, et al: Late mortality among 5-year survivors of childhood cancer: A summary from the Childhood Cancer Survivor Study. *J Clin Oncol* 27:2328-2338, 2009
9. Mulvihill JJ, Strong LC, Robison LL: Genetic disease in offspring of survivors of childhood and adolescent cancer. Presented at the annual meeting of the American Society of Human Genetics, October 12-16, 2001, San Diego, CA (abstr 1219)
10. Mulvihill JJ, Munro H, Whitton JA, et al: Genetic disease in offspring of survivors of childhood and adolescent cancer. Presented at the annual meeting of the American Society of Human Genetics, October 23-27, 2007, San Diego, CA (abstr 2002F)
11. Liang K-Y, Zeger SL: Longitudinal data analysis using generalized linear models. *Biometrika* 73:13-22, 1986
12. Therneau TM, Grambsch PM: *Modeling Survival Data: Extending the Cox Model*. New York, NY, Springer-Verlag, 2000, pp 170-229
13. Breslow NE, Clayton DG: Approximate inference in generalized linear mixed models. *J Am Statist Assoc* 88:9-25, 1993
14. Efron B, Tibshirani RJ: *An Introduction to the Bootstrap*. New York, NY, Chapman & Hall, 1993, pp 1-16, 47, 92-104
15. Surveillance, Epidemiology, and End Results Program: *Cancer Incidence and Survival Among Children and Adolescents: United States SEER Program 1975-1995*. Bethesda, MD, National Cancer Institute, 1999
16. Gooley TA, Leisenring W, Crowley J, et al: Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Stat Med* 18:695-706, 1999
17. Gaynor JJ, Feuer EJ, Tan CC, et al: On the use of cause-specific failure and conditional failure probabilities: Examples from clinical oncology data. *J Am Statist Assoc* 88:400-409, 1993
18. Kalbfleisch JD, Prentice RL: *The Statistical Analysis of Failure Time Data*. New York, NY, John Wiley & Sons, 1980, pp 168-169
19. Mertens AC, Yasui Y, Neglia JP, et al: Late mortality experience in five-year survivors of childhood and adolescent cancer: The Childhood Cancer Survivor Study. *J Clin Oncol* 19:3163-3172, 2001
20. Mertens AC, Liu Q, Neglia JP, et al: Cause-specific late mortality among 5-year survivors of childhood cancer: The Childhood Cancer Survivor Study. *J Natl Cancer Inst* 100:1368-1379, 2008
21. Little RJA, Rubin DB: *Statistical Analysis with Missing Data* (ed 2). Hoboken, NJ, John Wiley & Sons, 2002, pp 14-17, 255-259
22. Taylor JM, Munoz A, Bass SM, et al: Estimating the distribution of times from HIV seroconversion to AIDS using multiple imputation: Multicentre AIDS Cohort Study. *Stat Med* 9:505-514, 1990
23. Mertens AC, Yasui Y, Liu Y, et al: Pulmonary complications in survivors of childhood and adolescent cancer: A report from the Childhood Cancer Survivor Study. *Cancer* 95:2431-2441, 2002
24. Heitjan DF, Little RJA: Multiple imputation for the fatal accident reporting system. *Appl Stat* 40:13-29, 1991
25. Arnold AM, Kronmal RA: Multiple imputation of baseline data in the cardiovascular health study. *Am J Epidemiol* 157:74-84, 2003

