National Cancer Database NCDB

The National Cancer Database is an offshoot of an organization started in 1989 by the American College of Surgeons (ACoS) and the Commission on Cancer (CoC) that acts as 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 goal of the ACDB is to improve cancer care through setting standards, prevention, research, education, and the monitoring of comprehensive cancer care. The NCDB is housed under the ACoS.

To gain access to the NCDB, the requesting entity must be a hospital from a CoC accredited program. Once that initial criteria is established, the entity must agree to the data use agreement and provide a letter of support from CoC accredited program, all in addition to a completed application. When the ACDB was first established, any hospital could voluntarily report data to the NCDB, however, 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 hospitals who earned CoC approval.

To analyze the data contained within the NCDB, the user needs to use statistical analysis software due to the database only containing the data warehouse and lacks tools. The NCDB suggests using SAS, SPSS, Strata, SPLUS, or any other statistical analysis tool. The data within the NCDB is contained in PUFs (Participant User Files), of which are HIPPA compliant.

Due to the stringent access approval process, only CoC approved hospitals are permitted to access the data. The data is walled and only exists in detail with the correct credentials. The full dataset contains data on 21 million cancer patients diagnosed between 1985 and 2005, the data comes from 1,430 hospitals. However, some of the data is publicly accessible and includes some benchmark reports and includes: 14 most commonly diagnosed solid tumors in the United States. Public users are provided access to data from ten diagnosis years (2008-2017) and contains slightly more than 11 million cases.

The data contained within the NCDB is mostly raw, descriptive data. The PUF’s also contain quality-of-care-reports. However, the data is in the database for analysis by end-users and among the analyses done by end-users: Examined surgical and adjuvant therapy treatment trends, studies of rare cancers and subset analyses. Studies using NCDB data have addressed important clinical questions in areas where clinical trials may be challenging.

The National Cancer Database is a nationwide oncology outcomes database for around 1,500 commission-accredited cancer programs in the United States and Puerto Rico. The data is inaccessible for common users, but has particularly high data standards and is overseen by medical boards.

Surveillance, Epidemiology, and End Results (SEERS)

The Surveillance, Epidemiology, and End Results program of the National Cancer Institute is a source of epidemiologic information on the incidence and survival rates of cancer in the United States. The SEER program collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 34.6% of the population of the United States. The SEER program is the only comprehensive source of population-based information in the United States that includes stage of cancer at the time of diagnosis and patient survival data. SEER began collecting data in 1973 in a limited number of states and cities within the United States. The National Cancer Institute funds for the program come from the Centers for Disease Control and Prevention through the National Program of Cancer Registries and gets additional funding from participating states. SEER is supported by the Surveillance Research Program in NCI’s division of Cancer Control and Population Services.

To gain approval to access the database institutional users can complete a registration form, provide a signing official and then can acquire a SEER\*Stat username. Once this process is complete, the user must use the SEER\*Stat statistical software.

The data is somewhat publicly accessible and can be accessed from the SEER website using a web browser. The public can access and visualize the data through the SEER\*Explorer application, which is web based and connects to the SEER database, although this web interface does not allow public users access to the raw data. The publicly accessible and available as a public service in print and electronic formats.

The SEER program registries routinely collect data on patient demographics, primary tumor site, tumor morphology, and stage at diagnosis, first course of treatment, and follow-up to vital stats. The population data used in calculating cancer rates is obtained periodically from the Census Bureau. Quality control has been an integral part of SEER since its inception. Every year, studies are conducted in SEER areas to evaluate the quality and completeness of the data being reported.

SEER is an authoritative source for cancer statistics in the United States and provides information on cancer statistics in an effort to reduce the cancer burden among the U.S population.