

The National Cancer Data Base: Past, Present, and Future

David P. Winchester, MD¹, Andrew K. Stewart, MA², Jerri Linn Phillips, MA², and Elizabeth E. Ward, PhD³

¹Cancer Programs, American College of Surgeons, Chicago, IL; ²National Cancer Data Base, American College of Surgeons, Chicago, IL; ³Surveillance and Health Policy Research, American Cancer Society, Atlanta, GA

The National Cancer Data Base (NCDB) contains information on over 25 million cancer patients diagnosed and treated in cancer centers across the USA since 1985. The NCDB collects data on patient demographics, tumor stage and histopathology, treatment, and outcomes on more than 70% of the cancer cases diagnosed in the USA annually. Reporting centers range from small community hospitals to large academic medical centers and National Cancer Institute (NCI)-designated Comprehensive Cancer Centers. Since its inception in 1988 the NCDB has been cofunded by the American College of Surgeons (ACoS) and the American Cancer Society (ACS). The programmatic focus of the NCDB, since its founding, has been to support quality improvement at the local level.

The NCDB has evolved over time. Its focus has shifted from peer-reviewed observational studies on patterns of cancer care to the development of web-based audit and feedback reporting tools for Commission on Cancer (CoC)-accredited cancer programs to promote local quality assessment and improvement initiatives. These tools include descriptive reports that permit extensive user customization and report-card-style reports displaying performance rates for nationally recognized evidence-based quality-of-care measures endorsed by the National Quality Forum (NQF). The NCDB remains a widely recognized and valuable resource for a broad range of investigators. Although not population based, the NCDB reflects treatment patterns in a defined universe of CoC-accredited facilities that are required to maintain high-quality hospital cancer registries as well as meet other process standards. A number of retrospective studies have

been conducted in recent years, each spearheaded by multidisciplinary disease site teams and supported by NCDB analytic staff. The Health Services Research Group at the American Cancer Society has also used the data base to analyze disparities in care and outcomes related to insurance status of cancer patients, resulting in a number of publications and presentations. As a consequence, requests to initiate studies using data from the NCDB have far exceeded the available capacity of staff resources to support proposed projects. Anticipating a growth in the demand for access to the data base, the NCDB is working to develop a participant use file (PUF). The PUF will be a Health Insurance Portability and Accountability Act (HIPAA)-compliant de-identified data set available to interested investigators at CoC-accredited programs who have the local resources, such as statistical analysts, to conduct studies.

A review of the NCDB's strengths, weaknesses, and opportunities has been conducted and a framework for the future directions of the data base has been developed. To this end, a peer review of the NCDB was conducted in March 2008, followed by a stakeholder summit in January 2009. Recommendations from the NCDB March 2008 peer review centered on three critical elements of the NCDB's current and future activities: (1) maintenance and promotion of hospital participation in the CoC accreditation program, (2) continued emphasis on quality-improvement initiatives, and (3) validation of the accuracy and completeness of the data reported to the NCDB. The goal of the summit was to review and discuss the recommendations from the 2008 NCDB peer review, determine short-, medium-, and long-term objectives of NCDB stemming from a staff-developed project plan, and prioritize and develop strategies for these objectives. Summit attendees included senior clinical staff from the ACoS and ACS, NCDB staff, ACS Health Services Research staff, Commission on Cancer leadership representing each of the CoC

standing committees, and leadership figures from other constituent organizations such as the American Joint Committee on Cancer (AJCC), the American College of Surgeons Oncology Group (ACOSOG), and the NCI.

The peer-review report encouraged efforts to retain and expand participation in the CoC accreditation program, and emphasized that it was important for participating programs to perceive that they receive a high-value return from their participation. The NCDB was viewed as an integral part of the CoC, and furthermore, the development and implementation of new quality-improvement measures and tools through the NCDB for these cancer programs held great potential value. A review of the accreditation standards for cancer programs is underway, and revisions will focus on strategies to reduce burden of participation while increasing the value of participation. The CoC plans to refocus new standards around performance metrics and new mechanisms to help hospitals identify appropriate metrics to evaluate the care provided for its patients and to support the implementation of these quality assessment and monitoring activities.

Many hospital and cancer program administrators use data from the NCDB to monitor quality and compare their performance and outcomes with those of other providers. The extent to which NCDB can feed back aggregated data to hospitals and offer added value in the form of reporting tools and online access to data will help ensure hospitals' continued involvement with the data base. Expanding the information available to local institutions for comparative analyses increases the value of their investment in tumor registries. The CoC's Cancer Program Practice Profile Reports (CP³R) provide hospital performance rates for the five NQF evidence-based quality of cancer care measures for breast and colon cancer. In addition, this reporting tool also includes a rectal cancer measure jointly developed and specified by the CoC, the American Society of Clinical Oncology (ASCO), and the National Comprehensive Cancer Network (NCCN). These reports were recently incorporated into the accreditation review process, and will likely be expanded as new consensus measures are considered, developed, and specified. Soliciting clinical, methodological, and statistical expertise through panels of disease site teams will be critical as meaningful and feasible measures are identified and feedback reports are made available to CoC-accredited cancer programs

A new initiative of the NCDB is the Rapid Quality Reporting System (RQRS). The RQRS represents a significant step toward providing timely and high-value information to CoC-accredited programs. The RQRS permits close to "clinical real-time reporting," and issues alerts well within a timeframe that will allow decisions on anticipated, evidence-based care to be discussed, ordered, and provided efficiently and within guideline recommendations.

Users of the system are provided with rolling year-to-date assessment reports, daily updated online alerts, and timely comparison performance reports using the NQF cancer care measures. While the retrospective CP³R reporting model has demonstrated that linking assessment of clinical practice to metrics utilizing cancer registry data can result in noticeable shifts in the completeness of adjuvant therapy data in the registries, the RQRS brings to bear an assessment and monitoring tool closer to the time of the clinical encounter, and it is anticipated will drive highly reliable and timely data on ambulatory care into participating hospital registry data sets and subsequently into the NCDB.

The peer review encouraged the NCDB to engage in methodologically sound and well-conceived data validation studies. Concerted efforts to ascertain the representativeness of the data base were strongly encouraged. In addition, the panel urged that the NCDB undertake a validation of those items reported to the NCDB that are not collected by state or regional registries (e.g., insurance status, secondary diagnoses used to gauge comorbid disease status, and disease recurrence) that might otherwise be subject to routine quality control by central registries.

Currently, the NCDB manages large-scale data quality control through use of the EDITS software package developed under the auspices of the Centers for Disease Control and Prevention and widely used by hospital, state, and regional registries for over a decade. Using this software, the NCDB conducts extensive internal logic checking of reported cases using rules developed by the North American Association of Central Cancer Registries (NAACCR), in collaboration with the NCI/Surveillance, Epidemiology, and End Results (SEER) and Centers for Disease Control and Prevention (CDC)/National Program of Cancer Registries (NPCR) programs. Additionally, the CoC accreditation standards mandate clinical review of 10% of registry case abstracts. However, the results of these reviews are self-reported and not independently verified. Starting in 2009, the CoC has strengthened its chart review requirements as part of the accreditation site visit process. This includes a review of hospital charts and registry abstracts to verify that registry data correctly reflect the information documented in individual patient records, and that the information summarizing the patient's medical condition, care, and participation in treatment decision-making processes are adequately documented in hospital records. These chart reviews by CoC site surveyors use the CoC CP³R reported performance rates for the four NQF-endorsed accountability measures for breast and colon cancer care. These systematic reviews are intended to (1) ensure that reported performance rates are an accurate reflection of the care provided to patients at CoC-accredited programs, and (2) address the concern that hospital-based cancer registry data may not accurately

assess outpatient therapy data. By design, the CoC chart review process focuses on the completeness and accuracy of the reported ambulatory care. As the CoC reviews its program standards for 2011, it is certain that an increased focus on quality improvement will significantly influence the nature of surveys and that surveys will likely emphasize validation of data reported to the NCDB rather than a review of facility operational and structural characteristics.

Previous evaluations of the completeness and accuracy of cancer registry data in the literature adopted comparative chart review methodologies, and noted significant differences between registry-reported and provider-documented treatment information. There is early and increasing evidence that methodologies to access commercial insurance or Medicare claims data may yield accurate and efficient ways to obtain both in-patient and ambulatory therapy data. Supplementing registry data with information available from claims provides opportunities both to validate the sensitivity and specificity of registry data and to elicit greater levels of granularity with regard to treatment provided to cancer patients. Current efforts to link registry data with administrative claims suggest that, while data on ambulatory treatment is still underreported in cancer registries, it is more complete in contemporary data than in earlier studies. Results from a small number of studies may be expected in the near future.

Collaboration with other organizations actively engaged in monitoring the clinical management of cancer patients is an important key to supplementing and validating registry data. The American College of Surgeons Oncology Group (ACOSOG) now has the ability to compare data collected prospectively through National Cancer Institute (NCI) trials with data collected previously by cancer center registries. Another possibility is an emerging pilot project between the NCDB and the American Society of Clinical Oncology's (ASCO) Quality Oncology Practice Initiative (QOPI) to explore opportunities to access information describing outpatient delivery of systemic therapy.

The cancer program standards for CoC-accredited programs may also influence the accuracy, completeness, and timeliness of data reported to the NCDB. Some investigators have commented on the higher level of completeness and accuracy of data when comparing CoC-accredited cancer registries with registries located in non-accredited centers. Evidence from the CoC CP³R audit and feedback reporting application, described above, indicate that standards for abstracting timeliness may result in truncating follow-up for some treatment information by cancer registries. Additionally, although accreditation standards require a 90% follow-up rate over the most recent 5 years of abstracted data, recent NCDB analyses have found that the cumulative lost-to-follow-up rate for some patient cohorts approaches 25%. Such rates may limit

the types of analyses and conclusions that can be drawn from these data. Vital status follow-up from hospital registries has not been systematically reviewed or validated, and it appears that methods to determine vital status vary widely between registries. Many hospital registries do not have access to sources of vital status follow-up available to central registries, including linkage with state vital records and the National Death Index (NDI). A respecification of the follow-up activities for CoC-accredited programs is likely with the pending review of the 2011 accreditation standards. Additionally, collaborations with state and regional registries supported by federal agencies to begin to develop data-sharing policies and procedures to populate hospital registries with vital status follow-up data from administrative sources, as well as linkage between patient identifiers and the NDI to ascertain vital status and causes of death, will be explored.

The value of the NCDB as a national clinical surveillance tool for cancer has increased significantly since its inception 20 years ago. Clinical analysis using the NCDB to comment broadly on the state and variability of care provided to cancer patients in the USA must be facilitated and fostered. At publication, NCDB is in the closing stages of a project to develop a participant use file (PUF) that will contain case-level data for use by qualified investigators at CoC-accredited cancer programs. The PUF will contain a full complement of data items necessary to conduct a broad range of studies. Necessary measures will be taken to de-identify both reporting facilities and cases records in order to be compliant with HIPAA regulations and the long-standing business operations of the NCDB. Investigators will be expected to have access to sufficient statistical and technical expertise to conduct their own studies.

The CoC recognizes that a balance must be struck between maintaining and promoting as high a caliber of clinical studies as possible, while at the same disseminating the PUF to as broad a community of users as possible. The NCDB PUF can be expected to evolve as experience and expertise dictate. An incremental release strategy for the PUF will be necessary, as the CoC assesses procedures and policies developed to facilitate data distribution. During this time, the usefulness and potential limitations of the PUF dataset for clinical research, as well as the level of documentation and staff support required to support external investigators, will be evaluated with alpha test investigators. It is anticipated that a limited number of investigators from different institutions will be offered the opportunity to participate in this alpha test, facilitating interinstitutional dialogue and collaboration. An important objective of the alpha test is to foster communities of expertise and establish a core NCDB user community that can serve as an effective reference point for future users of the NCDB PUF.

While the scope and depth of the analytic and reporting activities of the NCDB has evolved since its inception, some of the potential for the NCDB to provide timely feedback on quality of care and to provide a national infrastructure for surveillance of patterns of care in the USA has not been fully realized. Several initiatives, including the RQRS, currently underway will accelerate progress towards these objectives. Increasing emphasis on evaluating the representativeness of patients and treatment patterns captured in the NCDB compared with all patients, and more systematically validating collected data, will likely lead to increasing acceptance of NCDB data in the cancer surveillance and health services research communities. The

size and sophistication of the NCDB infrastructure allows data to be collected, aggregated, and used to generate a wide range of reports and peer-reviewed manuscripts. Leveraging its unique relationship with providers through the CoC accreditation program, the data base will continue to establish itself as the primary source for developing and implementing quality metrics for cancer care improvement, and to retain and expand broad support within the clinical community. The NCDB is clearly recognized as a valuable resource and one that continues to position itself to maximize its value for both its contributors in CoC-accredited cancer programs and members of the clinical, research, and policy-making communities.