Monitoring public health reporting: data tracking in cancer registries

Abdulrahman M. Jabour MHSM, PhD [1], Brian E. Dixon MPA, PhD, FACMI, FHIMSS [2, 3, 4]

1. Department of Health Informatics, Jazan University School of Public health and Tropical Medicine. Jazan, Saudi Arabia.

2. Department of Epidemiology, Indiana University Richard M. Fairbanks School of Public Health, Indianapolis, Indiana, United States.

3. Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, Indiana, United States.

4. Center for Health Information and Communication, Department of Veterans Affairs, HSR&D Service, Roudebush VA Medical Center, Indianapolis, Indiana, United States.

Abstract:

Introduction: Timeliness of data availability is a key performance measure in cancer reporting. Previous studies evaluated timeliness of cancer reporting using a single metric, yet this metric obscures the details within each step of the reporting process. To enhance understanding of cancer reporting processes, we measured the timeliness of discrete cancer reporting steps and examined changes in timeliness across a decade.

Methods: We analyzed 76,259 cases of breast, colorectal and lung cancer reported to the Indiana State Cancer Registry between 2001 and 2011. We measured timeliness for three fundamental reporting steps: report completion time, report submission time, and report processing time. Timeliness was measured as the difference, in days, between timestamps recorded in the cancer registry at each step. We further examined the variation in reporting time among facilities.

Results: Identifying and gathering details about cases (report completion) accounts for the largest proportion of time during the cancer reporting process. Although submission time accounts for a lesser proportion of time, there is wide variation among facilities. One-seventh (7 out of 49) facilities accounted for 28.4% of the total cases reported, all of which took more than 100 days to submit the completed cases to the registry.

Conclusions: Measuring timeliness of the individual steps in reporting processes can enable cancer registry programs to target individual facilities as well as tasks that could be improved to reduce overall case reporting times. Process improvement could strengthen cancer control programs and enable more rapid discovery in cancer research.

Keywords: cancer reporting, cancer registry, data timeliness, cancer surveillance, rapid reporting system, public health reporting.

*Correspondence: ajabour@jazanu.edu.sa

DOI: 10.5210/ojphi.v10i3.9432

Introduction

Cancer registries capture and manage tumor-related data to identify incidents, monitor trends, and support cancer-related research. However, data can take months to become available after diagnosis, which limits the utility of cancer registries for population health surveillance. Previous work has shown that cancer data take an average of 253–426 days after diagnosis to become available at state cancer registries [1].

Timely reporting is an important aspect of cancer surveillance, care delivery, and research. Longer reporting times make monitoring the trends of cancer cases more difficult and limit accessibility to timely data for research [2]. As most intervention programs are data-driven, reporting delays hinder timely intervention in cancer care delivery. Assessments of prevention, screening, and treatment programs are more valuable if they are presented in a timely manner [3]. The need to improve timeliness in cancer reporting is indicated in a series of National Academy of Medicine (NAM) reports [4-6]. Timely data are critical for developing a Rapid Learning Systems (RLS) that could enhance the delivery of cancer care [6,7]. Although no clear consensus exists on the minimum timeliness required for an RLS, the need for improvement is documented [6,8,9]. Achieving timely data requires deeper understanding of the status of registries and the quality of their data.

Prior studies defined timeliness as a single metric, usually as the number of days from the date of diagnosis to the date when data are available for research. For example, Hunt (2004) outlines two common approaches to measuring timeliness [10]. The first approach calculates the proportion of cases abstracted during a certain period compared with the number of cases anticipated in that year [10]. The second approach uses date intervals (or date stamps) to calculate the difference between the date of diagnosis and the date of data entry. Common data points employed in this approach include date of diagnosis, date of admission or first contact, and date of transferring records to the registry [10,11]. One advantage of the second approach is the ability to analyze long-term trends, including assessment of changes over time. Another advantage is the opportunity to examine the different stages within the reporting process.

While delays in reporting cancer data are widely recognized, the distribution of reporting timeliness across the spectrum of cancer reporting processes is not fully understood. Previous studies often evaluated timeliness from a data quality perspective, in which the emphasis is placed on reporting time as an outcome to be measured and not a process to be understood. Therefore, timeliness was evaluated by measuring reporting time as a single process from start to finish, yet the precise duration of each step within the process remains unknown [12-15]. A review of timeliness in reporting of public health data concluded that public health surveillance studies lack detailed descriptions of reporting stages, such as data processing and analyzing [16]. This review also emphasized that both the collection and assessment of time interval data are important elements of any surveillance system or timeliness assessment [16].

Separating the cancer registry reporting process into multiple steps will enhance our understanding of reporting timeliness and represent the issue with respect to time at each sub-process as data move from health care facilities to the registry. Timestamps are a powerful tool for large scale quality monitoring. In this study, we employ timestamps to evaluate timeliness for the sub-processes within cancer reporting as well as the reporting pattern among facilities. Longitudinal data are utilized to assess changes in timeliness for each step and the variation among the different reporting facilities. Understanding variation within sub-processes will enable state registry personnel to direct their efforts to the tasks and facilities needed for targeted improvement.

2 Methods

2.1 Context and Data Retrieval

The study was approved by the Institutional Review Board at Indiana University. To examine timeliness of cancer case reporting, we employed a retrospective cohort study using data from the Indiana State Department of Health (ISDH) Cancer Registry. The ISDH Cancer Registry collects data on all malignant cancer cases required for reporting by federal regulation or the National Program of Cancer Registries [17]. The registry contains information about cancer cases needed for performing epidemiological, preventive, and control studies. Using the ISDH Cancer Registry, we extracted a cohort comprised of patients diagnosed with breast, colorectal, or lung cancer between 2001 and 2011. We retrieved 76,259 de-identified cases representing one of three cancer types: 28,782 breast cancer cases, 19,530 colorectal cancer cases, and 27,947 lung cancer cases.

2.2 Analysis

The cancer registry records dates and timestamps when various components of the reporting process occur. To measure timeliness, we calculated the difference, as the number of days, between each reporting process step (depicted in Figure-1). Descriptive statistics, such as the mean, median and interquartile (box plot), were calculated to assess overall trends as well as variation by reporting source using the Statistical Package for Social Sciences (SPSS) version 21.
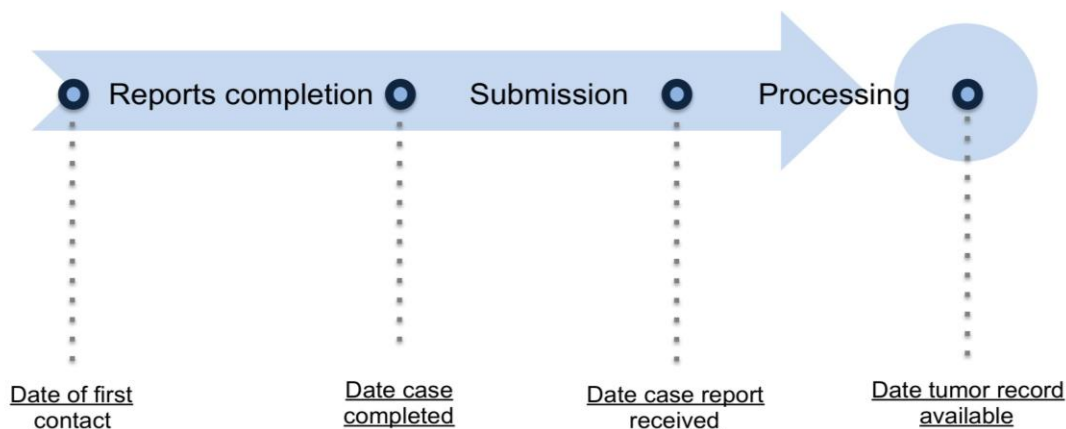

Figure 1: Cancer registry case reporting process depicted as a series of steps performed by registrars that occur at different points in time.

2.3 Reporting Steps

Cancer reporting involves a series of sequential steps. Completion of each step depends on the completion of the previous step; thus, delays in completion of one step will result in delaying the subsequent one and, subsequently, the overall reporting time. The fundamental reporting steps are: case-finding, abstracting, report submission, and report processing and editing.

2.3.1 Report completion

Report completion consists of two steps: case-finding and abstracting. Case-finding is the process of identifying new cancer patients for a given time period. The identification includes records associated with any cancer diagnosis, treatment term, or code indicating a reportable cancer condition [18-20]. Once a case is reviewed and identified as reportable, it is saved in a temporary database known as a "suspense file" in preparation for abstracting [18]. The "suspense file" serves as a temporary database for incomplete cases waiting for any additional exams or procedures to be done and entered into the EHR. Registrars often wait for weeks or months and recheck if the data needed are available before completing the abstracts.

Abstracting is the process of collecting and compiling information about each reportable patient in preparation for sending reports to the state cancer registry. Abstracts are comprehensive reports that include all the cancer-related information required by the receiving registry. The abstract may include data related to demographics, tumor information, staging, diagnostic studies, or treatment. Once the abstract is completed it is saved in preparation for reporting.

Reporting facilities, in general, save completed abstracts in order to report them to registries on a fixed schedule. Once all the required information is collected and the abstract meets the receiving registry's requirements, the abstract is considered complete and ready for reporting. In this study, the report completion time, which includes both case-finding and abstracting time, is defined as the time, measured as number of days, from the "date of first contact 580" in the Commission on Cancer (CoC) standard to the "the date case completed 2090" in the North American Association of Central Cancer Registries (NAACCR) standard (Figure 1). The date of first contact is the date when patients visit the reporting facility for the diagnosis and/or treatment of the tumor.

2.3.2 Report Submission

Completed abstracts are sent to the state registry at fixed intervals. State registries require facilities with high volume of cases to report at higher frequencies. The time from abstract completion to when the state registry receives the report was calculated by measuring the time, as number of days, from "the date case completed 2090" in NAACCR standard to "date case report received by the registry 2111" in the Centers for Disease Control and Prevention (CDC's) National Program of Cancer Registries (NPCR) standard (Figure 1).

2.3.3 Processing

State registries receive reports from multiple facilities. Reports may need to be checked for duplicates and edited before they are analyzed and reported or made accessible to researchers. The registry processing time was calculated by measuring the time, as number of days, from "date case

report received by the registry 2111" in NPCR standard to "date tumor record available at the registry 2113" in NPCR standard (Figure 1).

2.4 Variation among Facilities

To examine variation in reporting time among facilities, we measured the average time taken by individual facilities, annually, for report completion and report submission. Since the requirements for report submission differ based on facilities' annual caseloads, we focused on facilities with higher numbers of cases, due to their greater impact on the overall timeliness of cancer reporting. We included facilities with at least 1,000 cases for the entire study period, which equates to an average of 100 cases per year for the three cancer types included. Facilities with an average of 60 to 149 cases a year are expected to report their cases at least every quarter [19]. Facilities with higher number of cases are expected to report at a higher frequency. Facilities included in this part of the analysis are thereby either expected to report every quarter, every other month, or monthly.

3 Results

3.1 Reporting steps

Figure 2 summarizes annual timeliness for cancer reporting in Indiana. Figures 2-A to 2-C are boxplots showing annual timeliness for the three sub-processes involved in cancer reporting: report completion, report submission, and registry processing. Figure 2-D graphs the median reporting times for each process by year.

Report completion time represents the time taken for case finding and abstracting. The highest report completion time observed was during 2003 and 2004 (Figure 2-A). Report completion time was more consistent during the last five years (2006–2010) compared with the previous five years (2001–2005). A noticeable reduction in timeliness variability among case was observed after 2005.

Report submission time represents the time taken from report completion to report receipt by the state registry. With the exception of 2004, the median time was less than 50 days (Figure 2-B). However, the box plots show that the data were positively skewed. Although most cases (75th percentile) took fewer than 120 days, some cases in the upper whiskers took up to a year.
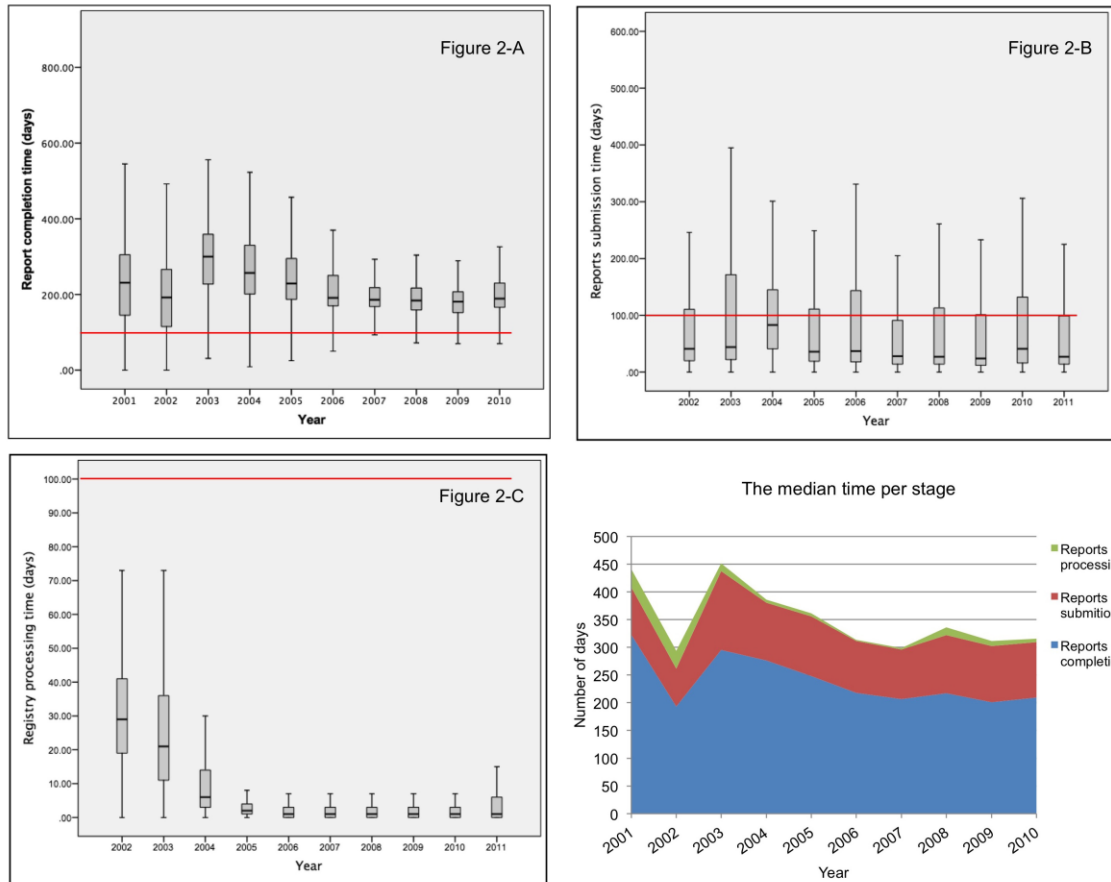
Figure 2-A: Box plot showing the distribution of time taken by facilities to complete reports.

Figure 2-B: Box plot showing the distribution of time taken by facilities to submit completed reports.

Figure 2-C: Box plot showing the distribution of time taken by the central registry to process received reports.

Figure 2-D: comparing the median time for each reporting stage.

The time taken by the state registry to process submissions into the registry was higher during the first two years (2002 and 2003), with medians of around 30 and 20 days, respectively. A noticeable reduction in processing time was observed in 2004 and beyond. The processing time was lowest between 2005 and 2010 (Figure 2-C).

When comparing the median reporting time for each reporting stage we observed that report completion time consumes the largest proportion of the total reporting time followed by report submission time and processing time (Figure 2-D). The stacked area chart also shows the median reporting time appears to be more consistent after 2005 compared to the first few years of the study period.

3.2 Variation among Facilities

The cases retrieved for the entire 10-year study period were reported from 49 unique facilities. Of the 49 facilities, 22 were classified as high volume facilities. The 22 facilities selected accounted for 90.5% of the cases.

Facility reporting time was investigated by distinguishing report completion time from report submission time. The time taken to complete reports varied among facilities (see Figure 3-A); whereas, most facilities' reported completion time ranged from around 175 to 225 days, a few exceeded 300 days.
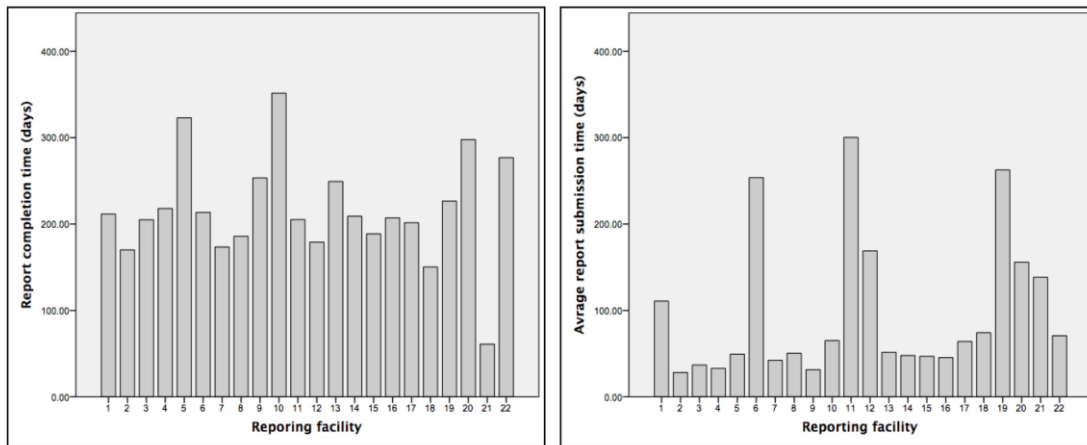

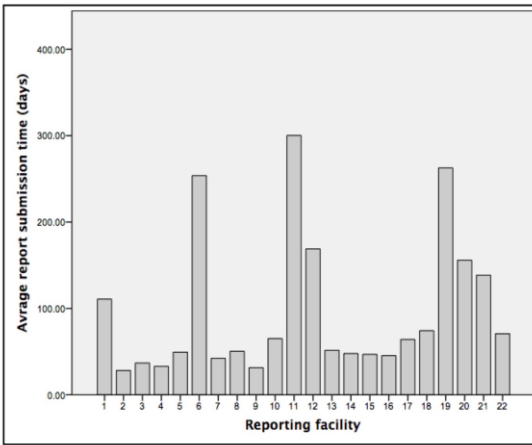Figure 3-A: Reported completion times at high volume facilities.

Figure 3-B: Report submission times at high volume facilities.

Report submission time is presented in Figure 3-B. A noticeable difference among facilities was found in report submission time. Whereas most facilities took less than 50 days on average to send completed reports to the state registry, 7 of the 22 facilities took more than twice that time (around 100 to 300 days). Those 7 facilities accounted for around 28.4% of the total number of reported cases.

4 Discussion

The delays in the timely reporting of cancer data have been discussed in many studies and NAM reports. Previous studies evaluated cancer-reporting time by measuring the total reporting time as a single process [12-15]. To our knowledge, no study has deconstructed the reporting time to examine the time taken per each step within the process. In this study we measured the time taken per each step of the reporting for each facility. One of the advantages of our method is that it provides a total coverage of the reporting sources, which could be utilized for targeted improvement at both the location (facility) and task (stage of reporting). This approach can be very practical for enhancing the quality of registry data and meeting registries certification criteria. Moreover, understanding the details of total time distribution can provide guidance for future research by enabling them to focus on the steps with the greater impact.

We found that report completion accounts for most of the total reporting time. We also found a large variation in report completion times among facilities. Variation can be attributed to many reasons such as the number of registrars working at the facility, or registrars' access to medical records [18]. Prior studies reported more than 50% of the registries and over 70% of the reporting facilities were experiencing shortages in cancer registrars [21-22].

Health information technologies can facilitate some of the reporting activities. Studies reported that the use of Electronic Health Records (EHR) can improve reporting time and support workflow by improving registrars' access to patient information, thus minimizing the time needed for data search and retrieval [22]. Another example is using the Electronic Pathology Reporting Systems (E-Path) to automate some of case finding tasks. E-path utilizes text mining to automate the identification and coding of pathology reports [23]. Around 90–95% of cases identified at the case-finding stage are identified through pathology reports and the time needed for review during case-finding can be reduced using E-path [23]. Studies also indicated that access to Health Information Exchange (HIE) can eliminate many of the barriers related information search and retrieval [18].

In most years, the mean report submission time was less than 50 days. However, there was a noticeable variation in the average reports submission time among facilities. Although a longer submission time was found only in 7 of the 49 facilities, those 7 facilities accounted for around 28.4% of the total number of reported cases. Maintaining timely reporting for all cases is important because cancer registry standards require high completion rates. The Surveillance, Epidemiology, and End Results Program (SEER) standard, for instance, requires 98% of cases to be reported within 22 months of the date of diagnosis [24]. Another example is the NAACCR standard that requires 90% of the cases to be reported within 23 months from the date of diagnosis for silver certification and 95% for golden certification [25].

Reports are usually transferred electronically from individual facilities to state registries. Unlike case-finding and abstracting, the task of submitting reports is easier and less time-consuming. After abstracting, completed reports are stored at a hospital's database to be submitted at fixed intervals (monthly, bi-monthly, quarterly, etc.). Encouraging facilities to submit completed reports more promptly can potentially improve overall registry performance, especially if facilities with longer submission times are targeted. Automated reports submission could also be employed to send each report individually upon completion instead of sending them as batches on a fixed schedule.

By separating the time taken by facilities to complete and send reports from the time taken to edit and process the reports at the state registry, we found that the registry takes a short amount of time for processing. Moreover, a remarkable reduction in registry processing time was found after 2004. In 2004, the time taken by the registry to process and edit cases received began to decrease from 20 to 30 days to fewer than 10 days. Additionally, the variation in processing time among cases decreased substantially after 2004 (less than 10 days), whereas in 2002 and 2003 processing time extended up to 75 days. The ISDH cancer registry went through some major changes in 2004 [26]. Although causation cannot be inferred, changes were concurrent with changes in the median reports processing time at the central registry. One of these changes was the implementation of a registry system that uses file transfer protocol (FTP) (standard for exchanging files online) [26]. It was implemented as an alternative to physically sending discs allowing real-time transfer of data as opposed to postal mail [26]. Surprisingly, using FTP did not seem to impact the report submission time. For the most part, report submission time appears to have been consistent over

the years. One explanation might be that submission time is constrained by reporting policies and requirements guidelines. Reporting facilities are required to submit completed reports at a fixed interval based on their annual caseloads. For instance, hospitals with an average of 1 to 59 cases annually are required to report their cases at least once each year while hospitals with an average of 300 or more cases annually are required to report them monthly [17].

4.1 Limitation

One limitation was the inability to distinguish the time needed for case-finding from the time needed for abstracting. Distinguishing the timeliness of these steps will allow a focus on the step with the longer reporting time. Knowing the timeliness for each step could provide insights into the potential impact of adopting technologies such as E-Path. E-path can automate many of the case finding activities but there are some challenges, such as the cost associated with the software implementation and infrastructure [27]. The ability to measure case-finding timeliness can be important for making such an investment decision.

4.2 Future research

Our study showed variation among facilities in both report completion and report submission times. As reporting time can be impacted by many factors such as facility type, resources, technologies, and reporting practice, the variation in these resources across facilities is unknown. For consistent reporting performance, further studies are recommended to examine the variation in these factors among facilities.

5 Conclusion

Our study presents a technique for statewide monitoring of public health reporting. Monitoring the reporting process is fundamental for quality of data. In 2013, the NAM recommended a framework to improve the quality of cancer care by investing in health care information technology as a core objective for translating evidence into clinical practice [6]. Establishing an RLS for cancer would enable the use of cancer registry data to inform decision-making and planning to achieve better treatment outcomes [6].

References:

1. Jabour AM, Dixon BE, Jones JF, Haggstrom DA. 2016. Data Quality at the Indiana State Cancer Registry: An Evaluation of Timeliness by Cancer Type and Year. *J Registry Managment*. 43(4), 168-73. PubMed

2. Midthune DN, Fay MP, Clegg LX, Feuer EJ. 2005. Modeling reporting delays and reporting corrections in cancer registry data. *J Am Stat Assoc*. 100(469), 61-70. https://doi.org/10.1198/016214504000001899

3. Meyer AM, Carpenter WR, Abernethy AP, Stürmer T, Kosorok MR. 2012. Data for cancer comparative effectiveness research. *Cancer*. 118(21), 5186-97. PubMed https://doi.org/10.1002/cncr.27552

4. Simone JV, Hewitt M. Enhancing data systems to improve the quality of cancer care, National Academies Press2000.

5. Hewitt M, Simone JV. Ensuring quality cancer care, National Academies Press1999.

6. Institute of Medicine, Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis (2013), 2013.

7. Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, et al. 2013. Rapid Learning health care in oncology'–an approach towards decision support systems enabling customised radiotherapy. *Radiother Oncol*. 109(1), 159-64. [PubMed](https://doi.org/10.1016/j.radonc.2013.07.007) https://doi.org/10.1016/j.radonc.2013.07.007

8. Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, et al. 2010. Rapid-learning system for cancer care. *J Clin Oncol*. 28(27), 4268-74. [PubMed](https://doi.org/10.1200/JCO.2010.28.5478) https://doi.org/10.1200/JCO.2010.28.5478

9. Meyer A-M, Carpenter WR, Abernethy AP, Stürmer T, Kosorok MR. 2012. Data for cancer comparative effectiveness research. *Cancer*. 118(21), 5186-97. [PubMed](https://doi.org/10.1002/cncr.27552) https://doi.org/10.1002/cncr.27552

10. A. National Cancer Registrars. Cancer Registry Management: Principles & Practice, Kendall/Hunt Publishing Company2004.

11. Jabour AM, Dixon BE, Jones JF, Haggstrom DA. 2016. Data Quality at the Indiana State Cancer Registry: An Evaluation of Timeliness by Cancer Type and Year. *J Registry Manag*. 43(4), 168-73. [PubMed]

12. Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, et al. 2009. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *Eur J Cancer*. 45(7), 1218-31. [PubMed] https://doi.org/10.1016/j.ejca.2008.10.037

13. Sigurdardottir LG, Jonasson JG, Stefansdottir S, Jonsdottir A, Olafsdottir GH, et al. 2012. Data quality at the Icelandic Cancer Registry: Comparability, validity, timeliness and completeness. *Acta Oncol (Madr)*. 51(7), 880-89. [PubMed](https://doi.org/10.3109/0284186X.2012.698751) https://doi.org/10.3109/0284186X.2012.698751

14. Tomic K, Sandin F, Wigertz A, Robinson D, Lambe M, et al. 2015. Evaluation of data quality in the National Prostate Cancer Register of Sweden. *Eur J Cancer*. 51(1), 101-11. [PubMed](https://doi.org/10.1016/j.ejca.2014.10.025) https://doi.org/10.1016/j.ejca.2014.10.025

15. Smith-Gagen J, Cress RD, Drake CM, Felter MC, Beaumont JJ. 2005. Factors associated with time to availability for cases reported to population-based cancer registries⋆. *Cancer Causes Control*. 16(4), 449-54. [PubMed](https://doi.org/10.1007/s10552-004-5030-0) https://doi.org/10.1007/s10552-004-5030-0

16. Jajosky RA, Groseclose SL. 2004. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health*. 4(1), 29. [PubMed](https://doi.org/10.1186/1471-2458-4-29) https://doi.org/10.1186/1471-2458-4-29

17. INDH., Policy and Procedure Manuals for Reporting Facilities. Indiana State Department of health Cancer Registry., 2015.

18. Jabour AM, Dixon BE, Jones JF, Haggstrom DA. 2018. Toward Timely Data for Cancer Research: Assessment and Reengineering of the Cancer Reporting Process. *JMIR Cancer*. 4(1). PubMed https://doi.org/10.2196/cancer.7515

19. Indiana State Department of Health, Policy and Procedure Manual 2016.

20. ACS. Registry Manuals, American College of Surgeons, 2015.

21. Tangka F, Subramanian S, Beebe MEC, Trebino D, Michaud F. 2010. Economic assessment of central cancer registry operations, Part III. *J Registry Manag*. 37(4), 152-55. PubMed

22. S.H. Houser, S. Colquitt, K. Clements, S. Hart-Hester, The impact of electronic health record usage on cancer registry systems in Alabama, Perspectives in health information management/AHIMA, American Health Information Management Association 9(Spring) (2012).

23. NAACCR. Electronic Pathology (E-Path) Reporting Guidelines, The North American Association of Central Cancer Registries, 2006.

24. NCI. Cancer Incidence Rates Adjusted for Reporting Delay. http://surveillance.cancer.gov/delay/%3E, 2014 03/01/2015).

25. North American Association of Central Cancer Registries. Certification Criteria. http://www.naaccr.org/Certification/CertificationLevels.aspx.%3E, 2018).

26. ISDH. Indiana State Cancer Registry News Briefs Issue 4, Indiana State Department of Health 2004.

27. NAACCR. Electronic Pathology Reporting Guidelines, The North American Association of Central Cancer Registries, 2011.