

Evaluating Early Case Capture of Pediatric Cancers in Seven Central Cancer Registries in the United States, 2013

MARY PUCKETT, PhD^{a,b}
ANTONIO NERI, MD, MPH^b
ELIZABETH ROHAN, PhD, MSW^b
CASTINE CLERKIN, MS, CTR^b
J. MICHAEL UNDERWOOD, PhD^b
A. BLYTHE RYERSON, PhD, MPH^b
SHERRI L. STEWART, PhD^b

ABSTRACT

Objective. Cancer is the second-leading cause of death in children, but incidence data are not available until two years after diagnosis, thereby delaying data dissemination and research. An early case capture (ECC) surveillance program was piloted in seven state cancer registries to register pediatric cancer cases within 30 days of diagnosis. We sought to determine the quality of ECC data and understand pilot implementation.

Methods. We used quantitative and qualitative methods to evaluate ECC. We assessed data quality by comparing demographic and clinical characteristics from the initial ECC submission to a resubmission of ECC pilot data and to the most recent year of routinely collected cancer data for each state individually and in aggregate. We conducted telephone focus groups with registry staff to determine ECC practices and difficulties in August and September 2013. Interviews were recorded, transcribed, and coded to identify themes.

Results. Comparing ECC initial submissions with submissions for all states, ECC data were nationally representative for age (9.7 vs. 9.9 years) and sex (673 of 1,324 [50.9%] vs. 42,609 of 80,547 [52.9%] male cases), but not for primary site (472 of 1,324 [35.7%] vs. 27,547 of 80,547 [34.2%] leukemia/lymphoma cases), behavior (1,219 of 1,324 [92.1%] vs. 71,525 of 80,547 [88.8%] malignant cases), race/ethnicity (781 of 1,324 [59.0%] vs. 64,518 of 80,547 [80.1%] white cases), or diagnostic confirmation (1,233 of 1,324 [93.2%] vs. 73,217 of 80,547 [90.9%] microscopically confirmed cases). When comparing initial ECC data with resubmission data, differences were seen in race/ethnicity (808 of 1,324 [61.1%] vs. 1,425 of 1,921 [74.2%] white cases), primary site (475 of 1,324 [35.9%] vs. 670 of 1,921 [34.9%] leukemia/lymphoma cases), and behavior (1,215 of 1,324 [91.8%] vs. 1,717 of 1,921 [89.4%] malignant cases). Common themes from focus group analysis included implementation challenges and facilitators, benefits of ECC, and utility of ECC data.

Conclusions. ECC provided data rapidly and reflected national data overall with differences in several data elements. ECC also expanded cancer reporting infrastructure and increased data completeness and timeliness. Although challenges related to timeliness and increased work burden remain, indications suggest that researchers may reliably use these data for pediatric cancer studies.

^aCenters for Disease Control and Prevention, Epidemic Intelligence Service, Atlanta, GA

^bCenters for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Cancer Prevention and Control, Atlanta, GA

Address correspondence to: Mary Puckett, PhD, Centers for Disease Control and Prevention, Epidemic Intelligence Service, 4770 Buford Hwy. NE, MS F-76, Atlanta, GA 30341; tel. 770-488-6451; e-mail <mpuckett1@cdc.gov>.

Approximately 13,500 children and young adults younger than 20 years of age are diagnosed with cancer each year.^{1,2} Yearly treatment costs range from \$19,000 to more than \$50,000 per patient.³ Although pediatric cancer survival rates improved by 50% from 1975 to 2006 because of advances in clinical trial-based treatment protocols,⁴ cancer is still the second-leading cause of death in children.¹ Additionally, because of high survival rates, pediatric survivors are at an increased need for follow-up, as they are at an increased risk of developing another cancer or treatment-related illness and mortality later in life.⁵⁻⁹

Cancer is a reportable disease, and the National Program of Cancer Registries (NPCR), managed by the Centers for Disease Control and Prevention (CDC), collects data on cancer cases for approximately 96% of the U.S. population.¹⁰ NPCR is an established surveillance system in which all invasive cancer cases are reported to state central cancer registries, usually managed by state health departments. However, reporting is often delayed by more than six months for reporting facilities to fully capture all diagnostic and treatment information. Data from these registries are typically not available for analysis and public health use until 24 to 36 months post-diagnosis.¹¹ This delay has been noted as a potential limitation of registry data.¹² NPCR programs have had limited success in using a rapid case ascertainment or early case capture (ECC) model, and these methods have not been used on a large scale across multiple NPCR registries.^{13,14} In response to this issue, and to The Carolyn Pryce Walker Act (2008),¹⁵ which sought to improve pediatric cancer tracking and include cases in a nationwide registry within weeks of diagnosis, NPCR piloted ECC for pediatric cases in seven existing state central cancer registries. Through the ECC pilot, facilities (e.g., hospitals, clinics, and laboratories) submit cases to states within 30 days—rather than six months—of diagnosis. States then report de-identified data to CDC biannually.

ECC has the potential to increase clinical trial enrollment and accelerate other surveillance and research activities across the cancer care continuum.¹⁴ In addition to timeliness, accuracy, representativeness, and acceptance by stakeholders are also important.¹⁶ This study assessed the quality and utility of data collected by the ECC pilot using well-established methods for surveillance system evaluation.¹⁷

METHODS

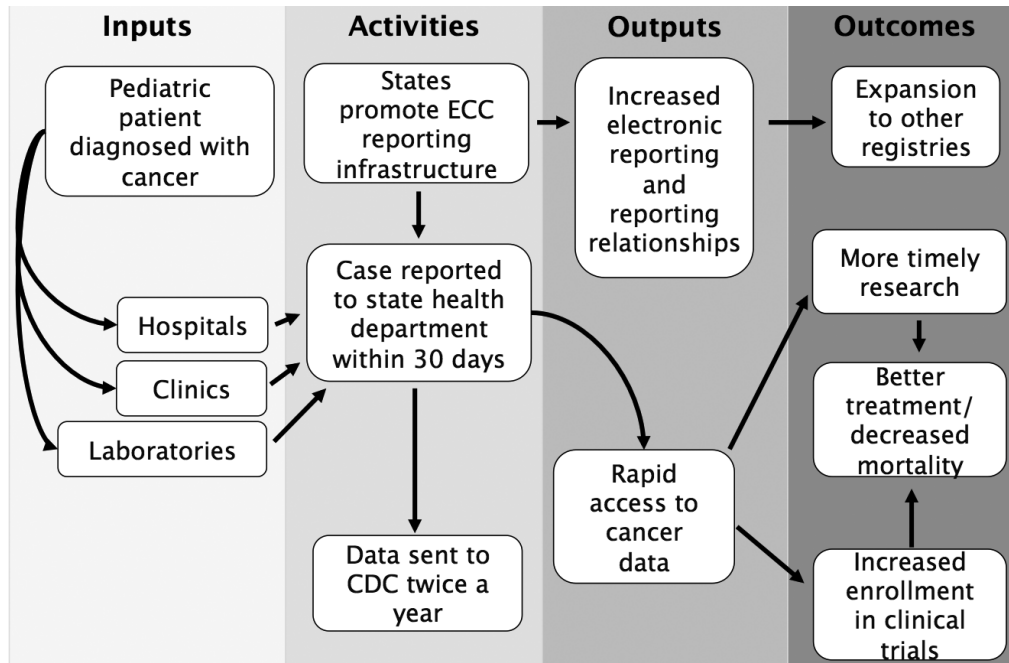
We used CDC's *Updated Guidelines for Evaluating Public Health Surveillance Systems* to evaluate ECC attributes. We assessed the systems' simplicity (structure and

ease of operation), flexibility (ability to adapt to changing information or requirements), data quality (completeness and validity of data obtained), acceptability (willingness of stakeholders to participate), representativeness (accuracy of disease distribution in the population), timeliness (duration between system steps), and stability (ability to continue operating in the future).¹⁷

The ECC pediatric cancer pilot included data from seven states that have central cancer registries funded by CDC's NPCR program: California, Kentucky, Louisiana, Minnesota, Nebraska, New York, and Oklahoma. The total award for ECC funding was approximately \$2 million per year for three years, with individual state registries receiving approximately \$240,000 to \$374,000 per year. Within each state, data were collected from all facilities that diagnosed or treated pediatric cancer patients. Because patients may have traveled to pediatric cancer facilities outside of their state of residence, each registry also pursued data reporting from out-of-state facilities that may have seen a large number of their resident patients. De-identified case-level data were then reported to CDC four months and 10 months following the data collection period. States were de-identified in all analyses to further maintain confidentiality in reporting (Figure).

We assessed data quality and representativeness by quantitative comparisons. Datasets included an initial ECC submission (submitted in October 2012) of January–June 2012 cases, which was compared with all other datasets, including a resubmission of ECC data (submitted in April 2013) diagnosed within the same time period and a U.S. Cancer Statistics (USCS) submission of cases diagnosed from January 2006 to December 2010, in all states. The USCS submission used in the comparison included stable, existing cancer data on the types of cases present in ECC. We analyzed it in aggregate (all states) and also restricted it to data from the participating ECC states for a more direct comparison.¹⁸

We compared the percentage of missing values and the distribution of selected variables for cases diagnosed from January to June 2012 in the initial ECC submission with those in the resubmission and USCS datasets. We measured the completeness of data by assessing the percentage of missing data in each dataset. The program standards for ECC data submission were that $\leq 1\%$ of records had missing data for age, sex, and county and $\leq 2\%$ of records had missing data for race/ethnicity. These standards, set for the ECC pilot project, were stricter than traditional NPCR reporting standards.¹¹ Because the central cancer registries continued to receive new ECC cases and additional data

Figure. Logic model for an early case capture pediatric cancer project, seven U.S. states, 2011–2013^a

^aCases reported on a more rapid time scale than traditional reporting methods were considered ECC cases. In this ECC pediatric cancer project, hospitals, clinics, and laboratories diagnose pediatric patients with cancer. State registries are responsible for promoting the infrastructure necessary for ECC reporting. States included California, Kentucky, Louisiana, Minnesota, Nebraska, New York, and Oklahoma.

ECC = early case capture

CDC = Centers for Disease Control and Prevention

for existing cases between the biannual submissions to CDC, it was expected that the resubmission dataset would have additional cancer cases reported and/or more complete information within a given record. Therefore, we assessed the number of new cases from January to June 2012 that were not included in the initial ECC submission and also measured the data item concordance between cases that were present in both the initial and subsequent ECC submission. Concordance was determined by assessing whether values present in the ECC submission were the same or different in the resubmission. We reported the percentage of cases with equal values for the variables assessed. We excluded data from one state from concordance measurements because of data linkage errors between initial ECC and resubmission datasets.

We measured the representativeness of ECC submission data to final registry data by comparing it with resubmission data and routinely reported NPCR data by demographic characteristics (sex, age at diagnosis, and race/ethnicity) and tumor characteristic distributions (primary site, behavior, and diagnostic confirmation) between the datasets using

bivariable analyses that controlled for state. We set statistical significance at $\alpha=0.05$. We imported data using SEER*stat 8.1.2,¹⁹ and we conducted all analyses using SAS[®] version 9.3.²⁰

We assessed the flexibility, simplicity, acceptability, and stability of the ECC system through qualitative methods to better understand participants' understanding of these complex situations and processes.²¹ Specifically, we conducted a qualitative assessment of the system through three telephone focus groups with staff members from each ECC site. From August to September 2013, we asked participants about several topics, including current practices, implementation difficulties, data use, system benefits, and major lessons learned. All interviews were transcribed verbatim. We used pseudonyms to de-identify transcripts. We used the constant comparative method with grounded theory techniques throughout the analysis.^{22,23} The three coders jointly developed and revised the codebook and independently coded the transcripts to identify common themes. Coders resolved conflicts by consensus; calculating interrater reliability was, therefore, unnecessary. We managed data using NVivo[®] version

10.²⁴ We also maintained an audit trail to record key analytic decisions.

RESULTS

The initial ECC submission reported data on 1,324 cases of pediatric cancer, and the resubmission dataset reported data on 1,921 cases of pediatric cancer, indicating that 597 cases had not been reported to the state central registries at initial submission, with 17.0% of the initial submission cases and 5.9% of the resubmission cases reported by pathology reports only. A total of 80,547 cases were among children aged 0–19 years in the 2006–2010 USCS dataset, of which 19,411 came from six of the seven ECC states. (One state had not submitted data to USCS and was excluded).

ECC cases were a median of 9.7 years of age. Most cases were male and white (average 52.3% and 61.1%, respectively). A median of 35.9% (range: 29.0%–40.6%) of all cases in each state were leukemia/lymphoma. Malignant tumors comprised 91.8% of reported cases. Benign (3.6%), uncertain/borderline (3.8%), and in situ (0.8%) tumors comprised the remaining 8.2% of the sample (Table 1).

Data completeness and concordance

Fewer than 1% of ECC records were missing information from the initial submission for several variables (e.g., age, sex, diagnostic confirmation, and histologic type). However, 18.3% of ECC cases were missing data on race, 21.2% were missing data on Hispanic ethnicity, and 3.1% were missing data on county of diagnosis. These values were outside the accepted standards for unknown values. All variables in the resubmission dataset met completeness standards, with the exception of race and Hispanic ethnicity (missing in 6.6% and 9.3% of cases, respectively) (Table 2).

Concordance between initial and resubmission data items for individual ECC cases in each dataset for all variables collected ranged from 95.6% to 100.0% (Table 2). Significant differences were seen for race/ethnicity ($p<0.001$), primary site ($p=0.002$), and behavior ($p=0.02$) (Table 3). When analyses were stratified by state, these results appeared to be driven by one state, which was not representative for race/ethnicity ($p<0.001$).

National data representativeness

We found significant differences between ECC data and USCS data for participating ECC states for sex ($p<0.001$), race/ethnicity ($p<0.001$), primary site ($p<0.001$), diagnostic confirmation ($p=0.04$), and behavior ($p<0.001$). Similarly, significant differences

were seen for race/ethnicity ($p<0.001$), primary site ($p<0.001$), behavior ($p<0.001$), and diagnostic confirmation ($p=0.01$) when comparing data for all states in USCS (Table 4).

Implementation challenges

One theme that emerged from registry staff interviews involved implementation challenges, particularly staffing issues such as a lack of qualified staff members for available positions and staff turnover at the reporting facilities and registries. For example, one staff member said:

Recruitment is always a challenge, recruiting the right staff with the right qualifications . . . is always tough, so we were just fortunate. . . . It took over a year before we found a technical person who we felt was competent and then could contribute sufficiently to the project. . . . You've got to have people in place and trained, and that's always a challenge.

Another interviewee described the challenges with turnover: "It's really difficult to continuously try and train somebody . . . and while the positions are vacant, have the staff of the registry fill in [during the interim]."

Another major challenge participants described was the increased work commitment required. Factors adding to the increased work commitment included the time and effort required to manually abstract cases, obtain race information, and remind reporting facilities of ECC requirements. Participants considered this work commitment to exceed the demands of normal registry operations.

Implementation facilitators

Staff members also reported circumstances that facilitated ECC activities, including leveraging existing resources within the registries and relationships with reporting facilities. One interviewee said, "I think one of the things that's been easy for us is we've had [other] studies in place, so I was able to [adapt] some of the communications based upon our other studies. I didn't have to come up with everything from scratch." Another interviewee noted:

I think it's really important to have good relationships with the hospitals . . . maybe I have a hospital . . . that sees a patient that . . . they're a little suspicious of the white blood count, and maybe they send them [out of state] . . . [but if we have a relationship, they] will give me a heads up and then I can contact [that] hospital.

Having these resources and relationships in place before beginning ECC allowed registries to complete

Table 1. Demographic characteristics of patients and characteristics of tumors in pediatric cancer cases (n=1,324) in seven early case capture U.S. states,^a January–June 2012^b

Characteristics	State							Overall
	A	B	C	D	E	F	G	
Demographic characteristics								
Mean age (in years)	10.2	9.1	9.0	9.3	12.0	9.4	10.9	9.7
Sex: male (percent)	52.6	60.0	47.8	64.1	70.2	48.2	48.4	52.3
Race (percent)								
White	53.5	82.5	69.6	78.9	86.0	52.7	81.3	61.1
Black	7.3	7.5	27.5	6.3	10.5	11.3	6.3	10.4
Asian/Pacific Islander	4.3	2.5	2.9	8.5	1.8	4.4	0.0	4.3
Other ^c	15.1	1.3	0.0	2.8	1.8	4.3	12.5	5.9
Unknown	19.8	6.3	0.0	3.5	0.0	27.4	0.0	18.3
Ethnicity (percent)								
Hispanic	52.2	3.8	4.4	4.9	5.3	12.1	14.1	17.2
Tumor characteristics								
Primary site (percent)								
Leukemia/lymphoma	37.1	32.5	29.0	38.2	38.6	35.5	40.6	35.9
Brain and other nervous system	14.2	23.8	30.4	25.2	21.1	16.5	18.8	18.5
Endocrine system	7.3	11.3	8.7	8.9	10.5	11.7	9.4	10.2
Soft tissue including heart	7.3	8.8	4.4	4.9	1.8	6.5	7.8	6.3
Bones and joints	5.6	2.5	4.4	3.3	12.3	5.5	6.3	5.4
Skin, excluding basal and squamous	5.6	6.3	7.3	4.9	3.5	4.5	0.0	4.7
Urinary system	5.2	1.3	7.3	6.5	5.3	4.1	4.7	4.6
Digestive system	2.2	8.8	4.4	2.4	5.3	1.5	3.1	2.6
Male genital system	6.5	2.5	0.0	1.6	0.0	3.6	3.1	3.5
Female genital system	3.5	2.5	1.5	3.3	0.0	1.8	1.6	2.2
Eye and orbit	1.7	0.0	1.5	0.0	0.0	3.3	3.1	2.2
Oral cavity and pharynx	2.6	0.0	1.5	0.0	0.0	0.8	0.0	1.0
Respiratory system	0.4	0.0	0.0	0.0	0.0	1.7	1.6	1.0
Female breast	0.4	0.0	0.0	0.0	0.0	1.8	0.0	1.0
Missing	0.4	0.0	0.0	0.8	1.8	1.2	0.0	0.8
Diagnostic confirmation (percent)								
Microscopically confirmed	100.0	90.0	97.1	97.2	93.0	90.4	98.4	93.7
Laboratory/clinical diagnosis	0.0	10.0	1.5	2.8	7.0	9.0	1.6	6.0
Unknown	0.0	0.0	1.5	0.0	0.0	0.6	0.0	0.4
Behavior (percent)								
Benign	3.5	2.5	5.8	2.1	3.5	3.5	6.3	3.6
Uncertain/borderline	2.2	7.5	4.4	7.0	1.8	3.4	3.1	3.8
In situ	0.4	0.0	4.4	1.4	0.0	0.7	0.0	0.8
Malignant (invasive)	94.0	90.0	85.5	89.4	94.7	92.2	90.6	91.8

^aCalifornia, Kentucky, Louisiana, Minnesota, Nebraska, New York, and Oklahoma. States were de-identified in all analyses to maintain confidentiality.

^bData were submitted in October 2012.

^cIncludes American Indian/Alaska Native and all other races not listed

ECC activities with less additional effort than would have otherwise been required.

ECC benefits

Staff cited multiple benefits of ECC implementation, including improving data in the overall registry, mostly through the introduction of electronic reporting systems to increase reporting speeds, and enhancing existing and/or establishing new reporting facility relationships. One staff member described this experience:

One of our other very large hospitals here . . . has not allowed us to go in and do our own case finding like we do at every other hospital. [Because of this project, they have] agreed to switch over [to electronic reporting]. And so, not only are we going to be able to benefit for the childhood cancer cases, but I think we're going to be able to find cases that case finders historically have missed.

Regarding establishing relationships, one staff member mentioned, "This ECC project has facilitated and

Table 2. Completeness and concordance of pediatric cancer patient data from January–June 2012 during initial early case capture (ECC) and resubmission datasets for six participating U.S. states, 2012–2013^{a,b}

Characteristics	Number missing in ECC initial submission (percent) ^c	Number missing in ECC resubmission (percent) ^d	Percent concordance ^e
Total	1,324 (100.0)	1,921 (100.0)	
State of diagnosis	36 (2.7)	0 (0.0)	100.0
County of diagnosis	41 (3.1)	36 (1.9)	99.9
Race	242 (18.3)	126 (6.6)	99.4
Spanish/Hispanic origin	281 (21.2)	178 (9.3)	99.4
Sex	2 (0.2)	1 (0.1)	100.0
Age at diagnosis	0 (0.0)	0 (0.0)	100.0
Date of birth	0 (0.0)	0 (0.0)	99.6
Primary site	33 (2.5)	0 (0.0)	98.6
Laterality	0 (0.0)	0 (0.0)	99.1
Diagnostic confirmation	5 (0.4)	3 (0.2)	97.9
Type of reporting source	0 (0.0)	0 (0.0)	95.6
Histologic type	0 (0.0)	0 (0.0)	98.3
Behavior code	0 (0.0)	0 (0.0)	99.6

^aCases reported on a more rapid time scale than traditional reporting methods were considered ECC cases. The percentage of missing data elements for ECC cases was calculated for the initial ECC submission and the resubmission of data for the reporting period of January–June 2012. Concordance was determined by assessing whether values present in the ECC submission were the same or different in the resubmission.

^bThe six states were California, Kentucky, Louisiana, Nebraska, New York, and Oklahoma.

^cECC initial (October 2012) and resubmission (April 2013) datasets were used.

^dPercentage of missing data in each submission for variables listed

^ePercentage of data concordance for variables listed. Missing values were excluded.

expedited our [ability to build] up a good relationship with the pediatric oncology community as well as researchers.” Overall, interviewers found that increasing and improving existing facility relationships allowed the registries to gain more complete and timely data.

Data utility

Although data had not been disseminated widely at the time of this study, registry staff members hypothesized various ways to use ECC data, such as to identify cancer clusters more rapidly, completely, and easily than without ECC reporting and to pool data from multiple registries, which could increase sample sizes for potential research uses of ECC data and allow for more robust data analysis.

DISCUSSION

When we compared initial ECC and resubmission data, data completeness and concordance were high. Furthermore, data item completeness increased between submissions. The high concordance of submission data suggests that even data reported at the earliest submission period may be reliable for variables such as age, sex, primary site, and tumor behavior. Although we saw significant differences in some variables, the magnitude of the differences was very small (<2 per-

centage points for most variables). Statistical differences are often found in large samples of cases (as in this study), and it can sometimes be more meaningful to focus on the actual differences than the statistical differences. Because the actual differences were relatively small, ECC data quality appears sufficient for clinical and research use; however, these differences should be considered when using ECC data. ECC was mostly representative of later submissions, as noted by comparisons of the ECC submissions with each other and with national data. The timeliness of data was high (i.e., cases were reported more quickly than with traditional reporting) because of rapid case reporting, and the acceptability of the ECC system by stakeholders was moderate, with benefits (e.g., increased use of electronic reporting) and challenges (e.g., finding and maintaining staff) to state registries.

Differences between ECC and USCS data can be attributed to ECC's heavy reliance on electronic reporting for case ascertainment. In particular, electronic pathology reports often lack race/ethnicity data. Additionally, because most electronic reports are pathology reports, they inherently miss cases that are not microscopically confirmed. However, our results indicated that ECC data are likely reliable for research into microscopically confirmed cases because the values reported for individual cases did not change between

Table 3. Comparison of pediatric cancer patient early case capture (ECC) datasets (January–June 2012) from initial submission and resubmission for six participating U.S. states, 2012–2013^{a-c}

Characteristics	Number of cases in the ECC initial submission (percent)	Number of cases in the ECC resubmission (percent)	P-value
Total	1,324 (100.0)	1,921 (100.0)	
Mean age (in years) ^{d,e}	9.7	9.6	0.97
Demographic characteristics			
Sex (male) ^d	692 (52.3)	1,021 (53.2)	0.60
Race ^{d,e}			<0.001
White	808 (61.1)	1,425 (74.2)	
Black	137 (10.4)	174 (9.1)	
Asian/Pacific Islander	56 (4.3)	128 (6.7)	
Other ^f	78 (5.9)	65 (3.4)	
Unknown	242 (18.3)	126 (6.6)	
Ethnicity ^{d,g}			<0.001
Hispanic	227 (17.2)	539 (28.1)	
Tumor characteristics			
Site ^{d,e}			0.002
Leukemia/lymphoma	475 (35.9)	670 (34.9)	
Brain and other nervous system	244 (18.5)	418 (21.8)	
Endocrine system	135 (10.2)	199 (10.4)	
Soft tissue including heart	83 (6.3)	99 (5.2)	
Bones and joints	71 (5.4)	115 (6.0)	
Skin, excluding basal and squamous	62 (4.7)	80 (4.2)	
Urinary system	60 (4.6)	78 (4.1)	
Digestive system	34 (2.6)	72 (3.8)	
Male genital system	46 (3.5)	57 (3.0)	
Female genital system	29 (2.2)	38 (2.0)	
Eye and orbit	29 (2.2)	24 (1.3)	
Oral cavity and pharynx	11 (0.9)	34 (1.8)	
Respiratory system	13 (1.0)	23 (1.2)	
Female breast	13 (1.0)	3 (0.2)	
Missing	10 (0.8)	5 (0.3)	
Diagnostic confirmation ^d			0.49
Microscopically confirmed	1,240 (93.7)	1,788 (93.1)	
Laboratory/clinical diagnosis	79 (6.0)	128 (6.7)	
Unknown	5 (0.4)	3 (0.2)	
Behavior ^d			0.02
Benign	47 (3.6)	113 (5.9)	
Uncertain/borderline	50 (3.8)	78 (4.1)	
In situ	10 (0.8)	11 (0.6)	
Malignant (invasive)	1,215 (91.8)	1,717 (89.4)	

^aCases reported on a more rapid time scale than traditional reporting were considered ECC cases. Initial and resubmission datasets covered the same time period for cases with different dates of reporting. Logistic regression was used to compare cases in initial and resubmission datasets and to determine *p*-values.

^bThe six states were California, Kentucky, Louisiana, Nebraska, New York, and Oklahoma.

^cECC initial (October 2012) and resubmission (April 2013) datasets were used.

^dBivariable comparisons of ECC submission data with full-year data controlled for state

^eOne state showed significant differences ($p < 0.001$) in state-stratified analysis.

^fIncludes American Indian/Alaska Native and all other races not listed

^gTwo states showed statistically significant differences in state-stratified analyses.

submissions, although they would likely preclude the use of race/ethnicity covariates in analyses. As we expected, ECC case completeness, as noted by the total number of cases included and the percentage of missing data elements, improved between the initial

submission and data resubmission. This improvement was likely a result of the additional time allowed for more complete case abstraction from sources other than pathology reports.

A major limitation of routinely collected cancer

Table 4. Comparison of demographic and tumor characteristics between the pediatric cancer patients in the initial early case capture (ECC) submission dataset (October 2012) and U.S. Cancer Statistics data (2006–2010) for six participating U.S. states and all U.S. states, November 2013^{a-c}

Characteristics	Number of ECC initial submissions (percent)	Number of USCS submissions in ECC states (percent)	P-value ^d	Number of USCS submissions in all states (percent)	P-value ^e
Total	1,324 (100.0)	19,411 (100.0)		80,547 (100.0)	
Mean age (in years)	9.7	9.9	0.09	9.9	0.33
Demographic characteristics					
Sex (male)	673 (50.9)	10,384 (53.5)	<0.001	42,609 (52.9)	0.17
Race			<0.001		<0.001
White	781 (59.0)	15,470 (79.7)		64,518 (80.1)	
Black	14 (10.9)	1,960 (10.1)		9,987 (12.4)	
Asian/Pacific Islander	50 (3.8)	1,378 (7.1)		3,221 (4.0)	
Other ^f	83 (6.3)	368 (1.9)		1,610 (2.0)	
Unknown	266 (20.1)	252 (1.3)		1,288 (1.6)	
Ethnicity (percent)			<0.001		<0.001
Hispanic	247 (18.7)	6,133 (31.6)		15,706 (19.5)	
Tumor characteristics					
Site (percent)			<0.001		<0.001
Leukemia/lymphoma	472 (35.7)	6,929 (35.7)		27,547 (34.2)	
Brain and other nervous system	235 (17.8)	4,095 (21.1)		18,284 (22.7)	
Endocrine system	137 (10.4)	2,154 (11.1)		837 (10.6)	
Soft tissue including heart	86 (6.5)	1,048 (5.4)		4,430 (5.5)	
Bones and joints	74 (5.6)	989 (5.1)		4,027 (5.0)	
Skin, excluding basal and squamous	62 (4.7)	698 (3.6)		3,302 (4.1)	
Urinary system	58 (4.4)	718 (3.7)		3,141 (3.9)	
Digestive system	34 (2.6)	659 (3.4)		2,658 (3.3)	
Male genital system	48 (3.7)	649 (3.3)		2,496 (3.1)	
Female genital system	27 (2.1)	368 (1.9)		1,530 (1.9)	
Eye and orbit	31 (2.4)	388 (2.0)		1,530 (1.9)	
Oral cavity and pharynx	14 (1.1)	329 (1.7)		1,208 (1.5)	
Respiratory system	14 (1.1)	291 (1.5)		1,288 (1.6)	
Female breast	14 (1.1)	19 (0.1)		161 (0.2)	
Missing	10 (0.8)	77 (0.4)		402 (0.5)	
Diagnostic confirmation			0.04		0.01
Microscopically confirmed	1,233 (93.2)	17,877 (92.1)		73,217 (90.9)	
Laboratory/clinical diagnosis	84 (6.4)	1,455 (7.5)		6,604 (8.2)	
Unknown	5 (0.4)	97 (0.5)		724 (0.9)	
Behavior			<0.001		<0.001
Benign	48 (3.7)	1,164 (6.0)		4,832 (6.0)	
Uncertain/borderline	45 (3.4)	795 (4.1)		3,382 (4.2)	
In situ	10 (0.8)	135 (0.7)		724 (0.9)	
Malignant (invasive)	1,219 (92.1)	17,314 (89.2)		71,525 (88.8)	

^aCases reported on a more rapid time scale than traditional reporting were considered ECC cases. Logistic regression was used to compare initial ECC cases with USCS data from ECC states and all USCS data from 2006–2010.

^bThe six states were California, Kentucky, Louisiana, Nebraska, New York, and Oklahoma.

^cU.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2012 incidence and mortality Web-based report. Atlanta: Department of Health and Human Services (US), National Cancer Institute; 2015. Also available from: www.cdc.gov/uscs [cited 2015 Sep 16].

^dP-value derived from bivariable comparison of ECC submission data with USCS data for six ECC states, adjusted by state

^eUnadjusted chi-squared analysis of ECC initial submission data to USCS data for all states

^fIncludes American Indian/Alaska Native and all other races not listed

USCS = U.S. Cancer Statistics

registry data is the two-year delay between diagnosis and data publishing.¹² The ECC pilot project aimed to address this issue in pediatric cancer reporting by requiring that cases be reported to state registries within 30 days of diagnosis. Several studies in individual NPCR states have previously assessed the feasibility and usefulness of this model. New Hampshire employs a two-phase system, including an initial rapid report within 45 days of diagnosis followed by a full report within 180 days of diagnosis. Data accuracy and timeliness were high, suggesting that such a reporting scheme may be feasible.²⁵ Similarly, a Florida study of electronic medical records (EMRs), which could decrease reporting times through electronic transmission, suggested that EMRs could increase data completeness.²⁶ Additionally, testing a Rapid Quality Reporting System in Georgia and New Jersey allowed for expedited case reporting and was seen as a positive addition by 75% of participants.²⁷ Finally, rapid case ascertainment studies in North Carolina and Florida have found the method to be cost-effective and beneficial overall.^{13,14} These results from individual NPCR states are generally consistent with the results of the pilot study, and, taken together, are encouraging.

Additionally, another U.S. surveillance system, the Surveillance, Epidemiology, and End Results (SEER) Program, has also successfully utilized ECC in special studies.²⁸⁻³⁰ However, several fundamental differences exist in the purpose and scope of SEER vs. NPCR. SEER is more research-oriented rather than public health-oriented, with a built-in funding structure to support ongoing, established research goals.¹⁰ The two programs follow similar standards for data collection for incident cases, but the SEER Program funds only a small number of registries that have demonstrated the ability to perform research activities and does not attempt to conduct the breadth of public health surveillance through health departments in the United States, as does NPCR.^{10,18} Two of the ECC states and portions of a third state included in this pilot project receive SEER funding in addition to NPCR funding. Existing infrastructure funding through SEER may have contributed to those registries' success in this study; however, it is not possible to measure the extent to which resources were leveraged (if at all) between these two surveillance systems at the local level.

A major facilitator of ECC was the incorporation of electronic pathology reporting and other electronic health records, which greatly increased the timeliness of data collection. Adoption of electronic reporting by all facilities would further strengthen the system. Although electronic reporting systems have greatly

increased reporting timeliness, clinically diagnosed cases (i.e., those without a pathology report) still present a challenge. Improved methods for obtaining clinically diagnosed cases more rapidly are needed. Some states have begun utilizing electronic radiology reports, which could increase the timeliness of case ascertainment for non-pathologically diagnosed cases, and more registries may benefit from the inclusion of such systems in reporting facilities.³⁰ Each NPCR registry required an additional \$240,000 to \$374,000 per year to participate in this ECC study. Operating costs would be expected to decrease after the initial investment in electronic reporting infrastructure, because maintaining such a reporting system would not be as expensive after the initial start-up costs. Yet, NPCR cancer registries would likely continue to need additional resources outside of their normal operating resources to sustain ECC activities.

ECC has unique benefits for data collection and research on pediatric cancers. ECC could provide more timely descriptive information about pediatric cancer incidence. Pediatric cancer is rare, and any data-driven public health interventions would require case assessments from a broad set of registries covering the entire United States. The NPCR surveillance system already allows for this type of assessment, and the results of this study suggest that the increased timeliness of these data through ECC would allow for more robust analyses, such as those required in cancer cluster investigations.

Improving pediatric cancer care through increasing clinical trial enrollment is another area where ECC data could be potentially useful. Clinical trial enrollment has been shown to substantially improve survival from pediatric cancers, including leukemias, lymphomas, and medulloblastomas.³¹ Because ECC identifies cancer cases more quickly than traditional reporting methods, ECC could lead to increased and timely enrollment of children and adolescents in a range of clinical trials and/or other clinical interventions across the cancer continuum. ECC data could also increase information available for follow-up of childhood cancer survivors by providing a more accurate or complete record, as recommended by the Children's Oncology Group.³² Although the pathway from ECC data collection to clinical trial enrollment requires planning and coordination among multiple entities, ECC cases could be used more immediately to recruit patients for survivorship studies that focus on the immediate effects of cancer treatment in children, or additional psychosocial or emotional support care that may be beneficial in the short term.

Strengths and limitations

This study's main strength is that it is the largest of its kind to date and includes data from a diverse group of states. Additionally, it is the first time that ECC has been conducted in a systematic way with a group of NPCR registries. However, several limitations were associated with this analysis. The subset of states included in this pilot may not be representative of other states or the entire United States. In addition, ECC data were compared with five-year aggregate national data rather than direct national comparisons, because a lag time in data reporting meant that national data from 2012, the time period of ECC reporting, were unavailable. Also, we focused our analysis on the first submission of ECC data to CDC, with cases diagnosed from January to June 2012, which could have resulted in selection bias if data from subsequent submissions changed over time. Finally, our qualitative analysis was limited to interviews with state cancer registry staff members and may have missed challenges or benefits seen by reporting facilities or other stakeholders.

CONCLUSION

Overall, ECC implementation provides pediatric cancer data on a more rapid time scale that is generally representative of state and national data. This pilot project benefited NPCR cancer registries through infrastructure improvements, particularly in electronic reporting enhancements and improved reporting facility relationships. Although ECC data could be improved, this pilot program resulted in data that could be used to increase the timely reporting of incident cases and clinical trial enrollment. Obtaining key stakeholder involvement and promoting awareness of the usability of these data for incidence reporting, as well as other uses, may assist with improving funding availability and broader geographic implementation of ECC within the NPCR. Future research could focus on methods for maximizing the use of ECC data, including economic analyses of program costs. In particular, the use of ECC data in clinical trials and interventions would greatly enhance the role of public health surveillance and strengthen the connection between communities of practitioners.

The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

- Centers for Disease Control and Prevention (US). CDC Wonder: United States cancer statistics. 2010 [cited 2014 Jan 14]. Available from: <http://wonder.cdc.gov/cancer.html>
- Li J, Thompson TD, Miller JW, Pollack LA, Stewart SL. Cancer incidence among children and adolescents in the United States, 2001–2003. *Pediatrics* 2008;121:E1470-7.
- Russell HV, Panchal J, VonVille H, Franzini L, Swint JM. Economic evaluation of pediatric cancer treatment: a systematic literature review. *Pediatrics* 2013;131:E273-87.
- Smith MA, Seibel NL, Altekruse SF, Ries LA, Melbert DL, O'Leary M, et al. Outcomes for children and adolescents with cancer: challenges for the twenty-first century. *J Clin Oncol* 2010;28:2625-34.
- Geenen MM, Cardous-Ubbink MC, Kremer LC, van den Bos C, van der Pal HJ, Heinen RC, et al. Medical assessment of adverse health outcomes in long-term survivors of childhood cancer. *JAMA* 2007;297:2705-15.
- Mertens AC, Yasui Y, Neglia JP, Potter JD, Nesbit ME, Ruccione K, et al. Late mortality experience in five-year survivors of childhood and adolescent cancer: the Childhood Cancer Survivor Study. *J Clin Oncol* 2001;19:3163-72.
- Neglia JP, Friedman DL, Yasui Y, Mertens AC, Hammond S, Stovall M, et al. Second malignant neoplasms in five-year survivors of childhood cancer: Childhood Cancer Survivor Study. *J Natl Cancer Inst* 2001;93:618-29.
- Hudson MM, Jones D, Boyett J, Sharp GB, Pui CH. Late mortality of long-term survivors of childhood cancer. *J Clin Oncol* 1997;15:2205-13.
- Henderson TO, Friedman DL, Meadows AT. Childhood cancer survivors: transition to adult-focused risk-based care. *Pediatrics* 2010;126:129-36.
- Wingo PA, Jamison PM, Hiatt RA, Weir HK, Gargiullo PM, Hutton M, et al. Building the infrastructure for nationwide cancer surveillance and control—a comparison between the National Program of Cancer Registries (NPCR) and the Surveillance, Epidemiology, and End Results (SEER) Program (United States). *Cancer Causes Control* 2003;14:175-93.
- Centers for Disease Control and Prevention (US). National Program of Cancer Registries program standards, 2012–2017 (updated January 2013) [cited 2015 Jan 14]. Available from: http://www.cdc.gov/cancer/npcr/pdf/npcr_standards.pdf
- Izquierdo JN, Schoenbach VJ. The potential and limitations of data from population-based state cancer registries. *Am J Public Health* 2000;90:695-8.
- Pal T, Permuth-Wey J, Betts JA, Krischer JP, Fiorica J, Arango H, et al. BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. *Cancer* 2005;104:2807-16.
- Aldrich TE, Vann D, Moorman PG, Newman B. Rapid reporting of cancer incidence in a population-based study of breast-cancer: one constructive use of a central cancer registry. *Breast Cancer Res Treat* 1995;35:61-4.
- H.R. 1553, 110th Congress, 2nd Session (2008).
- Thoburn KK, German RR, Lewis M, Nichols PJ, Ahmed F, Jackson-Thompson J. Case completeness and data accuracy in the Centers for Disease Control and Prevention's National Program of Cancer Registries. *Cancer* 2007;109:1607-16.
- German RR, Lee LM, Horan JM, Milstein RL, Pertowski C, Waller MN; Guidelines Working Group Centers for Disease Control and Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR Recomm Rep* 2001;50(RR-13):1-35.
- Centers for Disease Control and Prevention (US). United States cancer statistics. 2014 [cited 2014 Jan 14]. Available from: <http://apps.nccd.cdc.gov/uscs>
- National Cancer Institute (US), Surveillance Research Program. SEER*Stat: Version 8.1.2. Washington: National Cancer Institute; 2013.
- SAS Institute, Inc. SAS®: Version 9.3. Cary (NC): SAS Institute, Inc.; 2011.
- Riessman CK, editor. *Qualitative studies in social work research*. Thousand Oaks (CA): Sage Publications; 1994.
- Charmaz K. *Premises, principles, and practices in qualitative research: revisiting the foundations*. *Qual Health Res* 2004;14:976-93.

23. Strauss A, Corbin J. Basics of qualitative research: grounded theory procedures and techniques. 2nd ed. Thousand Oaks (CA): Sage Publications, Inc.; 1990.
24. QSR International Pty Ltd. NVivo: Version 10. Melbourne: QSR International Pty Ltd.; 2012.
25. Celaya MO, Riddle BL, Cherala SS, Armenti KR, Rees JR. Reliability of rapid reporting of cancers in New Hampshire. *J Registry Manag* 2010;37:107-11.
26. Hernandez MN, Voti L, Feldman JD, Tannenbaum SL, Scharber W, Mackinnon JA, et al. Cancer registry enrichment via linkage with hospital-based electronic medical records: a pilot investigation. *J Registry Manag* 2013;40:40-7.
27. Stewart AK, McNamara E, Gay EG, Banasiak J, Winchester DP. The Rapid Quality Reporting System—a new quality of care tool for CoC-accredited cancer programs. *J Registry Manag* 2011;38:61-3.
28. Hamilton AS, Hofer TP, Hawley ST, Morrell D, Leventhal M, Deapen D, et al. Latinas and breast cancer outcomes: population-based sampling, ethnic identity, and acculturation assessment. *Cancer Epidemiol Biomarkers Prev* 2009;18:2022-9.
29. Holly EA, Gautam M, Bracci PM. Comparison of interviewed and non-interviewed non-Hodgkin's lymphoma (NHL) patients in the San Francisco Bay Area. *Ann Epidemiol* 2002;12:419-25.
30. Yoon J, Malin JL, Tao ML, Tisnado DM, Adams JL, Timmer MJ, et al. Symptoms after breast cancer treatment: are they influenced by patient characteristics? *Breast Cancer Res Treat* 2008;108:153-65.
31. Bleyer WA, Tejeda H, Murphy SB, Robison LL, Ross JA, Pollock BH, et al. National cancer clinical trials: children have equal access; adolescents do not. *J Adolesc Health* 1997;21:366-73.
32. American Academy of Pediatrics Section on Hematology/Oncology Children's Oncology Group. Long-term follow-up care for pediatric cancer survivors. *Pediatrics* 2009;123:906-15.