RADIATION ONCOLOGY NEWS FOR ADMINISTRATORS

Volume 27 • Number 4 • 2017

UNDERSTANDING CANCER REGISTRY DATA COLLECTION AND DATA MINING

By Linda Reimers

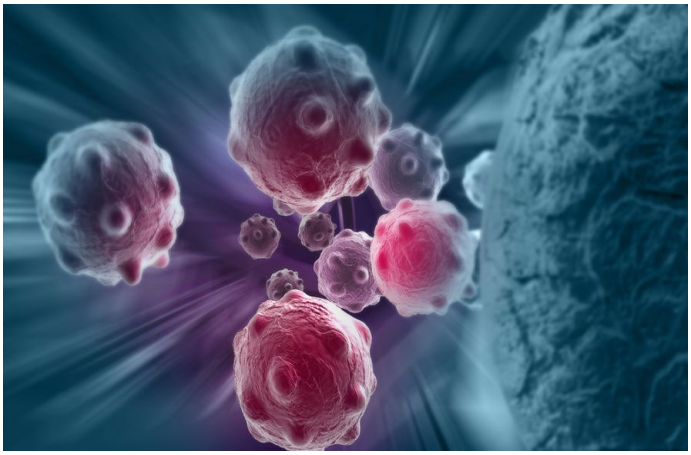
Linda Reimers

Cancer data collection is mandatory, yet the data are underutilized. Hospitals and other healthcare providers nationwide spend thousands of dollars annually collecting cancer-related information; however, some cancer program leaders fail to capitalize on the insights the data can provide. For example, data can be used to evaluate quality of care, to plan for programs and services, and to identify diagnostic and treatment pattern trends at one's facility.

Cancer data collection began in 1913;[1] the data set has since become more sophisticated and is used worldwide. In 1992, the United States Congress passed the Cancer Registries Amendment Act, which established the National Program of Cancer Registries (NPCR)[1] to support state cancer registries and thereby assure national data analysis. Initially, only 38 states had cancer registries; now, all 50 states and the District of Columbia have registries. The NPCR maintains American national cancer statistics.[1a] The North American Association of Central Cancer Registries (NAACCR) and the National Cancer Institute/Surveillance, Epidemiology, and End Results (SEER) are two other important registries. Clinicians, researchers, patients and patients' families use the American Cancer Society (ACS) Facts and Figures data, which are drawn from NPCR, NAACCR and SEER.

Who determines what data are collected?

International and national organizations set the standard for cancer registry data collection, including the World Health Organization (WHO); the Commission on Cancer (CoC); the American Joint Committee on Cancer (AJCC); the American Cancer Society; the National Cancer Institute/Surveillance, Epidemiology, and End Results; the National Program of Cancer Registrars; and the North American Association of Central Cancer Registries.



© vitanovski / 123RF Stock Photo

To provide comparative data that can be used for quality measurement, research, screening and prevention, the standard data set used is consistent from hospital to hospital and from state to state. By law, all clinicians and hospitals must report, at minimum, their cancer data to the state central cancer registry within six months from the date of first contact for every patient having a newly confirmed cancer diagnosis.

Who collects the data?

A certified tumor registrar (CTR) is a specialist in cancer data abstraction for reporting, study and research purposes. To become a certified CTR, candidates must meet various educational and work experience requirements. The national exam covers questions on data collection, data quality assurance, analysis and data usage, operations and management, cancer committee and cancer conference, and activities unique to central registries. Once certified, CTRs must complete continuing education requirements to remain current with the ongoing changes in oncology care.

[a] On the NPCR website, one can access official federal statistics on cancer incidence from registries with quality data, and cancer mortality statistics.

Commission on Cancer-accredited hospitals are required to use CTRs for their data abstraction. Hospitals without an accredited program, as well as physician offices, pathology labs and standalone surgical centers, typically use uncertified personnel who have some cancer registry training.

Currently, there are not enough CTRs to provide services for all programs.[b] According to 2016 statistics from the National Cancer Registrars Association (NCRA), the national certifying body, there were 5,143 domestic and 119 non-domestic CTRs.[2]

When are the data collected?

Initially, cases were abstracted at approximately five to six months following the date of diagnosis, which allowed for all diagnostic and treatment information to be collected at one specified point in time. To impact care in a timelier manner, abstraction is now performed concurrently or as soon after diagnosis as possible, with continual updates as various treatment modalities are completed.

Where does the cancer registry obtain its information?

All methods of diagnosis and the full scope of treatment are collected for each cancer diagnosis (i.e., primary site) a patient receives. An aggregated report is developed, drawing information from any combination of inpatient and outpatient care, freestanding radiation therapy and infusion centers, pathology labs and private physician practices.

What information is collected in a cancer registry?

The information collected is more comprehensive than clinicians and hospital administrators may realize. The following data are routinely collected for each case.

Demographics:

- Zip code
- County code
- State code
- Insurance status or Insurance class
- Marital status
- Race
- Age
- Gender

Clinical Diagnostic Factors:

- Physical exam
- Diagnostic biopsy results
- Radiology results
- Endoscopy results
- Laboratory testing results
- Comorbidities and complications
- Class of case (location of diagnosis and treatment modalities)

Some of the data are recorded as text that may be queried from the database[c] through a search function. Often clinical diagnostic data are used to ensure compliance with national prognostic indicator guidelines, use of equipment and appropriateness of services, and patterns of diagnostic workup. The co-morbidities and complications allow analysis for conditions such as chronic obstructive pulmonary disease (COPD), or an anastomotic leak following a surgical resection and implications related to outcomes. The class of case coding allows for analysis of in- and out-migration, strategic planning and patterns of care.

Predisposing Factors:

- Smoking history
- Alcohol history
- Asbestos exposure
- Occupation
- Family history of cancer
- Personal history of cancer

These data elements are primarily required for reporting to state central registries. Some hospitals may collect additional predisposing factors. The coding definitions must be clearly defined for the data to be manageable and meaningful.

Tumor Description:

- Cancer site
- Laterality
- Grade of tumor
- Histology of tumor
- Pathologic confirmation
- Confirmation of multiple tumors and histology's

The definitions for these data elements continue to be defined by the regulatory agencies and remain the same for any CTR abstracting cancer cases. These definitions allow for data analysis that is consistent, comparable and valuable. The grade indicates the resemblance of cancer tissue to normal tissue. Diagnostic confirmation indicates if a case is pathologically or clinically diagnosed. Following specific rules for ambiguous terms allows for accurate collection of clinically diagnosed

[b] The shortage of CTRs has been documented in several online articles. Moreover, according to the American Hospital Association (AHA), there are 5,564 hospitals in the U.S and large hospitals require several/many CTRs to meet their reporting needs.[3]

[c] Each accredited hospital has its own cancer registry database, while a hospital without an accredited program might only be submitting data to its state central cancer registry and not maintaining data for individual queries.

cases. The term "suspicious for" in a radiology report is considered reportable; however, if "worrisome" or "rule out" is stated, the case is considered not reportable unless additional workup confirms a diagnosis. Multiple tumors and histology's are also clearly defined by the regulatory agencies to indicate if a case is a new primary cancer, or a recurrence of a previously diagnosed cancer and the appropriate histology code assigned.

For abstraction, all hospitals accredited by the Commission on Cancer must employ either a CTR or a registrar who earns his/her CTR certification within three years of employment. Hiring an experienced CTR or providing a mechanism for formal training will facilitate collection of the most complete and accurate data. A non-certified registrar unfamiliar with the voluminous coding rules may provide inaccurate or inconsistent data to the reporting facility and external stakeholders, which is why cancer registry certification is crucial to ensure the integrity of cancer data.

Extent of Disease / Stage of Cancer:

- Size of tumor
- Extension of tumor (depth of invasion or extension to surrounding organs/tissue)
- Lymph node involvement
- Metastatic involvement
- Margin status
- Site-specific factors

Cancer staging determines the treatment of cancer and is one of the most frequently requested data fields for reporting. Most hospitals perform quality control on the accuracy and completeness of the extent of disease coding. However, quality control varies by facility: some facilities perform a weekly random sample review of cases abstracted; others may perform quality control monthly or quarterly prior to data submission; and others may wait for the state central registry to perform quality control once the data have been submitted. If errors are identified at the state level, the cases are returned to the facility to be corrected then resubmitted.

The regulatory agencies include edit sets for review and identification of potential errors; physicians are also appointed to complete quality control reviews of registry data. The field definitions and guidelines are very specific to ensure accuracy and standardization. The data elements collected continue to change as oncology research evolves. Examples of data elements that change include PSA level for prostate cancer and HER2Neu amplification and Oncotype DX for breast cancer.

First Course of Treatment:

- Start dates allow for sequencing of complete treatment plan and assist in evaluating time intervals from diagnosis to treatment and/or from treatment to recurrence.
- If treatment is not administered, the reason for no treatment is collected; or all treatment modalities to document patient refusal of treatment; contraindications that prevented treatment; or that treatment was recommended, but the facility was unable to confirm it was administered (i.e., patient went elsewhere for treatment).
- Includes all modalities: surgery, radiation, chemo, hormonal therapy, immunotherapy, etc.

Each treatment modality has standardized fields with specific definitions (see Table 1: Radiation Therapy Required Data Fields).

Outcome Data:

- Date of first recurrence
- Recurrence type (local, regional, distant/metastatic)
- Vital status
- Cancer status
- Subsequent new primary cancer

Customized Data Fields:

- Hospital defined
- Related to hospital quality initiatives
- Must be clearly defined

Customized data fields are created by hospitals/programs that are studying a specific initiative, treatment or process. Examples include documentation for dietary consultations, nursing interventions or appropriateness for a survivorship care plan.

How is the cancer registry database created?

Demographic and clinical information is translated into a symbolic statement based on the standardized rules (coding) used by all CTRs. On average, seven different reference manuals are used daily to follow coding definitions and regulations. Interoperability within a medical facility's electronic health records (EHR) has allowed some information (e.g., patient demographics) to be downloaded into the cancer registry databases. However, interoperability has not significantly impacted the way registrars do their jobs—anything that is downloaded still has to be reviewed by the registrar to ensure accuracy and completeness.

How is the quality of the data assessed?

Standard edit checks are created by each state cancer registry, the Commission on Cancer and NAACCR to validate the quality of data collected. Within the medical facility, quality control policies and procedures are required and implemented to ensure the quality and integrity of data. These methods include random sampling, physician and peer review, and benchmark comparison studies. Internal and external audits, along with software edit checks, are imperative for transparent data reporting and accountability of the statistics being published.

Facility Oncology Registry Data Standards (FORDS) Radiation Oncology Information located on pages 237–259 Items collected from Radiation Oncology Treatment Documentation	
Data Field	**Definition**
Date Radiation Started	Date patient began radiation
Date Radiation Ended	Date patient completed radiation
Location of Radiation Treatment	Whether all at the reporting facility or elsewhere, possible combination of regional dose at the reporting facility and boost elsewhere etc.
Radiation Treatment Volume	Anatomic target of the most clinically significant radiation therapy, breast, breast and lymph nodes, chest wall, brain, etc.
Regional Treatment Modality	Type of therapy delivered as the regional dose, photons, electrons, IMRT, 3D conformal, brachytherapy, stereotactic radiosurgery etc.
Regional Dose	Total regional dose coded in cGy
Boost Treatment Modality	Type of therapy delivered as the boost dose, photons, electrons, IMRT, 3D conformal, brachytherapy etc.
Boost Dose	Total boost dose coded in cGy
Number of Treatments to this Volume	Number of fractions inclusive of regional and boost doses
Other items collected from dictated reports	
Data Field	**Definition**
Reason for No Radiation	Reason radiation was not administered if it is the standard of care; patient refusal, contraindicated, etc.
Radiation Surgery Sequence	Sequence of radiation if surgery was performed

Table 1: Radiation Therapy Required Data Fields
Cancer registries are required to capture radiation treatment information for each patient receiving radiation therapy as standard care for their cancer diagnosis, or the reason why a patient refused radiation therapy. The Commission on Cancer's Facility Oncology Registry Data Standards (FORDS) Manual outlines the definitions for collecting and reporting radiation treatment information. Similar definitions are established for surgery, chemotherapy, hormonal therapy and immunotherapy.

Conclusion

Cancer data collection is mandated nationwide; however, the data are not fully utilized. The vast amount of data collected are invaluable to physicians, clinical researchers, epidemiologists, clinical practitioners and hospital administrators. Cancer registry data play a part in patient education, screening and early detection efforts, but the data also allow clinicians to evaluate the effectiveness of diagnostic and treatment advances, to discover areas where additional research, development and improvement are needed, and to improve patient outcomes. Ultimately, the data collected may be central to finding a cure for cancer.

References

1. Centers for Disease Control and Prevention. Cancer Registries Amendment Act. October 13, 2015.
2. National Cancer Registrars Association. 2017.
3. American Hospital Association. Fast Facts on US Hospitals. January 2017.

About the Author: Linda Reimers, BS, CTR, is the director of business development for Registry Partners, a national provider of data abstraction, registry management and consulting services.

Editor's Note: Thank you to Dawn DeBolt, BS, RHIA, CTR, for her editorial contributions to this article.