



Published in final edited form as:

J Registry Manag. 2022 ; 49(1): 10–16.

Operational Characteristics of Central Cancer Registries that Support the Generation of High-Quality Surveillance Data

Patrick Edwards, MSc^a, Amarilys Bernacot, MPH^a, Florence K. L. Tangka, PhD^b, Paran Pordell, MPH^b, Jenny Beizer, BA^a, Reda Wilson, MPH^b, Wendy Blumenthal, MPH^b, Sandra F. Jones^b, Maggie Cole-Beebe, PhD^a, Sujha Subramanian, PhD^a

^aRTI International, Research Triangle Park, North Carolina.

^bCenters for Disease Control and Prevention, Atlanta, Georgia.

Abstract

Objectives: We aim to assess external and internal attributes and operations of the Centers for Disease Control and Prevention (CDC)'s National Program of Cancer Registries (NPCR) central cancer registries by their consistency in meeting national data quality standards.

Methods: The NPCR 2017 Program Evaluation Instrument (PEI) data were used to assess registry operational attributes, including adoption of electronic reporting, compliance with reporting, staffing, and software used among 46 NPCR registries. These factors were stratified by (1) registries that met the NPCR 12-month standards for all years 2014–2017; (2) registries that met the NPCR 12-month standards at least once in 2014–2017 and met the NPCR 24-month standards for all years 2014–2017; and (3) registries that did not meet the NPCR 24-month standards for all years 2014–2017. Statistical tests helped identify significant differences among registries that consistently, sometimes, or seldom/never achieved data standards.

Results: Registries that always met the standards had a higher level of electronic reporting and a higher compliance with reporting among hospitals than registries that sometimes or seldom/never met the standards. Although not a statistically significant finding, the same registries also had a higher proportion of staffing positions filled, a higher proportion of certified tumor registrars, and more quality assurance and information technology staff.

Conclusions: This information may be used to understand the importance of various factors and characteristics, including the adoption of electronic reporting, that may be associated with a registry's ability to consistently meet NPCR standards. The findings may be helpful in identifying best practices for processing high-quality cancer data.

Keywords

cancer data; cancer registries; data modernization

Introduction

Cancer was the second-leading cause of death in the United States in 2018, with more than 1.7 million new cancer diagnoses and about 600,000 cancer deaths.¹ Central cancer registries (referred to henceforth as *registries*) provide information on cancer incidence and trends that can be used to monitor the burden of disease and develop and evaluate local and national cancer prevention and control interventions.^{2–4} Registries provide data that include information on new cancer cases, such as the type, stage, location of the cancer, treatment, and outcomes.²

The Centers for Disease Control and Prevention (CDC) provides technical and financial support to central population-based cancer registries in 46 states, the District of Columbia, Puerto Rico, the US Virgin Islands, and the US-Affiliated Pacific Islands (USAPI), and has been actively engaged in providing training and software support and in evaluating registries' progress and data quality.^{2, 5} This data quality is monitored by CDC through registries' ability to meet established National Program of Cancer Registries (NPCR) standards. These standards are based on various metrics that monitor data quality, completeness, and timeliness at 12 and 24 months following the date of diagnosis.⁶

To date, the operational characteristics of registries by their ability to meet the NPCR quality standards have not been systematically evaluated. The aim of this study is to assess external and internal attributes and operations of registries by their consistency in meeting NPCR data quality standards. By identifying characteristics and operational processes that result in high-quality data, the findings from this study may help inform efforts to support registries.

Methods

Overview of Approach

We conducted a program evaluation using secondary quantitative data from a prior program evaluation survey to assess registry attributes by their consistency in meeting NPCR data quality standards. NPCR data quality are assessed based on whether the registry meets the National Data Quality Standards (formerly known as the *24-month standards*) and the Advanced National Data Quality Standards (formerly known as the *12-month standards*). Registries were grouped into 1 of 3 levels: registries that met the NPCR 12-month standards for all years 2014–2017 (labeled in figures as “Always”); registries that met the NPCR 12-month standards at least once in 2014–2017 and met the NPCR 24-month standards for all years 2014–2017 (labeled in figures as “Sometimes”), and registries that did not meet the NPCR 24-month standards for all years 2014–2017 (labeled in figures as “Seldom/Never”). Additionally, we used data collected by the CDC and publicly available data to establish a general overview of registry characteristics. Registries were split into thirds for volume of cases (number of cancers reported by the registry), size of area covered by registry (square miles covered by the state), and presence of rural areas (population density) to indicate registries that represented the top third (high), middle third (medium), and bottom third (low) of each attribute. We also include geographic location (Midwest, Northeast, South, or West) and funding source (based on federal, state and other funding received).

Quantitative Data Compilation and Analysis

Data from the NPCR 2017 Program Evaluation Instrument (PEI) were extracted to support contextual information about the participating registries. The PEI, which has Office of Management and Budget (OMB) approval (#0920–0706) and is generally conducted every 2 years, collects information via a web-based survey instrument on registries' operational attributes, progress towards meeting program standards, and information on advanced activities performed.⁷ The 2017 PEI data represented the latest completed year of the PEI survey during the time of the study and provided baseline information about participating registries for the evaluation. We used Stata software to analyze data on 46 registries compiled from the PEI. Registries' general characteristics, along with data from the PEI, were stratified by their consistency in meeting NPCR's quality standards.⁶

In the PEI, registries reported the number of sources required to report, the number of sources compliant with reporting, and the number of sources reporting electronically for various hospital and laboratory sources. The PEI defines electronic reporting for registries as "the collection and transfer of data from source documents by hospitals, physician offices, clinics or laboratories in a standardized, coded format that does not require manual data entry at the central cancer registry (CCR) level to create an abstracted record." *Compliance* with reporting was defined as the total number of facilities required to report that actually reported cancer data at the end of 2017. For each registry, we calculated the percentage of electronic reporting and the percentage of reporting compliance separately for each source type (pathology laboratories and hospitals). Using our defined stratifications, we then calculated the overall average percentage of electronic reporting and compliance among registries within each designation of consistency in meeting NPCR quality standards. Analysis of variance tests were conducted to assess statistical significance among registries of different consistencies in meeting the NPCR data quality standards.

The PEI also collected data on registries' staffing characteristics, including their number of contracted and noncontracted staff by various positions, along with their number of filled and vacant positions. Registries reported the number of staff who were a certified tumor registrar (CTR) and number of staff in key registry positions, including quality control staff and computer or information technology (IT)-related staff. The number of full-time equivalents (FTEs) that registries reported across various staffing positions was used to generate average percentage of positions filled, positions that were noncontracted staff, and positions that had CTR credentials. We also generated the median number of FTEs of various specified position types stratified by the registries' consistency in meeting data quality standards. Kruskal-Wallis equality-of-populations rank tests were performed to assess statistical significance of the staffing results.

Registries also reported information about data content and format, including the primary software systems used to process and manage cancer data and use of select Registry Plus suite software components used to acquire and link data. This study reports on registries' use of Electronic Mapping, Reporting, and Coding (eMaRC) Plus, Web Plus, and Link Plus,⁸ as these were the most-reported Registry Plus software applications. The percentage of registries using various software among each registry stratification level was generated. Lastly, registries reported on any systems in place for early case capture (rapid case

ascertainment), along with whether early case capture was being performed for all cases or a subset of cases. Early case capture is when a registry reports cases to the central cancer registry much earlier than required by law (ie, 180 days), usually within 30 days of diagnosis date or date first seen at an institution. The percentages of registries with systems in place to perform early case capture and those that performed early case capture on all cases were also calculated. Fisher's exact tests were performed to identify statistical significance of the software usage as well as that of early case capture numbers by registry consistency in meeting NPCR data quality standards.

Results

General characteristics of the 46 registries are presented in Table 1. Registry characteristics varied across their consistency in meeting NPCR 12- and 24-month reporting standards. Out of the 12 registries that always met NPCR 12- and 24-month reporting standards, half were considered lower case volume registries, with the other half split between medium or higher volumes of cases. Compared to the 2 groups of registries that never or sometimes met the NPCR standards, the group of registries that always met the NPCR standards had the highest number of registries located in the Northeast, the largest proportion of registries with a smaller geographic area covered, and the largest proportion of registries with a low presence of rural areas. Overall, 9 total registries received sources of additional federal and state funding beyond the standard NPCR or state funding.

The average level of electronic data adoption and the compliance by registries of various consistencies in meeting NPCR data reporting standards are presented in Figure 1. On average, almost half of laboratory facilities among registries that always met standards reported data electronically, compared to 40% and 43% among registries that sometimes or seldom/never met the standards. On average, about 89% of hospital sources reported electronically among registries that always met the standards, compared to 83% and 63% hospital electronic reporting among registries that sometimes or seldom/never met the standards.

Registries that always met data reporting standards had an average reporting compliance of 91% among laboratories, compared to 86% and 79% among registries that sometimes or seldom/never met the standards. In hospital sources, average compliance was 98% among registries that always met the standards, compared to 92% and 76% among registries that sometimes or seldom/never met the standards.

The average percentage of staffing along with the median staffing for selected positions is presented in Figure 2 by registries' consistency in meeting the standards. Among registries that always met the standards, 93% of registry positions were filled, 71% of total positions were noncontracted, and 53% of total staff had CTR credentials. Among registries that sometimes met standards, 92% of positions were filled, 75% of total staff were noncontracted, and 46% of staff had CTR credentials. Among registries that seldom/ never met standards, 85% of positions were filled, 97% of total staff were noncontracted, and 40% of staff had their CTR credentials. Registries that always met the standards had a median of 4 CTR quality-control staff, compared to 3 CTR quality-control staff among registries that

sometimes met the standards, and 2 among registries that seldom/ never met the standards. However, registries that seldom/ never met the standards had a median of 2 quality-control staff who did not have CTR credentials, which was higher than the other registry groups. Registries that always or sometimes met the standards also had a higher median number of computer/information technology-related staff than registries that seldom/never met the standards. None of the Kruskal-Wallis equality-of-populations rank test results for staffing were statistically significant.

The primary software used by registries for data processing, along with other Registry Plus software used, are presented in Figure 3. These data are stratified by registries' consistency in meeting data reporting standards. Among registries that always met data reporting standards, about a third primarily used CDC software (CRS Plus) to process and manage data, while 58% used a commercial vendor, and 8% used software developed by the registry (in-house software). Among registries that sometimes met the standards, about 52% primarily used CRS Plus, while 30% used a commercial vendor, and 17% used in-house software. Among registries that seldom/never met standards, about 45% primarily used CRS Plus, 45% used a commercial vendor, and 9% used in-house software.

The majority of registries, including all registries that always met the standards, used eMaRC Plus software (part of Registry Plus). This electronic pathology reporting program allows registries to receive and process Health Level Seven (HL7)⁸ data from laboratories and HL7 Clinical Data Architecture (CDA)-formatted data from physicians' offices. About 75% of all registries used Web Plus, allowing them to receive electronic data securely over the internet. This software is frequently used to acquire hospital data, contributing the bulk of registries' case volume. Most registries also reported using Link Plus, a linkage tool to help registries detect duplicates in their databases and to link cancer data with external data sources. Registries that always met the data standards had a slightly higher than average level of using Web Plus and Link Plus than registries that sometimes or seldom/never met the standards. Fisher's exact test results for the software usage by registry consistency in meeting NPCR data quality standards were not statistically significant.

Figure 4 presents the percentage of registries with a system in place for early case capture along with the percentage of registries that perform early case capture on all cases. Among registries that always met the standards, half had a system in place for early case capture and about 25% performed early case capture on all cases. For registries that sometimes met the standards, 30% had a system in place for early case capture and 4% performed early case capture on all cases. Of registries that seldom/never met the standards, 18% had a system in place for early case capture and none performed early case capture on all cases. Fisher's exact test results were not statistically significant.

Discussion

This study examined key attributes and staffing among registries of differing consistencies in meeting NPCR data quality standards. While practices such as electronic reporting have previously been hypothesized to support registries in meeting the NPCR standards, this is the first study to report data on specific external and internal factors, including the

adoption of electronic reporting, among registries with varying consistencies in meeting the standards. The exploratory study presented in this manuscript suggests that the format of data coming into the registry plays a significant role in the registry's performance, in that high adopters of electronic reporting more consistently met the NPCR 12- and 24-month data standards. Similarly, registries with a higher compliance of reporting also more consistently met the NPCR standards. This relationship was most clearly demonstrated by hospitals that reported electronically, which reports the bulk of registries' case volumes. High-quality hospital registries are therefore an important data source within the national cancer surveillance infrastructure; they are major facilitators in central cancer registries' receipt of electronic data and their reporting of high-quality data for data users and cancer prevention and control efforts.⁴ The ability of cancer registries to meet data quality standards as a result of hospital electronic reporting and hospital compliance with reporting were both statistically significant findings. Furthermore, a recent qualitative study found that electronic reporting helped registries to improve data completeness, timeliness, and quality, which are all components in achieving the NPCR 12- and 24-month data standards.⁵

Registries' ability to hire and train qualified staff may also be an important facilitator in their consistency in meeting the data quality standards. Staffing represents the largest component of a registry's budget, with past research showing that staff compensation can range from about 65% to 92% of a registry's budget and resources used.^{9,10} Using the NPCR PEI data allowed our study to take a more detailed look at the number of staff with specific roles and qualifications across registries. Although not statistically significant, our study identified that registries that always met the NPCR standards had a higher proportion of staff with CTR credentials than registries that sometimes or seldom/never met the standards. This included a higher number of CTR quality-control staff, who are critical in ensuring the processing of high-quality data. Past research suggests that IT and technical support teams likely play a critical role in registries' ability to adopt electronic reporting and in using data processing software, which supports the importance of having a dedicated IT role and ensuring adequate technical assistance for registries to meet the data reporting standards.^{5, 11} Similar to registries with IT support, registries with a system in place for early case capture may also be more likely to meet data standards, but it is important to note that these findings were not statistically significant in our study.

The type of software used by registries did not appear to have a significant role in registries' ability to meet data reporting standards. Registries that consistently met the standards were likelier to use a commercial vendor versus the CDC software; however, this may reflect their more advanced IT resources and capabilities, to which not all registries have access. Software, hardware, networks, and other electronic architecture play a major role in registry operations and their adoption of electronic reporting, and therefore facilitate a registry's ability to meet NPCR data quality standards. While many software issues, such as the lack of automation,⁵ have been noted in past studies, CDC is actively involved in developing solutions, such as natural language processing, to better automate cancer data collection and reporting. It is also developing cloud-based computing platforms to better standardize the electronic data that registries receive from pathology laboratories and physicians' offices.^{12,13}

In this study, based on the factors considered, we find that the adoption of electronic reporting and potentially optimal staffing mix could be key areas for facilitating the achievement of NPCR data standards. By incorporating broader registry characteristics and quality of data submitted to registries, such as in information gathered in future PEIs or interviews, future studies may help reveal additional factors that can support registries to achieve data quality standards. Furthermore, as we have found in prior studies that registries vary in the cost of their operations, the roles of funding and cost of electronic data adoption needs to be further investigated.¹⁰

This study has several limitations. By using the NPCR 12- and 24-month standards only as the outcomes of interest, we may have missed factors that can impact other goals that registries often aim to achieve, such as research output and data usability. Another study limitation is that the narrow retrospective time frame used may have limited our ability to see the effect of improvements over time on registries' ability to meet the NPCR standards. Furthermore, for some categories such as type of registry staff, we had small sample sizes that may not have allowed for meaningful comparisons. While much of these data were collected during 2014–2017, certain factors, like a registry's adoption of electronic reporting, may become even more critical during periods of natural disaster, disease outbreaks similar to the COVID-19 pandemic, or in situations that require staff to work remotely for extended periods of time. Despite these limitations, this study has revealed critical factors that can support registries in meeting NPCR data standards and has identified key areas, such as the role of funding and cost of electronic data adoption, that can be explored in future studies.

Conclusion

This study explored the variations in registry attributes and processes and their roles in impacting the achievement of NPCR 12- and 24-month data standards. Though many registries differed in characteristics and in software used, the adoption of electronic reporting compliance with reporting and potentially the staffing mix emerged as the primary drivers of meeting NPCR data standards. NPCR has been at the forefront of data modernization initiatives at the CDC, and the shift among cancer registries to real-time data collection has propelled the need to develop new solutions to automate processes, develop natural language processing, and initiate the use of cloud-based platforms to modernize the collection of high-quality cancer data. This study underscores the importance of scaling up electronic reporting along with ensuring registries have the resources and staffing required to consistently meet the NPCR data standards.

Acknowledgments

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Funding for this study is provided by Centers for Disease Control and Prevention to RTI International through contract # HHSD2002013M5396. Use of trade names and commercial sources is for identification only and does not imply endorsement by the US Department of Health and Human Services.

References

1. US Cancer Statistics Working Group. US Cancer Statistics Data Visualizations Tool, based on 2020 submission data (1999–2018): US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. June 2021. www.cdc.gov/cancer/dataviz
2. About the program. Centers for Disease Control and Prevention National Program of Cancer Registries website. Accessed August 31, 2020. <https://www.cdc.gov/cancer/npcr/about.htm>
3. Weir HK, Thompson TD, Soman A, Moller B, Leadbetter S, White MC. Meeting the Healthy People 2020 objectives to reduce cancer mortality. *Prev Chronic Dis.* 2015;12:E104. [PubMed: 26133647]
4. White MC, Babcock F, Hayes NS, et al. The history and use of cancer registry data by public health cancer control programs in the United States. *Cancer.* 2017;123 (suppl 24):4969–4976. [PubMed: 29205307]
5. Tangka FK, Edwards P, Pordell P, et al. Factors affecting the adoption of electronic data reporting and outcomes among selected central cancer registries of the National Program of Cancer Registries. *JCO Clin Cancer Inform.* 2021; 5:921–932. [PubMed: 34464161]
6. Centers for Disease Control and Prevention. CDC Stacks: National Program of Cancer Registries program standards, 2012–2017 (updated January 2013). Published August 9, 2013. Accessed June 30, 2021. <https://stacks.cdc.gov/view/cdc/44020>
7. Centers for Disease Control and Prevention National Program of Cancer Registries. Appendix H: Performance Evaluation Instrument (PEI) Overview. Accessed August 3, 2021. 2021. <https://www.cdc.gov/cancer/dcpc/about/foa-dp22-2202/npcr/pdf/DP22-2202-NPCRAAppendix-H-PEI-Overview-508.pdf>
8. Registry plus software programs for cancer registries. Centers for Disease Control and Prevention National Program of Cancer Registries (NPCR) website. Accessed August 3, 2021. <https://www.cdc.gov/cancer/npcr/tools/registryplus/index.htm>
9. Subramanian S, Tangka F, Green J, Weir H, Michaud F, Ekwueme D. Economic assessment of central cancer registry operations. Part I: Methods and conceptual framework. *J Registry Manag.* 2007;34(3):75–80.
10. Tangka FK, Subramanian S, Beebe MC, et al. Cost of operating central cancer registries and factors that affect cost: findings from an economic evaluation of Centers for Disease Control and Prevention National Program of Cancer Registries. *J Public Health Manag Pract.* 2016;22(5):452–460. [PubMed: 26642226]
11. Bianconi F, Brunori V, Valigi P, La Rosa F, Stracci F. Information technology as tools for cancer registry and regional cancer network integration. *IEEE Transactions Systems Man Cybernetics.* 2012;42(6):1410–1424.
12. Blumenthal W, Alimi TO, Jones SF, et al. Using informatics to improve cancer surveillance. *J Am Med Inform Assoc.* 2020;27(9):1488–1495. [PubMed: 32941600]
13. Jones DE, Alimi TO, Pordell P, et al. Pursuing data modernization in cancer surveillance by developing a cloud-based computing platform: real-time cancer case collection. *JCO Clin Cancer Inform.* 2021;5:24–29. [PubMed: 33411623]

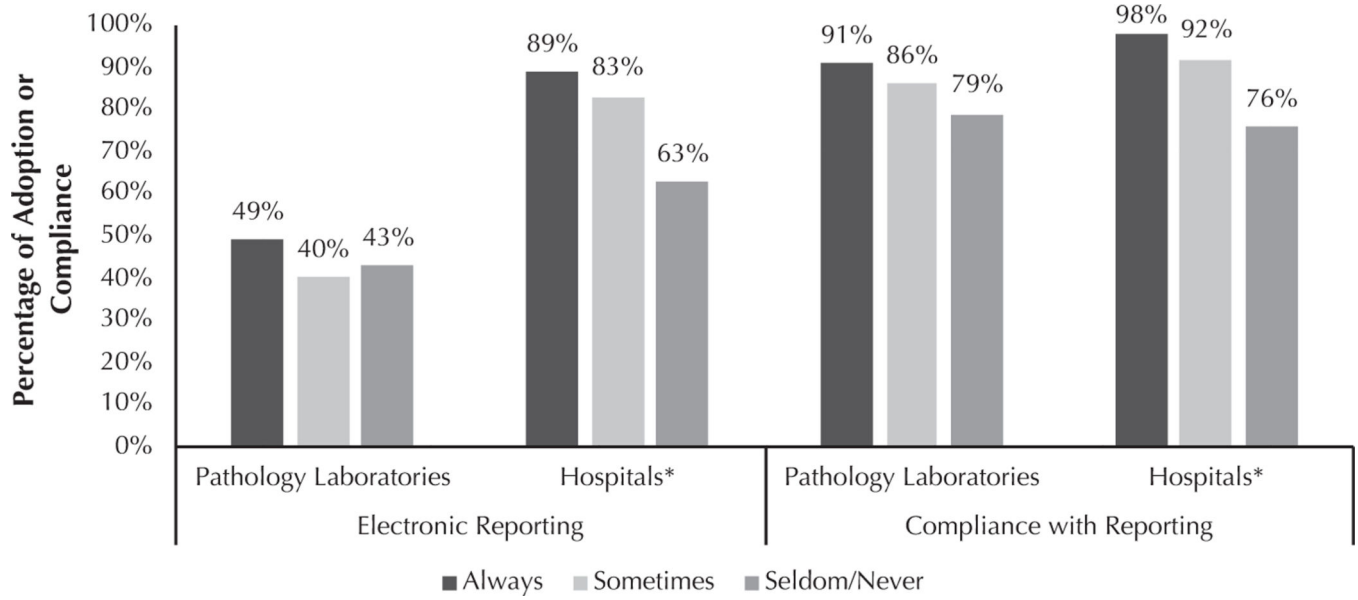


Figure 1. Average Electronic Data Adoption and Compliance with Reporting in 2017 by Achievement of National Program of Cancer Registries Data Quality and Reporting Standards. An asterisk (*) denotes statistical significance.

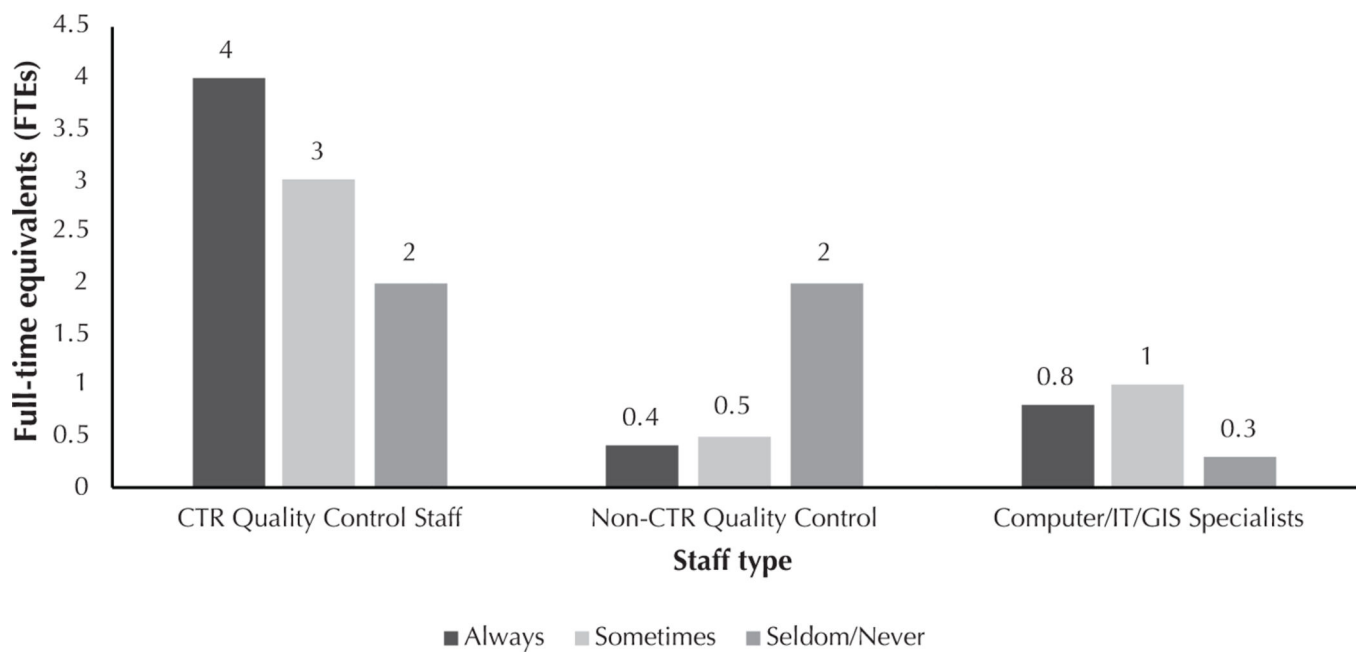
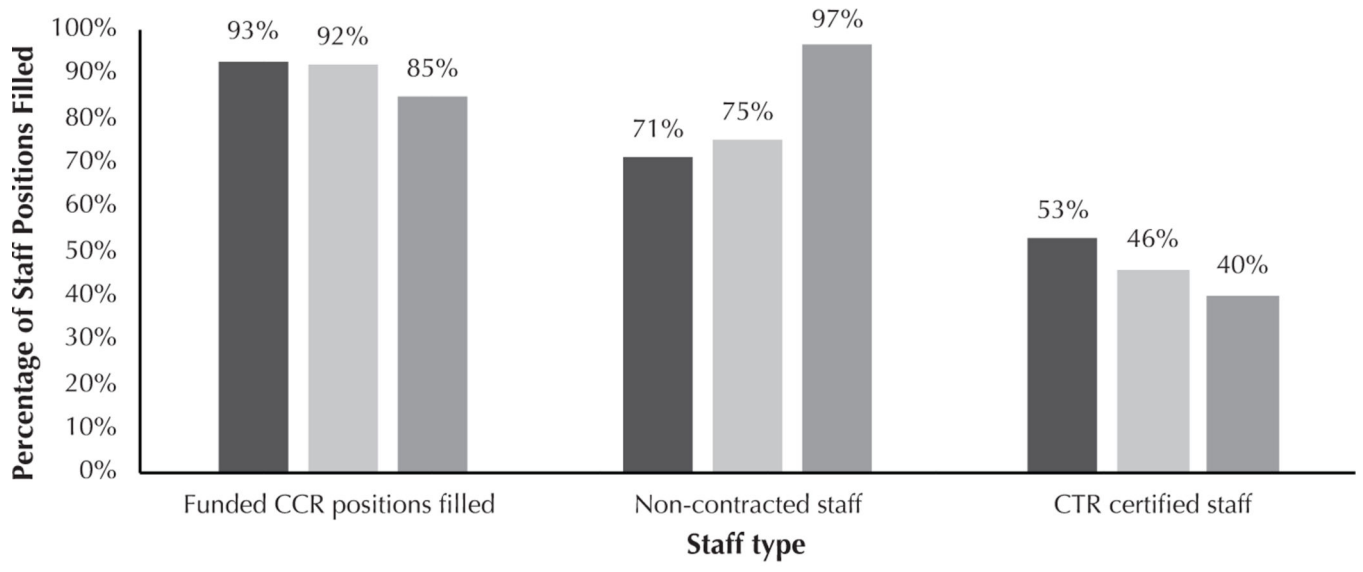


Figure 2. Average Percent Staffing and Median Staffing for Selected Positions by Achievement of National Program of Cancer Registries Data Quality and Reporting Standards, 2017
 CCR, central cancer registry; CTR, certified tumor registrar; GIS, geographic information system; IT, information technology.

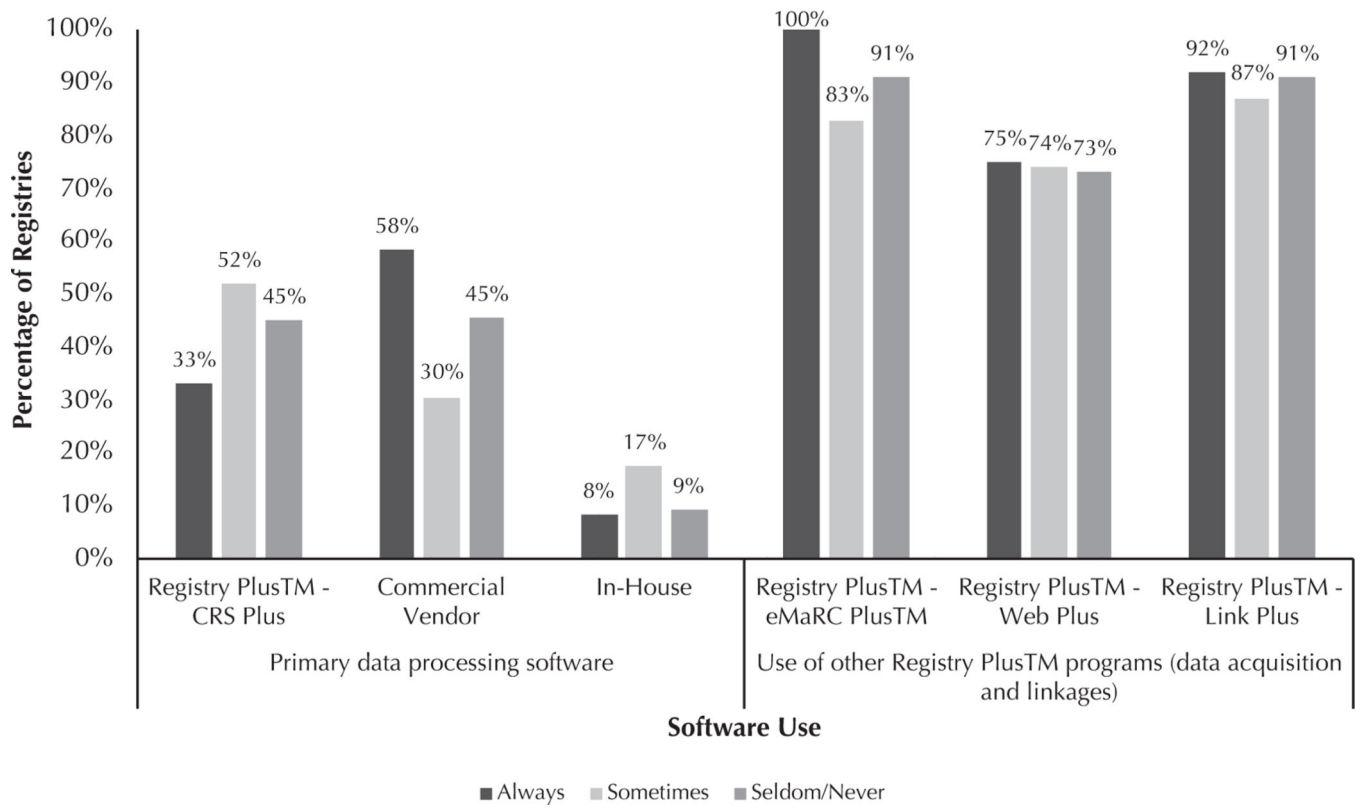


Figure 3.
 Percentage of Registry Software Use by Achievement of National Program of Cancer Registries Data Quality and Reporting Standards, 2017

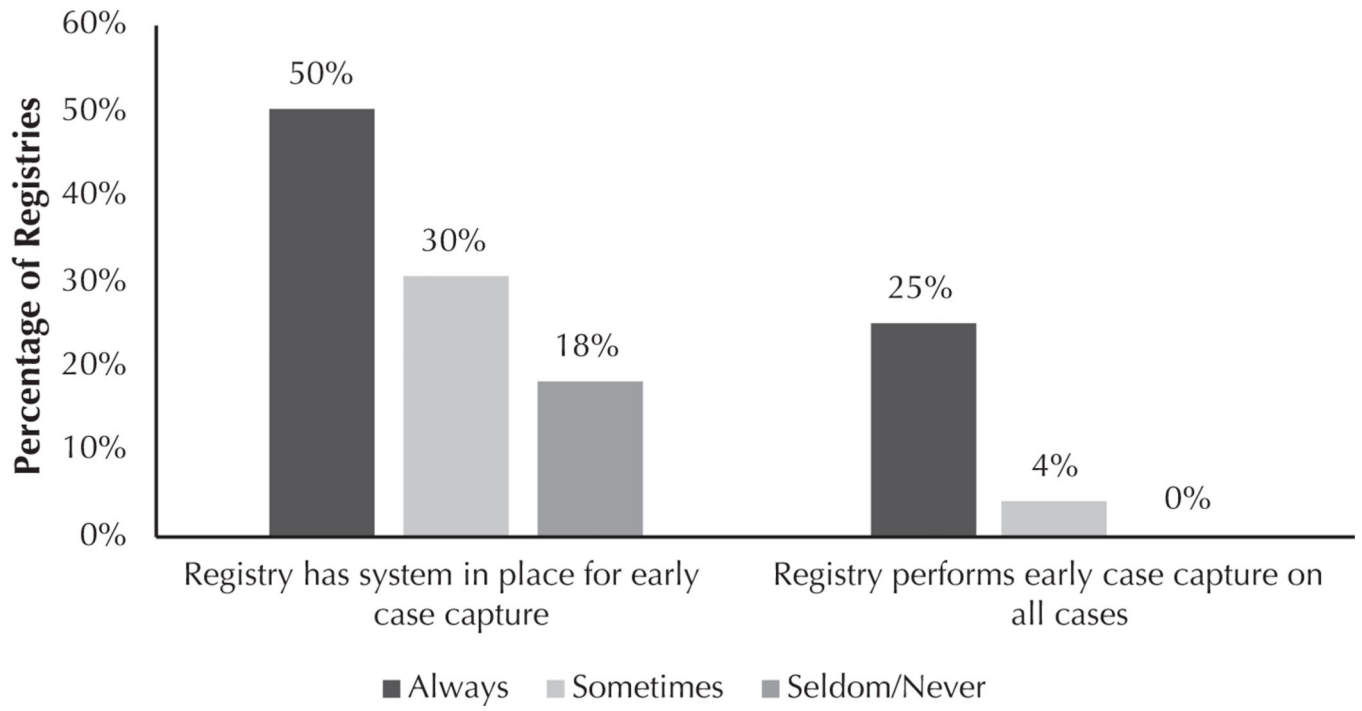


Figure 4. Percentage of Registries with Early Case Capture by Achievement of National Program of Cancer Registries Data Quality and Reporting Standards, 2017

Table 1. Overview of NPCR Registries by Achievement of NPCR Data Quality and Reporting Standards in 2014–2017

Achievement of NPCR Standards	All	Always	Sometimes	Seldom/Never
Attributes of Registries	n (%)	n (%)	n (%)	n (%)
Total	46 (100)	12 (27)	23 (51)	11 (22)
Volume of Cases				
High (> 26,558 cases)	16 (35)	4 (33)	9 (39)	3 (27)
Medium (10,455 to 26,558 cases)	15 (32)	2 (17)	8 (35)	5 (45)
Low (<10,455 cases)	15 (33)	6 (50)	6 (26)	3 (27)
Geographic Area				
Midwest	11 (23)	2 (17)	6 (26)	3 (27)
Northeast	8 (17)	5 (42)	3 (13)	0 (0)
Southern	17 (36)	3 (25)	10 (43)	4 (36)
West	10 (23)	2 (17)	4 (17)	4 (36)
Size of Area Covered by Registry				
High (> 70,684 square miles)	15 (34)	3 (25)	6 (26)	6 (55)
Medium (44,453 to 70,684 square miles)	16 (34)	4 (33)	9 (39)	3 (27)
Low (< 44,453 square miles)	15 (32)	5 (42)	8 (35)	2 (18)
Presence of Rural Areas				
High (> 147 residents per square mile)	15 (34)	4 (33)	5 (22)	6 (55)
Medium (58 to 147 residents per square mile)	15 (32)	2 (17)	11 (48)	3 (27)
Low (< 58 residents per square mile)	16 (34)	6 (50)	7 (30)	2 (18)
Sources of Additional Funding				
Other federal and state	9 (20)	3 (25)	6 (26)	0 (0)
State only	35 (76)	9 (75)	15 (65)	11 (100)
Other additional funding	2 (4)	0 (0)	2 (9)	0 (0)

NPCR, National Program of Cancer Registries.