The National Cancer Data Base: A Powerful Initiative to Improve Cancer Care in the United States

Karl Y. Bilimoria, MD, Andrew K. Stewart, MA, David P. Winchester, MD, and Clifford Y. Ko, MD, MS, MSHS

The National Cancer Data Base (NCDB) is a nationwide oncology outcomes database that currently collects information on approximately 70% of all new invasive cancer diagnoses in the United States each year and serves as a powerful clinical surveillance and quality improvement mechanism for cancer programs participating in the American College of Surgeons (ACoS) Commission on Cancer (CoC) approvals program. Currently, the NCDB receives over one million cancer case reports annually from more than 1,430 hospitals. The NCDB now contains data on more than 21 million cancer patients diagnosed between 1985 and 2005, and is recognized as the largest clinical registry in the world. NCDB data are used to explore trends in cancer care, to examine regional and national benchmarks, and to serve as the basis for quality improvement activities (http://www.facs.org/cancer/ncdb).

The purposes of this review are: (1) to describe the NCDB and the data collected; (2) to discuss how the NCDB can be used to study clinical outcomes and the quality of cancer care in the United States; and (3) to describe the clinical care improvement tools provided by the NCDB and the CoC to participating hospitals.

BACKGROUND: THE COMMISSION ON CANCER AND THE NCDB

The CoC is a multidisciplinary consortium of professional organizations that strive to improve cancer care through setting standards, prevention, research, education, and the monitoring of comprehensive cancer care. The CoC also accredits hospitals as cancer centers, based on their ability to provide a broad range of cancer-related services and specialists. The NCDB was started in 1989 by the ACoS and the CoC in conjunction with the American Cancer Society. The NCDB is housed at the ACoS headquarters in Chicago, Illinois. When the database was established any hospital could voluntarily report data to the NCDB, and during the mid 1990s, more than 2,000 hospitals were contributing cases. In 1996, all CoC-approved hospitals were required to report cancer cases to the NCDB, and during the mid 1990s, more than 2,000 hospitals were contributing cases. In 1996, all CoC-approved hospitals were required to report cancer cases to the NCDB, and in 2001, participation and the associated advantages of reporting to the NCDB were limited to those hospitals who had earned CoC approval.

PARTICIPATING HOSPITALS

The CoC classifies approved hospitals into teaching/research hospitals, comprehensive community cancer centers, and community cancer centers. Teaching/research hospitals must be primarily affiliated with a medical school or National Cancer Institute (NCI)-designated comprehensive cancer center. Community cancer center designations are
based on a combination of annual case volume and the ability to provide a wide range of cancer-specific personnel and services. Of the 1,430 hospitals that report to the NCDB, approximately 20% are classified as teaching/research hospitals, 39% are comprehensive community cancer centers, 35% are community cancer centers, and 6% are other cancer centers that include Veterans Health Administration (VA) hospitals (Fig. 1).

Reporting hospitals are located in 49 US states and also Puerto Rico. The cancer programs that report to the NCDB include 19 of 20 adult National Comprehensive Cancer Network (NCCN) hospitals, 33 of 37 NCI-designated cancer centers, and 69 of 121 major inpatient VA hospitals. Some hospitals may elect not to gain CoC approval and thus not contribute data to the NCDB, but the reasons for this are unclear. The 1,430 hospitals that have approved cancer programs and report data to the NCDB account for approximately one third of the inpatient hospitals in the United States. While the NCDB does collect the majority of incident cases in the United States, a hospital selection bias may exist as only CoC-approved hospitals contribute data to the NCDB.

**HOW CASES ARE REPORTED**

In 1973 federal legislation made invasive cancer a reportable disease, and in 1992 federal legislation further mandated that all cancer cases be reported to state cancer registries. The CoC requires its approved programs to abstract and follow all malignant tumors diagnosed and/or initially treated at the hospital. Annually, CoC-approved hospitals report over one million cancer cases to the NCDB. Approximately 940,000 cases are new invasive cancer diagnoses and account for about 70% of all new cancer diagnoses in the United States, based on annual incidence estimates from the American Cancer Society (Table 1). In addition, the NCDB takes receipt of over 50,000 in situ cancers and 18,000 benign and borderline brain and central nervous system tumors. Patients diagnosed and treated in a physician’s office setting, those who undergo consultation to confirm a diagnosis or formulate a treatment plan, and cases reviewed by pathology, but who never enter the hospital for any aspect of their care, are not reported to the NCDB.

Data reporting to the NCDB is highly standardized and similar to other state health departments and federal cancer registry data systems, including the NCI’s Surveillance Epidemiology and End Results (SEER) program and the Centers for Disease Control’s National Program of Cancer Registries (NPCR). Data reported from CoC-approved hospitals are abstracted from patient charts by Certified Tumor Registrars (CTR) who undergo training specific to cancer registry operations. In recent years, the training and certification requirements for CTRs have increased significantly, focusing on many of the nuances of cancer care. If data points are missing, registrars can query the treating physicians to obtain the necessary data to complete the record.

All data submitted to the NCDB undergo a battery of data integrity checks.
quality-assurance measures was increased significantly for the diagnosis year 2002 and expanded further for 2005 diagnoses. Case records that fail to meet a standardized set of requirements are identified and returned to the hospital. Internal studies comparing the quality of data reporting from large teaching/research hospitals and small community centers revealed no distinct differences in the completeness or accuracy of reported data.13

Each year, the NCDB undergoes extensive internal quality monitoring and validity reviews. Aggregate data are examined and verified in advance of the release of annual reports created for each cancer site, and the release of data for analytic purposes.14 In addition to these extensive automated electronic data-quality checks, site surveys from the CoC evaluate each hospital’s data and data-collection processes once every three years, to ensure that high-quality case information is being reported to the NCDB in an accurate and timely fashion. The NCDB has also undergone multiple data-quality reviews.15–18

Once diagnosed and/or treated at a hospital with a CoC-approved cancer program, the patient will be followed by the hospital registrar and the case will be reported to the NCDB regardless of whether the patient receives part of their care at a non-CoC hospital. For example, a patient is diagnosed with a malignancy at hospital A which does not report to the NCDB, undergoes surgery at hospital B which reports to the NCDB, and then returns to hospital A for adjuvant therapy. The CoC requires that the cancer registry at hospital B record the information regarding diagnosis from hospital A and enter follow-up data from hospital A regarding subsequent treatment, recurrence, and survival (active surveillance). If both hospitals were CoC approved, two case records for the same patient would be reported to the NCDB, termed a ‘clinical duplicate.’ These duplicate cases are identified using a computer algorithm, and the case

| Cancer Site | Number of invasive malignancies reported to the NCDB | Estimated number of annual cancer cases in the United States | Proportion of invasive malignancies reported to the NCDB |
|-------------|-----------------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------|
| Anus        | 3,476                                               | 3,990                                                       | 87.1%                                                  |
| Bladder     | 40,647                                              | 63,210                                                      | 64.3%                                                  |
| Bones and joints | 2,178                       | 2,570                                                       | 84.7%                                                  |
| Breast      | 141,741                                             | 212,930                                                     | 66.6%                                                  |
| Cervix      | 9,912                                               | 10,370                                                      | 95.6%                                                  |
| Colon       | 65,576                                              | 104,950                                                     | 62.5%                                                  |
| Esophagus   | 11,296                                              | 14,520                                                      | 77.8%                                                  |
| Gallbladder and other biliary | 6,205                               | 7,480                                                       | 83.0%                                                  |
| Kidney, renal pelvis, ureter | 33,513                             | 38,670                                                      | 86.7%                                                  |
| Larynx      | 9,541                                               | 9,880                                                       | 96.6%                                                  |
| Leukemia    | 20,366                                              | 34,810                                                      | 58.3%                                                  |
| Liver and intrahepatic bile ducts | 11,440                              | 17,550                                                      | 65.2%                                                  |
| Lung and bronchus | 141,757                           | 172,570                                                     | 82.1%                                                  |
| Lymphoma    | 54,034                                              | 63,740                                                      | 84.8%                                                  |
| Melanoma    | 28,865                                              | 59,580                                                      | 48.4%                                                  |
| Oral cavity and pharynx | 23,824                             | 29,370                                                      | 81.1%                                                  |
| Other sites | 64,303                                              | 88,610                                                      | 72.8%                                                  |
| Ovary       | 15,225                                              | 22,220                                                      | 68.5%                                                  |
| Pancreas    | 23,724                                              | 32,180                                                      | 73.7%                                                  |
| Prostate    | 117,995                                             | 232,090                                                     | 50.8%                                                  |
| Rectum      | 30,676                                              | 40,340                                                      | 76.0%                                                  |
| Small intestine | 4,001                              | 5,420                                                       | 73.8%                                                  |
| Soft tissue | 7,350                                               | 9,420                                                       | 78.0%                                                  |
| Stomach     | 13,882                                              | 21,860                                                      | 63.5%                                                  |
| Testis      | 5,729                                               | 8,010                                                       | 71.5%                                                  |
| Thyroid     | 23,648                                              | 25,690                                                      | 92.1%                                                  |
| Uterus      | 30,571                                              | 40,880                                                      | 74.8%                                                  |
| TOTAL       | 941,675                                             | 1,372,910                                                   | 68.6%                                                  |

1 Estimated number of invasive malignancies in the United States for 2005 based American Cancer Society estimates.63 The American Cancer Society has since revised their estimation methodology by basing their calculations on data from state cancer registries rather than on SEER regions. Estimates of cancer incidence have increased with this new methodology.

2 American Cancer Society estimates for bladder cancer include in situ disease.

Note: The NCDB requires reporting of all in situ cancers for most cancer sites; however, these are not included in the table above. For 2005 diagnoses, more than 50,000 in situ malignancies were reported to the NCDB. The NCDB also receives benign, as well as malignant, brain tumors.
record containing the most complete data is retained in the analytic dataset.

DATA ITEMS COLLECTED

The NCDB collects data on patient factors, tumor characteristics, staging details, surgical and adjuvant treatments, and outcomes, using nationally standardized coding guidelines. All CoC-approved hospitals submitting data to the NCDB are required to use the Facility Oncology Registry Data Standards (FORDS) manual to guide abstracting and coding. Additional guidelines for coding histopathological characteristics and stage of disease are established by the World Health Organization (WHO) in its International Classification of Disease for Oncology (ICD-O) and the American Joint Committee on Cancer (AJCC), respectively.

Tumor characteristics include size, the presence and location of nodal and distant metastases, number of nodes positive and examined, grade, margin status, and both clinical and pathological staging information. Clinical stage is a pretreatment assessment that is recorded prior to surgery, and if the patient undergoes resection, a pathological stage is recorded separately. Beginning with cases diagnosed in 2004, the NCDB started collecting Collaborative Stage data elements, in addition to physician-recorded AJCC clinical and pathological tumor–node–metastasis (TNM) elements and overall stage. For the purpose of assessing comorbidity, the International Classification of Disease, 9th Edition (ICD-9-CM), secondary diagnosis codes, excluding cancer diagnoses and post-operative complications, are recorded for each patient.

The NCDB also collects a wide range of data regarding surgical treatment and adjuvant therapy, including radiation, chemotherapy, hormone therapy, and other systemic treatments. Patients are also followed to assess outcomes including readmission within 30 days, recurrence, and survival. CoC-approved hospitals are expected to provide follow-up on at least 90% of all known living patients annually.

Through linkage to tertiary data sources, area-based socioeconomic status and estimates of travel distance/time to the reporting hospital can be obtained based on each patient’s home zip code at the time of diagnosis. In addition, hospital type and structural characteristics can be determined through linkage with the CoC’s facility information profile system (FIPS) file of self-reported hospital structural features, resources, and services related to oncology.

DATA CONFIDENTIALITY

The ACoS and the NCDB take great care to protect patient and hospital anonymity. The American College of Surgeons has a business associate agreement with each of its CoC-approved hospitals that prohibits the CoC from revealing hospital-specific information. Data reported to the NCDB are retrospective in nature. No direct patient identifiers are collected. Analyses are reported only at the aggregate level to assist hospital cancer programs with quality assurance and improvement. Reported results are in compliance with the privacy requirements of the Health Insurance Portability and Accountability Act (HIPAA).

CLINICAL OUTCOMES STUDIES

The NCDB offers a large, nationally applicable sample to examine oncology practice patterns and outcomes in the United States. Single institutions studies are typically from large centers that focus on particular malignancies. With information from a wide range of hospitals in diverse geographic locations, a better understanding of the variability in care across the country can be analyzed, and, in particular, smaller, non-specialty hospitals can be examined in order to compare differences by hospital type and volume of patients. This is particularly important since the majority of cancer patients in the United States are not treated at high-volume, specialized centers.

Initially, the large number of patients available for analysis through the NCDB resulted in a considerable body of descriptive literature on the clinical and epidemiologic characteristics of common cancers in the United States. Subsequent studies have been more hypothesis driven and have addressed a broad range of cancers and oncological issues. Several studies have used the NCDB to examine surgical and adjuvant therapy treatment trends and utilization. The large sample size offered by the NCDB allows studies of rare cancers (i.e., pancreatic neuroendocrine tumors, metaplastic breast cancer, and melanoma in children) and subset analyses (i.e., stage I pancreatic cancer patients undergoing resection, 1–2 cm papillary thyroid cancers, T3N0 colon cancer patients) with sufficient statistical power to make reasonable inferences. As AJCC staging information is collected, numerous reports have validated staging systems, and the AJCC has relied heavily on the NCDB to provide data for the
Cancer Staging Manual and to guide staging-system modifications. Furthermore, studies using NCDB data have addressed important clinical questions in areas where clinical trials may be challenging (i.e., local excision of rectal tumors, and extent of surgery for papillary thyroid cancer).

The broad range of hospital types reporting to the NCDB have also facilitated an array of health services research. The NCDB has been used to compare differences in treatment and outcome by hospital type and volume of patients. Data from the NCDB have also been used to examine concordance with established quality measures and standards of clinical care. Multiple studies have also examined cancer care for minorities and disadvantaged populations, and a recent article examined how clinical trials could be developed to specifically target minority populations. Statistical methodology pertaining to outcomes research has also been examined for multiple cancer sites. In addition, the NCDB can be used to obtain details beyond those typically collected, by requiring cancer registrars to ascertain additional data points on a specific cohort of patients.

Although there are limitations associated with the use of cancer registry data, appropriate, hypothesis-driven questions can be effectively investigated. Administrative datasets are based on data with financial motivations, such as claims or reimbursement, whereas cancer registries are designed and collect data with the intent of studying cancer epidemiology, treatment trends and utilization, outcomes, and the quality of cancer care. In addition, cautious investigations regarding the efficacy of treatments can also be examined. Results from these studies are useful and important in order to generate hypotheses for future prospective studies and clinical trials.

QUALITY IMPROVEMENT

Although the NCDB is a tremendous repository of oncologic data, perhaps the most important aspect is its ability to benchmark hospitals, and feed information back to the public and to participating hospitals. The NCDB offers public benchmark reports for the eleven most commonly diagnosed solid-organ tumors in the United States. These reports currently provide data on more than 5.9 million patients, allowing users to define queries based on patient gender, age, ethnicity, histology, stage, first-course therapy, type of surgical resection, hospital type, and geographic location. The NCDB also offers five-year survival reports stratified by AJCC staging.

In early 2007, the National Quality Forum (NQF) endorsed five quality measures developed in a joint effort between the ACoS, the National Comprehensive Cancer Network (NCCN), and the American Society of Clinical Oncology (ASCO). Four of these measures focus on adjuvant therapy for breast and colorectal cancer, and one measure concerns the resection and pathological examination of 12 or more regional lymph nodes for colon cancer. Hospitals with CoC-approved cancer programs reporting to the NCDB are evaluated and benchmarked according to their performance on these measures. Hospitals can examine their performance on each measure in comparison to the other 1,430 CoC hospitals, allowing only the reporting hospital to see where they rank in comparison to the rest of the state, region, or country (Fig. 2). Moreover, this data feedback mechanism allows hospitals to evaluate deficient records, to ensure the accuracy and completeness of their data, and make corrections and additions as needed.

These quality-of-care reports are intended to provide hospitals with the ability to measure and compare standards of cancer care at their facility to that of a large number of hospitals across the country, without any associated accountability. Over the next year, the NCDB will be developing and incorporating additional quality measures for multiple cancers, including common (i.e., lung) and complex (i.e., pancreas) malignancies.

FUTURE DIRECTIONS

The power of the NCDB clearly rests in its ability to receive data, analyze it, and send actionable information back to hospitals for the purposes of quality improvement. These quality improvement activities will play a significant role in the future for the NCDB. In addition to the breast and colorectal initiatives, quality measures must be developed for other cancer sites. These measures should not be used for accountability, but rather only to provide information to hospitals with the aim of improving patient care and outcomes. In the future, federal regulatory agencies, payers, purchasers, and consumers may look to data from cancer registries to assess surgical quality, determine reimbursement levels, and judge providers on the basis of process measures and patient outcomes.

The NCDB will soon have a de-identified dataset available to investigators at CoC-approved hospitals.

Ann. Surg. Oncol. Vol. 15, No. 3, 2008
This file will provide the opportunity to examine a large dataset containing a number of data items that are not available through other cancer registries including surgical margin status, systemic therapy (chemotherapy and hormonal therapy), detailed radiation treatment data, and clinical and pathological AJCC staging.

CONCLUSIONS

The NCDB is a large, powerful database providing multiple opportunities for clinical studies and quality-improvement endeavors. Importantly, the NCDB can be used to benchmark hospitals on performance measures and serve as an impetus for quality improvement initiatives at the hospital level. Cancer registration activities are expensive for hospitals, but the benefits are clearly demonstrated by the NCDB's extensive efforts to feed important information back to participating institutions. The NCDB is the only cancer registry that provides feedback data to hospitals in such a manner. It is the responsibility of surgeons, and their colleagues in other oncologic specialties, to work collectively, and with their hospital's cancer registrars, to continuously review and improve the accuracy, completeness and quality of the hospital's cancer registry data, and thus the quality of the data reported to the NCDB. The NCDB is a valuable tool to improve the quality of cancer care in the United States.

ACKNOWLEDGEMENTS

The authors would like to recognize the staff of the National Cancer Data Base: James M. Banasiak, E. Greer Gay, Kathy Mallin, Bryan Palis, Florin Petrescu, Jerri Linn Phillips, and Kimberly R. Porter.

REFERENCES

1. National Cancer Data Base. http://www.facs.org/cancer/ncdb/index.html. Cited 15 May 2007.
2. Commission on Cancer: Approvals Categories. http://www.facs.org/cancer/coc/categories.html. Cited 16 July 2007.
3. Winchester DP, Stewart AK, Bura C, et al. The National Cancer Data Base: A clinical surveillance and quality improvement tool. J Surg Oncol 2004; 85:1–3.
4. Commission on Cancer.http://www.facs.org/cancer. Cited 15 July 2007.
5. Commission on Cancer. Cancer Program Approval. http://www.facs.org/cancer/whatis.html. Cited 1 July 2007.
6. Cancer Registries Amendment Act, Public Law 102–515. Congressional Record 1992; 138, 106 STAT 3372–7.
7. Facility Oncology Registry Data Standards. Chicago: Commission on Cancer; 2004.
8. Jemal A, Murray T, Ward E, et al. Cancer statistics, 2005. CA Cancer J Clin 2005; 55:10–30.
9. National Cancer Institute: SEER Surveillance Epidemiology and End Results. http://seer.cancer.gov/. Cited 26 January 2007.
10. National Program of Cancer Registries, Cancer Surveillance System. http://www.cdc.gov/cancer/nprc/training/cks.htm. Cited 11 August 2007.
11. National Cancer Registrars Association: Certified Tumor Registrar certification. http://ncra-usa.org/certification/index.htm. Cited 12 July 2007.
12. Centers of Disease Control and Prevention: National Program of Cancer Registries, EDITs. http://www.cdc.gov/cancer/nprc/tools/edit/. Cited 1 July 2007.
13. A. K. Stewart, E. G. Gay, L. Patel-Parekh, et al. Provider feedback improves reporting on quality measures: National
profile reports for adjuvant chemotherapy for stage III colon cancer. Presented at the American Society of Clinical Oncology Annual Meeting, Chicago, 2007.

14. NCDB Public Benchmark Reports. http://www.facs.org/ncdbbenchmarks8.cfm. Cited 1 July 2007.

15. Sylvester M, Blankenship C, Carter A, et al. Quality control: The American College of Surgeons Commission on Cancer Standards, National Cancer Data Base, and Cancer Liaison Program. *Reg Mgmt* 2001; 28:68–74.

16. Eberle C, Phillips J, Tary P, et al. Quality management in the National Cancer Data Base: A re-abstraction study of the Midwest region. *J Reg Mgmt* 1997; 24:93–7.

17. Fremgen A, Jessup J, Menck HR. Hospital use of NCDB data to assess quality of cancer patient care. *J Reg Mgmt* 1995; 22:69–71.

18. Mulinar K, Phillips J, Fritz A, et al. Quality of oncology data: Findings from the Commission on Cancer PCE study. *J Reg Mgmt* 2001; 28:24–34.

19. Standards for Cancer Registries. Data Standards and Data Dictionary, 8th Ed., Version 10.1, Vol. II, Springfield, IL: North American Association of Central Cancer Registries; 2003.

20. International Classification of Disease for Oncology, 3rd ed. Geneva: World Health Organization; 2000.

21. AJCC Cancer Staging Manual. 6th ed. Chicago, IL: Springer; 2002.

22. Health Insurance Portability and Accountability Act (HIPAA). 1996. Standards for Privacy of Individually Identifiable Health Information; Final Rule (45 CFR Parts 160 and 164).

23. National Cancer Data Base: Bibliography. http://www.facs.org/cancer/ncdb/biblclin.html. Cited 12 July 2007.

24. Bilimoria KY, Tomlinson JS, Merkow RP, et al. Melanoma in children and teenagers: an analysis of patients from the National Cancer Data Base. *J Clin Oncol* 2007; 25:1363–8.

25. Bilimoria KY, Bentrem DJ, Ko CY, et al. National failure to operate on early stage pancreatic cancer. *Ann Surg Oncol* 2007; 24:173–80.

26. Bilimoria KY, Bentrem DJ, Ko CY, et al. Validation of the 6th edition AJCC Pancreatic Cancer Staging System: report from the National Cancer Database. *Cancer* 2007; 110:738–44.
56. Bilimoria KY, Stewart A, Tomlinson JS, et al. Impact of adjuvant radiation on survival: A note of caution when using cancer registry data to evaluate adjuvant treatments. *Ann Surg Oncol* 2007; in press.
57. Miller DC, Spencer BA, Ritchey J, et al. Treatment choice and quality of care for men with localized prostate cancer. *Med Care* 2007; 45:401–9.
58. Warren JL, Harlan LC. Can cancer registry data be used to study cancer treatment? *Med Care* 2003; 41(9):1003–5.
59. Bilimoria KY, Bentrem DJ, Ko CY, et al. Extent of surgery affects survival for papillary thyroid cancer. *Ann Surg* 2007; 246:375–81; discussion 381–4.
60. NCDB Public Survival Reports. http://www.facs.org/ncdbbr/survival8.cfm. Cited 1 July 2007.
61. Commission on Cancer (CoC). http://www.facs.org/cancer/qualitymeasures.html. Cited 26 April 2007.
62. Quality of Cancer Care Performance Measures. National Quality Forum website. http://216.122.138.39/ongoing/cancer/index.asp. Cited 17 January 2007.
63. American Cancer Society, Cancer Facts and Figures. http://www.cancer.org/docroot/STT/content/STT_1x_Cancer_Facts__Figures_2007.asp. Cited 21 January 2007.