North American Association of Central
Cancer Registries, Inc.

Standards for Cancer Registries
Volume III

Standards for Completeness, Quality, Analysis, and Management of Data

Edited by
Lori A. Havener

October 2004

SPONSORING ORGANIZATIONS

American Cancer Society
American College of Surgeons
American Joint Committee on Cancer
Canadian Association of Provincial Cancer Agencies
Centers for Disease Control and Prevention
Health Canada
National Cancer Institute
National Cancer Registrars Association
Statistics Canada

Edited by:
Lori A. Havener, CTR
Program Manager of Standards
NAACCR, Inc.

Comments and suggestions on this and other NAACCR standards documents are welcome and can be forwarded to the editor or to any member of the NAACCR Board of Directors.

The other volumes in the series, Standards for Cancer Registries, are:

- *Volume I: Data Exchange Standards and Record Description*. Intended for programmers, this provides the record layout and specifications for the standard for data exchange, including correction and analysis formats.

- *Volume II: Data Standards and Data Dictionary*. Intended for central registries, this provides detailed specifications and codes for each data item in the data exchange record layout.

Copies of all standards documents can be viewed or downloaded from NAACCR's website at www.naaccr.org.

Suggested Citation:
Havener L (Ed). Standards for Cancer Registries Volume III: Standards for Completeness, Quality, Analysis, and Management of Data. Springfield (IL): North American Association of Central Cancer Registries, October 2004, 180 pp.

Acknowledgment:
Thank you to Susan Sullivan for administrative support and to the NAACCR Standards Volume III Workgroup of the Registry Operations Committee for their dedication and long hours of review to prepare this document.

Support for editorial services of this publication was funded in part with Federal funds from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN261200444001C and ADB No. N02-PC-44401. Production and distribution of this publication was supported in part by Grant/Cooperative Agreement Number U75/CCU523346-02 from the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

Table of Contents

NAACCR Board of Directors
2004-2005

NAACCR Registry Operations Committee 2004-2005

Chairs:
Susan T. Gershman, MS, MPH, PhD, CTR
Massachusetts Cancer Registry
Massachusetts Department of Public Health
Telephone: (617) 624-5646
E-mail: susan.gershman@state.ma.us

Susan Bolick-Aldrich, MSPH, CTR
South Carolina Central Cancer Registry
South Carolina Department of Health and Environmental Control
Telephone: (803) 898-3626
E-mail: bolicks@dhec.sc.gov

NAACCR Staff:
Jan Conway
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: jconway@naaccr.org

Ilona Kay Gebhard, BA, CTR
NAACCR, Inc.
Telephone: (505) 332-0470
E-mail: kgebhard@naaccr.org

Lori A. Havener, CTR
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: lhavener@naaccr.org

Committee Members:
Judy Boone
Los Angeles Cancer Surveillance Program
Telephone: (323) 442-2330
E-mail: jboone@hsc.usc.edu

Maria O. Celaya
New Hampshire State Cancer Registry
Telephone: (603) 653-1032
E-mail: maria.o.celaya@dartmouth.edu

Gayle Greer Clutter, CTR
Centers for Disease Control and Prevention
Telephone: (770) 488-4570
E-mail: gcc6@cdc.gov

Rose M. Garner, CTR
IMPAC Medical Systems, Inc.
Telephone: (301) 432-0425
E-mail: rgarner@impac.com

Carol Hahn-Johnson, BS, CTR
Cancer Statistics Branch
Division of Cancer Control and Population Sciences
Surveillance Epidemiology and End Results Program
National Cancer Institute
Telephone: (301) 402-6226
E-mail: johnsoca@mail.nih.gov

Joanne Harris, CTR
Cancer Surveillance
Karmanos Cancer Institute
Telephone: (313) 833-0715, Ext. 2345
E-mail: harrisj@med.wayne.edu

Stuart Herna
Florida Cancer Data System
University of Miami School of Medicine
Telephone: (305) 243-4600
E-mail: s.herna@miami.edu

Jennifer McCoy, MPA, CTR
Oregon State Cancer Registry
Telephone: (503) 731-4858
E-mail: jennifer.m.mcoy@state.or.us

Stacey Neloms, MPH
Maryland Cancer Registry
Maryland Department of Health and Mental Hygiene
Telephone: (410) 767-5521
E-mail: nelomss@dhmh.state.md.us

Lilia O'Connor
California Department of Health Services
Cancer Surveillance Section
Telephone: (916) 779-0355
E-mail: loconnor@ccr.ca.gov

Robin D. Otto, RRA, CTR
Pennsylvania Department of Health
Telephone: (717) 783-2548
E-mail: rootto@health.state.pa.us

Reita Pardee
Kentucky Cancer Registry
Telephone: (859) 219-0773
E-mail: rpardee@kcr.uky.edu

Lisa Pareti, BS, RHIT, CTR
Louisiana Tumor Registry
Telephone: (504) 568-3036
E-mail: lparet@lsuhsc.edu

Lina Patel-Parekh, MHA
Commission on Cancer
American College of Surgeons
Telephone: (312) 202-5401
E-mail: lpatel-parekh@facs.org

Randi Rycroft, BA, BS, CTR
Colorado Central Cancer Registry
Telephone: (303) 692-2540
E-mail: randi.rycroft@state.co.us

Bette R. Smith, RHIT, CTR
Ohio Cancer Incidence Surveillance System
Bureau of Health
Telephone: (614) 644-7058
E-mail: bsmith@gw.odh.state.oh.us

Kathleen Thoburn
New York State Cancer Registry
Telephone: (518) 474-2255
E-mail: kkt01@health.state.ny.us

Kim Vriends
Prince Edward Island Cancer Registry
Telephone: (902) 894-2167
E-mail: kavriends@ihis.org

Brad Wohler
Florida Cancer Data System
University of Miami School of Medicine
Telephone: (305) 243-5527
E-mail: bwohler@med.miami.edu

NAACCR Volume III Workgroup
2004-2005

Chair:
Lori A. Havener, CTR
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: lhavener@naaccr.org

Workgroup Members:
Ingrid Friesen
Statistics Canada
Telephone: (306) 652-3876
E-mail: ifriesen@sasktel.net

Ilona Kay Gebhard, BA, CTR
NAACCR, Inc.
Telephone: (505) 332-0470
E-mail: kgebhard@naaccr.org

Susan T. Gershman, MS, MPH, PhD, CTR
Massachusetts Cancer Registry
Massachusetts Department of Public Health
Telephone: (617) 624-5646
E-mail: susan.gershman@state.ma.us

Carol Hahn-Johnson, BS, CTR
Cancer Statistics Branch
Division of Cancer Control and Population Sciences
Surveillance Epidemiology and End Results Program
National Cancer Institute
Telephone: (301) 402-6226
E-mail: johnsoca@nih.gov

Ryan Intlekofer, RN, CTR
Centers for Disease Control and Prevention
Telephone: (770) 488-1075
E-mail: rbi1@cdc.gov

Lilia O'Connor, MBA, RHIT, CTR
California Dept. of Health Services
Cancer Surveillance Section
Telephone: (916) 779-0355
E-mail: loconnor@ccr.ca.gov

Lina Patel-Parekh, MHA
American College of Surgeons
Commission on Cancer
Telephone: (312) 202-5401
E-mail: lpatel-parekh@facs.org

Bette R. Smith, RHIT, CTR
Ohio Cancer Incidence Surveillance System
Bureau of Health
Telephone: (614) 644-7058
E-mail: bsmith@gw.odh.state.oh.us

Preface

One of the primary goals of the NAACCR Registry Operations Committee (ROC) has been to review, update, and revise *Standards for Cancer Registries Volume III: Standards for Completeness, Quality, Analysis, and Management of Data* of the NAACCR standards documents. The *Procedure Guidelines for Cancer Registries* being developed by the ROC focuses on individual operational activities at the central registry level. The intent is to supplement Volume III by providing detailed guidelines for specific operations activities.

The revisions in the 2004 Edition focus on:

- Updating references and tables to refer to ICD-O-3 and ICD-10;
- Revisions to reflect updates to all chapters;
- Reporting of benign brain tumors;
- Table 2: Actual Percent Unknown for Selected Data Items;
- Table 4: Standard Site Analysis Categories with ICD-O-3 Codes;
- Table 6: Standard Site Analysis Categories for Mortality Data (ICD-10);
- Table 7: SGC Codes for Canadian Provinces and Territories;
- Table 8: Standard Treatment Analysis Categories; and
- Updated reference information to include the most recent materials.

The Committee is hopeful that the revised document will more accurately reflect activities and resources within the central registry population.

Susan Bolick-Aldrich, MSPH, CTR
Co-Chair, Registry Operations Committee

Susan T. Gershman, MS, MPH, PhD, CTR
Co-Chair, Registry Operations Committee

Chapter 1: Introduction

The mission of the North American Association of Central Cancer Registries, Inc. (NAACCR) is to promote uniform data standards for cancer registration; provide education and training; certify population-based registries; aggregate and publish data from central cancer registries; and promote the use of cancer surveillance data and systems for cancer control and epidemiologic research. The NAACCR Standards for Cancer Registries volumes were prepared to develop and promote uniform data standards. These publications compile consensus standards among the North American cancer registry community as represented by the NAACCR membership. The purpose of these standards is to increase the quality, comparability, and utility of cancer incidence data in North America.

The NAACCR membership is comprised of central registries throughout the United States and Canada, national organizations, and individuals collaborating to reduce the burden of cancer in North America (see Appendix A). Central cancer registries in North America are a diverse group and have been established at different times and for different purposes. Some are intended to provide only basic descriptive epidemiological data; others provide a base for epidemiological and biomolecular research. Some registries emphasize cancer control and patient management, and others focus on end results and survival.

Establishment of standards is of major importance in enhancing the usefulness of central cancer registry data. Collaborative studies and data comparisons are feasible as data become more directly comparable. NAACCR promotes activities pertinent to effective and efficient cancer registry operations. These activities include, but are not limited to, the Procedure Guidelines for Cancer Registries (Series I-V), training programs and education CDs, and *ad hoc* workgroup reports (e.g., *A Review of the Definition for Multiple Primary Cancers in the United States).* Additional information may be found at www.naaccr.org.

No single set of standards can address all points of diversity in local needs or take all local idiosyncrasies into account. These standards were formulated based on the following principles:

- ***Model:*** The model central cancer registry addressed by these standards collects complete population-based data for a defined geographic area, including treatment and stage data, and may or may not collect patient follow-up. It collects information from hospitals, other health care facilities, and physicians.

- ***Strictness:*** The standards presented in this document vary in how strongly they are recommended. Below are the three levels of application:
  - **MUST:** Experience has shown that certain central registry characteristics are necessary for effective and efficient operation of a cancer registry. These are identified by **MUST** in the standards. Although there may be registries that function without these characteristics, it is the present consensus that any new registry should adopt these standards.
  - **SHOULD:** Experience has shown that other characteristics are strongly recommended, but not absolutely required as the **MUST** above. These are designated by **SHOULD** in the standards. Some of the problems addressed by **SHOULD** can be solved in alternate ways depending on local conditions, needs, and resources.
  - **MAY:** Other characteristics that are highly desirable, but not necessary, are designated as **MAYs.**

Detailed discussions of methods have been omitted from this document when they are available elsewhere. However, a reference is provided for the source of this information.

Chapter 2: Access to Source Data and Completeness of Reporting

2.1. STRUCTURAL REQUIREMENTS

2.1.1. Legislation and Regulations

Authority for a population-based cancer registry to collect data on cancer incidence is established through: (1) legislation for cancer reporting with or without regulations; and/or (2) regulations/rules developed under general authorization for the reporting of diseases, as specified by state or provincial health authorities. *Legislation or statute* refers to a form of law enacted by a state, provincial/territorial legislature, Congress, or Parliament. *Regulation or rule* refers to a form of law created by administrative agencies of a government.

Legislative authority **SHOULD** include specific components that relate to central registry development and function, as well as specific directives for the publication of regulations detailing these components. Often, authority is granted to the jurisdiction's health department, which, in turn, may delegate authority to another agency. In other instances, authority is granted directly to another agency, such as a university or foundation.

The purpose of this section is to provide guidance to departments of health or other agencies seeking to develop, evaluate, or improve both cancer registry legislation and regulations in their state or province/ territory.

Comprehensive central cancer registry legislation and regulations cover a number of issues, including: reporting requirements, patient record access, enforceability, data quality and data standards, confidentiality and disclosure of data, liability, and specification of funding source. Section 2.1.1.1 provides a further explanation of these issues, and Appendix B provides an example of reporting legislation.

Through NAACCR, central cancer registries have worked toward improving data quality and increasing comparability across geographic areas. Reducing variability in cancer reporting by state and province/territory is part of the NAACCR agenda.

In Canada, provincial and territorial cancer registries have joined with Statistics Canada, a federal agency, to form the Council of Canadian Cancer Registries (CCCR), which supervises the operation of the national level Canadian Cancer Registry (CCR) dataset (reference year 1992) and provides guidelines and advice for provincial/territorial central registries. The agreement establishing the Council permits all parties to put in place operational arrangements for quality assessment and control. Overviews of the development of cancer registration in Canada and the patterns of cancer occurrence can be found in *The Making of the Canadian Cancer Registry: Cancer Incidence in Canada and its Regions 1969-1988.*

The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute is an authoritative source of information on cancer incidence and survival in the United States. SEER began collecting data on cases on January 1, 1973, in the states of Connecticut, Iowa, New Mexico, Utah, and Hawaii, and the metropolitan areas of Detroit and San Francisco-Oakland. In 1974-1975, the metropolitan area of Atlanta and the 13-county Seattle-Puget Sound area were added. In 1978, 10 predominantly black rural counties in Georgia were added, followed in 1980 by the addition of American Indians residing in Arizona. Three additional geographic areas participated in the SEER Program prior to 1990: New Orleans, Louisiana (1974-1977, rejoined 2001); New Jersey (1979-1989, rejoined 2001); and Puerto Rico (1973-1989). The National Cancer Institute also funds a cancer registry that, with technical assistance from SEER, collects information on cancer cases among Alaska Native populations residing in Alaska. In 1992, the SEER Program was expanded to increase coverage of minority populations, especially Hispanics, by adding Los

Angeles County and four counties in the San Jose-Monterey area south of San Francisco. In 2001, the SEER Program expanded coverage to include Kentucky and Greater California; in addition, New Jersey and Louisiana once again became participants.

In 1992, the U.S. Congress passed the Cancer Registries Amendment Act (PL 102-515) for the purpose of establishing "a national program of cancer registries (NPCR)," through a system of cooperative agreements with states, territories, and the District of Columbia to support the operation of population-based statewide cancer registries (see Appendix C; see Appendix D for the Benign Brain Tumor Cancer Registries Amendment Act). Prior to funding, the national legislation requires assurances from states that the state will "provide for the authorization under State law of the statewide cancer registry, including publication of regulations." The national legislation states that reporting requirements, patient record access, data quality and standards, confidentiality and disclosure of data, and liability all are areas that must be addressed through state legislation and regulations. NPCR is administered through the Centers for Disease Control and Prevention (CDC) and addresses three specific goals for its registries: completeness, timeliness, and quality.

2.1.1.1. Standards for Reporting Requirements

Legislation and/or regulations **MUST** authorize a central cancer registry, and a mechanism **MUST** be in place to define reportable tumors, a reference date for registry operation, residency requirements for reportable tumors, who has the authority and responsibility for implementing and maintaining the database, who is responsible for reporting the data (i.e., physicians, hospitals, pathology laboratories, etc.), what geographic area is covered, timeliness of reporting, the type and format of data to be reported, and to whom and under what circumstances the central registry has authority to release the data. The legislation or regulations **SHOULD** address penalties for non-compliance.

Components of the legislation and/or regulations regarding reporting requirements include:

- All terminology used in the text of the law **MUST** be defined.

- "Cancer" **SHOULD** include all neoplasms with a behavior code of 2 or 3 (*in situ* or malignant), listed in the most recent edition of the *International Classification of Disease for Oncology* (ICD-O). Exceptions **MAY** include basal and squamous cell carcinomas of the skin and *in situ* carcinoma of the cervix uteri. Benign brain tumors are reportable starting with tumors diagnosed January 1, 2004. Some central registries may collect additional benign tumors; these should be defined in their legislation or regulations.

- "Reference date refers to the date coverage starts in a specified population at risk. The reference date is not the date the central registry is organized or actually performs the work. Tumors diagnosed on or after the reference date **MUST** be included. However, tumors diagnosed prior to the reference date **MAY** be included (see Section 2.3.1.). The reference date **SHOULD** be January 1 of a calendar year, but may be another date.

- All cancers occurring in the geographic region covered by the central registry **SHOULD** be reportable. The Registry **SHOULD** include all residents and non-residents to allow:
  - Sharing of tumor records with other population-based registries.
  - Facilitation of death clearance and other record linkages.
  - Preparation of reports for individual facilities that report all their tumors.

- Laws and regulations, for conciseness and the flexibility to make changes over time, **SHOULD** reference the more detailed documents containing reporting requirements, such as:
  - Required reporting format.
  - Registry data collection and coding manuals.
  - Outside standard references, including ICD-O, and where appropriate, data acquisition manuals.

The central registry **SHOULD** have the authority to make changes to reporting requirements as needed without additional legislation or regulations.

- The central registry **MUST** be population-based. To assure maximum coverage of the designated population, tumors **SHOULD** be reported by, or tumor information obtained from:
  - Hospitals or other facilities providing screening, diagnostic, palliative, or therapeutic services to patients with reportable tumors.
  - Physicians, surgeons, and all other health care providers who diagnose or provide treatment for patients with reportable tumors.

    Exception: Patients previously admitted to a hospital or other facility that provides screening, diagnostic, or therapeutic services and reported by those facilities.

- All cancers **SHOULD** be reported to the state or provincial health department or to another agent designated by the legislation or regulations. The legislation or regulations **SHOULD** state that tumor reports be reported to the central registry no later than 180 days from the date of admission or diagnosis. Submitted tumor reports **MUST** follow data definitions and **SHOULD** be in the NAACCR record layout (NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

The 180-day standard is consistent with the requirements of NPCR and the American College of Surgeons (ACoS) Commission on Cancer (COC) Approvals Program for hospital cancer programs. In Canada, the standard is consistent with the requirement that the source reports and physician billings be submitted within 6 months.

Under the following conditions, provisions **SHOULD** authorize the central registry to require more rapid reporting of specific tumors, as specified by law or regulations:

- Evidence exists that an epidemiologic investigation based on recently diagnosed tumors of a specific histology will assist in the further understanding of the disease.

- A specific, peer-reviewed study protocol is available for performing the epidemiologic investigation.

- Funding is available to cover the additional costs of rapid case ascertainment.

2.1.1.2. Standards for Patient Record Access

Legislation and/or regulations **SHOULD** provide access to records of health care providers and facilities that identify tumor records or establish characteristics of the tumor, treatment of the tumor, or the medical status of any identified tumor record by authorized representatives of the central registry. Access is necessary for meeting both initial reporting requirements and subsequent quality assurance activities.

Legislation and/or regulations **SHOULD** document that the authorized representative of the central registry may access information and report it in the appropriate format if a health care facility or provider fails to report in the required format.

Public Health reporting under the authority of state statutes and regulations is permitted by the Health Insurance Portability and Accountability Act (HIPAA). The Privacy Rule contains a specific provision authorizing covered entities to disclose protected health information as required by law.

2.1.1.3. Standards for Enforceability

The legislation and/or regulations **SHOULD** articulate specific penalties for:

- Failure to report tumor data.
  - The facility/provider **MAY** be required to reimburse the health department or the authorized representative for the health department's cost of obtaining and reporting data.

- Failure to grant access to all records that would identify tumor records or define tumor characteristics, treatment of the tumor, or the medical status of any identified tumor records.
  - Willful failure to grant access to records **MAY** be punishable under the law. Forms of punishment **MAY** include a fine(s) for each day access is refused (the legislation and/or regulations **MAY** specify where collected fines will be deposited—for example, the state's general fund) or revocation or suspension of a hospital's license.

2.1.1.4. Standards for Data Quality and Data Standards

The legislation and/or regulations **MUST** articulate that data reported to the central registry **MUST** meet standards of completeness, timeliness, and quality as mandated by the authorized agency for the registry.

2.1.1.5. Standards for Confidentiality and Disclosure of Data

The legislation and/or regulations **MUST** specify the confidential nature of the data and provide for confidentiality protection of all patient data. The confidentiality directives of the legislation and/or regulations **MUST** address how the data are to be released, to whom, and for what purpose. The legislation and/or regulations **SHOULD** articulate that aggregate data **SHOULD** be available to the public through published reports or through data access policies, but that access to confidential data or "raw data" is restricted. The guidelines **SHOULD NOT** be so strict that approved researchers are denied access to the raw data (see the *Data Use and Confidentiality Task Force Report*).

- Central registries **SHOULD** make all reported data available in aggregate form for use by central registry staff and authorized researchers for analyses and reports about the incidence, prevalence, management, survival, and risk factors associated with the state and province cancer experience.

- Central registries **MAY** exchange patient-specific data with the reporting facility, any other cancer-control agency, or clinical facility for the purpose of obtaining information necessary to complete the tumor record, provided these agencies and facilities comply with the registry's confidentiality policies.

- Central registries **MAY** exchange patient-specific data with other cancer registries for the purpose of complete case ascertainment if reciprocal data-sharing agreements that include confidentiality provisions are implemented.

- Central registries **MAY** grant researchers access to confidential information concerning individual tumor patients, provided the researchers comply with the registry's confidentiality policies and have the approval of the registry Institutional Review Board (IRB).

- Violation of any confidentiality provisions established by the state and province/territory **SHOULD** be punishable under the law.

2.1.1.6. Standards for Liability

The legislation and/or regulations **MUST** provide for the protection of individuals and institutions in compliance with the law. This includes provisions specifying that no person or institution will be held liable in any civil action for the reporting of tumor patient information to the central registry. Central registry staff **MUST** be protected from liability for the release of the tumor record information to entities that agree to all requirements of the confidentiality policies.

2.1.1.7. Standards for Specification of Funding Source

The legislation or regulations **SHOULD** specify the funding source(s) for the central registry (e.g., cigarette tax or general revenue). If the registry is not adequately funded, the original intent of the legislation to develop and maintain a central cancer registry is not met.

2.1.2. Reportability Definitions

Precise definitions of tumors that are reportable to the central registry **MUST** be developed and publicized. Standardized, written definitions help ensure consistent reporting by abstractors across facilities and over time. The basis for the definitions will be the reportability provisions of the enabling legislation or regulations, but more detailed definitions will be needed that reference the following:

- Reportable and non-reportable diagnoses and the reference standard (see ICD-O and also NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

- Multiple primary rules (NAACCR endorses the SEER rules as the *de facto* standard in the United States for both central and hospital-based registries).

- Reportability of non-residents and residents (see Section 2.2.5.).

- Reference date (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

- Diagnostic confirmation.

- Class of case.

- Type of admission to the reporting facility.

- Ambiguous terminology.

2.1.2.1. Standards for Reportable Diagnoses

- The central registry's reportable list **SHOULD** reference the International Classification of Disease for Oncology. At a minimum, all neoplasms with a behavior code of 2 or 3 in ICD-O-3 **SHOULD** be designated reportable. Effective January 1, 2004, and later, benign and borderline intracranial and central nervous system tumors **SHOULD** be designated reportable. The exceptions are basal and squamous cell cancer of non-genital skin and carcinoma *in situ* of the cervix uteri (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

- NAACCR recommends that population-based registries discontinue routine collection of data on pre-invasive cervical neoplasia unless there is a local need and interest and sufficient resources are available to collect all histologically confirmed high grade squamous intraepithelial lesions and its equivalent terms (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

- Any benign neoplasms or neoplasms of uncertain behavior that are reportable **SHOULD** clearly be identified with reference to their ICD-O codes. This includes benign and borderline intracranial and central nervous system tumors diagnosed January 1, 2004, and later.

- A copy of the reportable list and other rules **SHOULD** be provided to:
  - All reporting facilities or practitioners required to report;
  - All cancer registrars in the coverage area;
  - All medical records or cancer registrar training programs or schools in the area;
  - All cancer registry software providers serving the registry's area.

Professional organizations or the central registry **SHOULD** offer workshops on the reporting requirements for cancer registrars and non-registrars.

2.1.2.2. Standards for Multiple Primary Rules

To compare cancer rates between two registries, it is important that identical rules are used for counting multiple tumors in the patient—whether in the same organ, opposite sides of paired organs, different sub-sites, or different sites, and whether at the same or different times. NAACCR endorses the SEER Program rules as the *de facto* standard in the United States for both central and hospital-based registries.

SEER rules are not identical to the international standard recommended by the International Agency for Research on Cancer (IARC) and the International Association of Cancer Registries (IACR). The IARC rules have the effect of reporting fewer tumors than those that are reported using SEER rules.

The Canadian Cancer Registry (CCR) rules are different from the SEER rules for counting multiple tumors in the patient. The CCR rules are followed by most of the provincial/territorial cancer registries, but some Registries follow the IARC rules, some follow the SEER rules, and some have developed their own rules. Therefore, when data are published in Canada, the IARC rules are used to count multiple primaries, since this is the lowest common denominator. The CCR rules do not assess the time of diagnosis, nor the behavior at this time. These rules currently are being reviewed. Further details can be found in the CCR Input Data Dictionary.

2.1.2.3. Standards for Diagnostic Confirmation

To obtain complete incidence reporting and to have the central registry's data accurately reflect the burden of cancer in the population at risk, clinically diagnosed tumors as well as microscopically confirmed tumors **SHOULD** be designated as reportable.

Microscopically confirmed tumors include all tumors with positive histopathology, including examinations of bone marrow and peripheral blood; and all tumors with positive cytopathology, including peritoneal or pleural fluid, fine needle aspirations of cells, and bronchial washings.

Clinically diagnosed tumors include those without microscopic confirmation (i.e., those whose diagnoses are based only on diagnostic imaging, laboratory tests, or other clinical examinations).

2.1.2.4. Standards for Class of Case

To assure that all incident cases are reported, the registry **SHOULD** stipulate that tumors that are "non-analytic" for the reporting facility are reportable to the central registry when they meet the other requirements of reportability and date of diagnosis.

"Non-analytic" refers to a categorization used in hospital-based registries to identify tumors excluded from survival analyses, most prominently those first diagnosed or treated somewhere other than the reporting hospital (see COC *Facility Oncology Registry Data Standards* [FORDS] [Revised 2004]).

2.1.2.5. Standards for Type of Admission

To assure that all incident cases are reported and that reporting is consistent across the central registry's coverage area, the registry **SHOULD** stipulate that tumors are to be reported regardless of type of admission to the reporting facility (i.e., all tumors in the following situations are to be reported), when they meet other criteria for inclusion:

- Both inpatient and outpatient cases.
- Patients seen only in the emergency room (includes patients who are dead on arrival).
- Tumors diagnosed at autopsy.
- Patients seen for consultation only.
- Surgery centers.
- Physicians.
- Stand-alone centers.
- Pathology laboratories (includes cases in which only specimens were reviewed at the reporting facility).

However, the registry **MAY** specify a reduced reporting requirement or a separate notification mechanism (e.g., a "short form") for some of these situations. This can provide a cross check on reporting from the primary source.

2.1.2.6. Standards for Ambiguous Terminology

Diagnoses and descriptions of patients' conditions often are described in the medical record with ambiguous terms such as "possible" and "rule out." For comparability with national databases, the central registry **SHOULD** adopt rules for interpreting ambiguous terms identical to those used by SEER, COC, and the CCCR. These rules are included in their code manuals. Guidelines for ambiguous terminology also can be found in NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary.*

2.1.3. Staffing Guidelines for Data Collection

Central registry staffing needs **SHOULD** be based on the estimated annual caseload. Existing central registries can predict the annual caseload based on the experience of previous years, noting trends and projecting increases or decreases. New central registries will need to collect some baseline data to estimate the number of tumor records expected during their first year of data collection. This section refers only to staff that will be employed by the central registry and does not address the staffing needs at hospitals or other facilities that might be required to report to the central registry.

2.1.3.1. Estimating the Annual Caseload

The significant estimate to obtain is the number of case reports (not incidence cases) that the central registry staff will be responsible for identifying and abstracting. This includes hospitals that are not submitting their own reports and may also include the following: federal facilities not subject to state and provincial/territorial reporting requirements; non-hospital sources such as clinics, physician offices, pathology laboratories, nursing homes, and coroners offices; or facilities outside the registry's area. Staff also will be required to process death clearances.

- *Existing Central Cancer Registries:* In existing central registries, the annual caseload **MAY** be predicted based on previous years and the following trends:
  - Increase or decrease in the defined population.
  - New treatment facilities or the closures and mergers of existing facilities.
  - Increase or decrease in physicians treating cancer patients.
  - National standards for estimating completeness.

- *New or Expanding Central Cancer Registries:* New and existing central registries that are expanding their coverage into new areas **MAY** collect baseline data from the following sources:
  - Diagnostic indices at reporting hospitals **MAY** be reviewed and a count made of the number of discharges with a primary or secondary tumor diagnosis, noting the number of tumors that are readmissions and subtracting these from the total.
  - Pathology records **MAY** be used as an alternative method of estimating the annual caseload. Pathology reports (including biopsies, autopsies, cytology, bone marrow examinations, and consultation slides) **SHOULD** be reviewed for those reports that contain a reportable diagnosis. Hospital pathology departments and independent pathology laboratories **SHOULD** be surveyed. Five percent **SHOULD** be added to this figure to account for tumors that are not diagnosed microscopically.

2.1.3.2. Estimating Number of Data Collection Staff

This estimate applies only to tumor records the central registry is directly responsible for abstracting, and does not apply to tumor records submitted by hospitals or other reporting sources. Although no firm standard can be stated, some central registries use the rule of thumb that one abstractor can be responsible for 1,000 to 1,200 abstracts per year, including casefinding, abstracting, and coding directly onto a computer. This will

vary because of the factors listed above. For example, if extensive travel is needed, fewer tumor records can be abstracted. The rule of thumb does not take into account either patient follow-up or the various support activities (data processing and clerical) for data collection that require staff.

- *Availability, Completeness, and Extent of Patient Records:* Consideration **SHOULD** be given to medical record completeness and the types of reporting facilities in the central registry's area (e.g., teaching hospitals, research facilities, health maintenance organizations [HMOs], and clinics). The more comprehensive the patient records and the more complex the care given to patients, the more time that is required to collect registry data.

- *Dataset:* The data **MAY** be limited to items needed for incidence only, **MAY** include treatment and follow-up, or **MAY** further include items for a special study of a specific disease process.

- *Reporting Facilities:* The location of reporting facilities in relation to the central registry impacts the amount of time required for staff to collect data. Data collection staff **MAY** be required to travel great distances to collect the required data. The type of facilities reporting data also needs to be considered. The data required and available from freestanding clinics, surgery centers, group practices, prison hospitals, and military facilities **MAY** vary, and the central registry staff **MAY** need to visit some facilities that are not required to report.

2.1.3.3. Data Collection Method

Information technology has been changing the way data collection processes are carried out in the central registry, and computerization has improved registry productivity. Compared with manual methods, the use of portable computers and standardized data collection software for abstracting and coding increases the number of tumor records each staff member can collect. Similarly, recent developments in electronic pathology reporting for casefinding and Web-based cancer reporting are improving productivity and changing staffing patterns.

2.1.3.4. Training

The type of data collected and the format used dictates the technical expertise necessary for complete case ascertainment. Some on-the-job training may depend on the educational background and experience of the data collection staff. Standards for training are addressed in Section 2.2.10.

2.1.3.5. Standards for Type of Staff

Staffing levels **MUST** be adequate to assure compliance with mandated reporting requirements for timeliness, completeness, and accuracy of data collection.

Data collection staff **MUST** know general anatomy and physiology, the disease process of cancer, casefinding procedures; and basic coding, disease classification, and staging schemes.

Certified Tumor Registrars (CTR) or those who are CTR-eligible **SHOULD** be used for performing data collection activities. A CTR **SHOULD** be used for supervising data collection activities.

2.1.3.6. Standards for Continuing Education

Continuing education **SHOULD** be provided to data collection staff to assure that they have up-to-date knowledge about diagnostic and treatment modalities and are able to retain certification status. The National Cancer Registrars Association (NCRA) maintains the continuing education information related to CTRs (20

continuing education hours must be completed in a 2-year cycle). The central registry **MAY** offer training, or staff **MAY** be given time and travel funds to attend programs offered outside the registry. Continuing education **SHOULD** be available in the following areas:

- Tumor diagnosis, staging, and treatment.
- Data management.
- Epidemiology and statistics.
- Hardware and software applications.

Data collection staff **SHOULD** be supplied with appropriate references and literature to provide ongoing continuing education and to answer questions that arise. Current medical reference books **SHOULD** be immediately available in the areas of anatomy and physiology, tumor diagnosis and management, and basic medicine and pathology. Pertinent journals and other periodicals also **SHOULD** be readily available. Staff **SHOULD** be informed about the Cancer Information Service at 1-800-4-CANCER. The central registry **MAY** provide access to electronic bulletin board services and online resources such as the National Library of Medicine's MEDLARS databases. These include PDQ, CANCERLIT and MEDLINE. Other resources for continuing education include ACoS, NAACCR, NCRA, NPCR, SEER websites (see Appendix E). These services will provide the staff with rapid access to the most current information and educational opportunities.

Central registry staff **SHOULD** be encouraged and funded to participate in local and national professional associations such as state/provincial/territorial registrars' associations, the NCRA Annual Educational Conference, the NAACCR Annual Meeting, the Annual Canadian Cancer Registry Technical Workshop (CCR TWS), and the Canadian Health Information Management Association. The registry budget **SHOULD** include funds for participation by one or more persons at annual association meetings. The registry **SHOULD** consider sending staff to special symposia, conferences, and courses.

2.2. PROCESS STANDARDS

2.2.1. Hospital Reporting

Participation of all hospitals in the reporting area that diagnose, evaluate, or treat cancer is essential to ensure completeness of reporting.

2.2.1.1. Standards

The central registry **SHOULD** gain access to 100 percent of the hospitals in its reporting area to ensure completeness of reporting at the hospital level. Letters of agreement **MAY** be useful for both the hospital and the central registry. These letters **SHOULD** specify the responsibilities of the hospital, the responsibilities of the central registry, and the timeframe for reporting. Also, state and provincial/territorial reporting laws that allow the central registry to enforce reporting, and any such enforcement procedures **SHOULD** be included in the letters of agreement.

State, provincial, or federal laws pertaining to patient privacy may exist that apply to specialty hospitals, such as mental health facilities, chemical dependency facilities, and hospitals in state penitentiaries. This issue **SHOULD** be considered when initiating tumor reporting discussions with these specialty hospitals.

2.2.1.2. Standards for Federal Facilities

Federal facilities, such as military hospitals, Veterans Administration hospitals, and hospitals in federal penitentiaries, are not subject to state reporting laws. Therefore, the central registry **SHOULD** actively pursue obtaining voluntary participation of such facilities. The central registry **SHOULD** identify staff at the federal facility to assist in working with the administration to achieve voluntary participation. Once the administration agrees to voluntary participation, a letter of agreement **SHOULD** be signed. Historical documentation of the federal facility's voluntary participation can aid the central registry in the future as the facility's administration experiences turnover.

2.2.2. Non-Hospital Sources Reporting

Because of the shift in health care toward ambulatory or outpatient services, it is expected that the number of patients seen for diagnosis, evaluation, or treatment in outpatient settings will continue to increase. Capturing these tumor records through an extended reporting system is important to ensure the completeness of tumor registration. Central registries **SHOULD** expand their coverage to non-hospital sources to facilitate complete reporting (e.g., independent pathology laboratories).

This section refers to facilities that provide medical services to patients. The vital statistics agency in the registry's area also is an important source of case ascertainment, and it is covered separately in Section 2.2.9.

2.2.2.1. Standards

The central registry **MUST** develop mechanisms to locate and obtain information on tumors diagnosed or treated entirely outside of hospital settings (for further information, see *Procedure Guidelines for Cancer Registries Series IV: Cancer Case Ascertainment*). The usefulness of specific sources will vary across geographic areas and over time. However, experience has shown that at a minimum, the central registry **SHOULD** obtain tumor records from the following types of facilities:

- Independent pathology laboratories (histopathology and hematology laboratories).
- Ambulatory surgery centers.
- Radiation therapy centers.
- Outpatient oncology centers.

Although reportable tumors **MAY** be identified in pathology laboratories, the laboratory records often contain insufficient information for preparing a complete abstract. Information on the patient's residence and/or health insurance number, for example, rarely is present. These cases usually **MUST** be followed back to the treating physician or facility (for additional information, see Section 2.2.3).

The expansion of case ascertainment procedures into all types of non-hospital facilities would ensure complete reporting; however, the central registry's ability to do so **MAY** be limited by its financial resources. Therefore, the registry **SHOULD** consider the following items when evaluating the expansion of casefinding into non-hospital facilities such as chemotherapy treatment facilities, coroner's offices, private clinics, nursing homes, and hospices:

- The cost of accessing each type of facility. The cost depends on the reporting law and which types of facilities and practitioners are required to report. The cost also depends on whether the reporting process is manual or electronic.

- The quality of the data and the number of new incidence cases that would be obtained from each type of facility.

- The impact on the future use of the data if a decision is made not to collect data from a specific type of facility.

- The impact of these requirements on each type of facility.

2.2.3. Physician Reporting

Because not all persons diagnosed with a tumor are hospitalized for diagnosis, evaluation, or treatment, a mechanism for registering tumor records from physicians' offices is necessary for complete case ascertainment. The central registry **MAY** rely on active reporting by physicians, or **MAY** have its own staff obtain the data from physicians' offices. The registry generally will require patient or tumor information from an individual physician only when no report is obtained from a hospital or other reporting facility. However, the central registry also might need to obtain demographic or treatment information on tumors reported initially by other sources.

2.2.3.1. Standards

The central registry **SHOULD** perform the following:

- Follow-back to physicians' offices to obtain reports on otherwise unreported tumors identified in pathology laboratories, through consult-only reports from hospitals, or from death certificates.

- Develop an appropriate method to identify tumors and obtain information from hematologists, dermatologists, dermatopathologists, oncologists, gynecologists, and urologists. These specialties are most likely to diagnose malignancies that will not be identified through the active casefinding methods used at hospitals and laboratories.

- Develop registration methods for physicians.

2.2.4. Liaison(s) With Outside Agencies and the Medical Community

Even though tumor reporting may be required by law, the efficient and effective operation of the central registry rests on the continued good will of physicians, staff at reporting facilities, and governmental agencies with whom the registry works on a day-to-day basis. Broad support from the general public, voluntary agencies, and community special interest groups can also be important to the central registry's continued existence. In a complementary sense, the registry often will need medical and other advice from the wider community. Formal mechanisms **SHOULD** be in place for these liaison and advisory functions. The central registry **SHOULD** actively cultivate liaisons with a wide variety of agencies and professional groups. Methods **MAY** include attendance and/or presentations at group meetings, use of newsletters, collaboration on various projects, and serving on committees.

2.2.4.1. Standards for Medical Advisors

The central registry **MUST** designate medical advisors—physicians who agree to serve, usually without compensation, to consult with the registry staff as needed on questions of medical data interpretation, diagnosis and management, and/or classification of issues. The central registry generally will require at least one pathologist and one clinical oncologist advisor. Identifying physicians who have an interest in and understanding of the needs of registries is crucial. Maintaining a long-term relationship with the advisors is especially helpful in achieving continuity and consistency. The mechanism of obtaining advice **MAY** range from informal telephone calls to discuss questions to regularly scheduled meetings of the advisor(s) with key registry staff.

The central registry **SHOULD** designate an individual on its staff to handle requests to the advisors and **MUST** document all decisions made through consultation with the advisors.

2.2.4.2. Standards for Community Advisory Boards

The central registry **MAY** institute an advisory board. In some cases, an advisory board may be required as part of the registry's formal governance; in other cases, the board's role will be strictly advisory. Composition of the board will be unique to the community served, but should be broad-based and represent medical interests, academic researchers, public health and government agencies, cancer registrars, voluntary agencies such as the Cancer Society, and national advocacy or special interest groups.

2.2.5. Out-of-State and Province/Territory Coverage, Case Sharing, and Coverage of Non-Residents

Identification of residents of the central registry's coverage area diagnosed in other areas is essential for complete population-based reporting. Collecting these tumor records from surrounding state/provincial/ territorial registries often is possible because many registries collect information on non-residents if they are diagnosed and/or treated in their area. Additionally, to obtain pathology reports of residents in their areas, central registries **MAY** contact national pathology laboratories, although many laboratories do not maintain residency information on their patients. Hospitals and pathology laboratories located in bordering states and provinces/territories often exchange data for the purpose of obtaining complete coverage.

2.2.5.1. Standards

The central registry **MUST** include all reportable tumors occurring in residents of its coverage area, regardless of where the tumors are diagnosed or treated.

The central registry also **SHOULD** include all residents and non-residents diagnosed or treated in its coverage area to allow for sharing of tumor records with other population-based registries, facilitate death clearance and other record linkages, and allow for preparation of reports to individual facilities that include all of their tumor records. The registry **SHOULD** record the complete address at diagnosis for its non-resident tumor records as well as resident tumor records in a form that allows electronic sharing of the full address.

The central registry **SHOULD** provide information on a non-resident to the population-based registry covering the patient's place of residence when the required components listed below are in place. The shared information **SHOULD** include confidential and non-confidential data and abstracted text summaries as described in the current NAACCR *Standards for Cancer Registries Volume II*: *Data Standards and Data Dictionary*.

The central registry **SHOULD** analyze the results of case sharing and data exchange (NAACCR's *Procedure Guidelines for Cancer Registries Series I: Interstate Data Exchange)*.

2.2.5.2. Required Components

The following components generally will be required for the performance of case sharing between registries:

- *Case Sharing Agreements:* These are written agreements between registries covering the usage and confidentiality of exchanged data. These **MAY** be informal agreements simply requesting data and affirming the confidential nature of the data or they may be longer, more formal legal documents, depending upon the laws governing release of data. For an example, see Appendix F (for more information on case sharing agreements, see NAACCR's *Procedure Guidelines for Cancer Registries Series I: Interstate Data Exchange*).

- *Exchange Media:* Data **MAY** be exchanged between central registries across a variety of media. In order of preference, they are: electronic files of data on diskette, CD ROM, or tape; electronic files of data transferred by e-mail or website; copies of paper abstracts; or printed reports generated from computer systems (see Section 2.2.8.1. for information on data security).

- *Exchange Format:* The nationally accepted format for tumor data exchange is the current NAACCR data exchange format, as it is comprehensive and contains standard data items and definitions (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*). Use of the standard format means that each registry's computer system needs to read and write only one format, instead of reading and writing a different format for each registry with which data are being shared.

If the sending registry uses a different version of the NAACCR data exchange format, the receiving registry may need to convert the data into its format for entry into its system.

- *Staff:* Personnel to read and convert data, and coding and data-entry staff are needed to convert information received from diskettes and paper abstracts.

- *Data Compatibility:* Data definitions and codes sometimes vary among central registries. However, central registries **SHOULD** ensure that all transmitted data follow the standard definitions of the NAACCR data exchange format.

2.2.6. Reporting Requirements

To encourage compliance with tumor reporting requirements, the central registry **SHOULD** notify facilities and practitioners that are required to report of their obligations. The registry **MAY** be required to do so by law or regulation (for more information, see *Procedure Guidelines for Cancer Registries Series IV: Cancer Case Ascertainment*).

2.2.6.1. Standards

The tumor reporting notification **SHOULD** include:

- A brief description of the central registry's history and purpose.

- A description and copy of the cancer reporting law.

- The rationale for the central registry's access to the source data.

- The data items to be collected.
- The procedures for reporting.
- All relevant considerations for data handling and ensuring data security and confidentiality.
- A brief statement that the HIPAA privacy regulations do not restrict the disclosure of patient information by a health care provider to a central registry so long as the central registry is a "public health authority."

The following notification activities **SHOULD** be carried out:

- Support of the central registry and its reporting methods from appropriate groups **MAY** be sought. Examples include medical societies, specialty colleges or boards, community groups, and the American Cancer Society. Citing such support or endorsements in the various communications to medical professionals may encourage their compliance.
- Announcements **MAY** be made through professional organizations or societies regarding their members' tumor-reporting responsibilities. The mechanisms **MAY** include newsletters, direct mailings, journal articles, and presentations at scheduled meetings.
- In addition, the exact details of all expectations of and options available to the facilities and practitioners **SHOULD** be communicated by targeted contacts. The means for accomplishing these steps include:
  - Direct mailings to individuals.
  - Meetings with groups, such as staff of large clinics or specialty laboratories.
  - Presentations at scheduled meetings, such as hospital staff meetings or local medical society meetings.
  - Regional presentations and orientation workshops organized by the central registry.
- These communications **SHOULD** be directed to:
  - All relevant physicians (e.g., pathologists, medical oncologists, dermatologists, general surgeons and surgical specialists, and radiation oncologists).
  - All related facility personnel (e.g., hospital administrators, health information service administrators, and cancer registry managers).

In cases for which time and money permit, consideration **SHOULD** be provided for the implementation of procedures to access the source data (e.g., 6 months or longer). Specific deadlines **SHOULD** be provided to conform to the central registry's reference date.

2.2.7. Monitoring Use of and Changes in Reporting Facilities and Practitioners

Population-based registries **MUST** be able to document that they capture tumors from the entire population at risk for their area. To do so, they **MUST** be able to document where residents of their population receive tumor diagnoses and how the registry identifies these tumor reports.

Central registries **SHOULD** monitor changes in the number and location of facilities and practitioners and where their area's patients are being diagnosed and treated. Facility openings, closings, and mergers and the establishment of new screening programs all can impact workload and procedures for the registry by

influencing the number of tumors diagnosed and the number and location of sources the central registry needs to cover.

2.2.7.1. Standards

The central registry **MUST** be aware of the flow of patients to areas outside the registry coverage area for diagnosis and treatment; the closing of hospitals and clinics and the opening of new ones, including screening and treatment centers; mergers of facilities that impact the operation of hospital registries and the central registry; and shifts in utilization of screening, diagnostic, or treatment facilities that would impact where patients are diagnosed and treated. The central registry **MAY** obtain information from governmental licensing agencies and also **SHOULD** conduct periodic surveys and review telephone directories, local newspapers, professional association publications, and the Internet.

2.2.8. Confidentiality Policies and Procedures: Issues in Data Collection and Management

Confidentiality policies and procedures are required in all phases of central registry operations (see the *NAACCR 2002 Workshop Report, Data Security and Confidentiality*) to:

- Protect the privacy of the individual patient.
- Protect the privacy of the reporting facilities.
- Provide public assurance that the data will not be abused.
- Abide by any confidentiality-protecting legislation or administrative rules that may apply.

Aspects of confidentiality policies and procedures that relate to authorized use of and release of data are addressed in Section 4.1.1.

- *Definition of Confidential Data:* Although the tumor reporting laws and regulations under which the registry operates may define patient-specific data as confidential, central registries also **SHOULD** treat any information that specifically identifies a health care professional or an institution as confidential. Information that characterizes the caseload of a specific institution or health care professional also **SHOULD** be considered proprietary and confidential.
- *The Registry's Responsibilities:* It is the responsibility of every registry to protect its data from unauthorized access and release. The central registry **MUST** maintain the same standards of confidentiality as customarily apply to the doctor-patient relationship; this obligation extends indefinitely, even after a patient's death.

The costs of inappropriate release of confidential data are many. Inappropriate release of data could damage an individual whose diagnosis of cancer was made public. Support and cooperation of facilities submitting data to the central registry could be severely compromised.

Data security policies and procedures **MUST** address general confidentiality practices, electronic data security, paper record security, educational programs for central registry staff, network security, etc. (see the *NAACCR 2002 Workshop Report, Data Security and Confidentiality*).

2.2.8.1. Standards for Policies and Procedures for Data Security

The following components generally will be required to assure data security:

- The director of the central registry **MUST** be responsible for data security.
- Suitable locks and alarm systems **MUST** be installed to control access to the central registry, and the director **SHOULD** maintain a list of persons authorized to enter the registry.
- Central registry staff **MUST** be responsible for the confidentiality of all data encountered during the collection of tumor data.
- Confidential data **MUST NOT** be transmitted by any means (mail, telephone, fax, electronic) without the explicit authority from the director or a staff member to whom such authority has been delegated.
- Central registries **SHOULD** consider the use of registered mail, overnight mail, or courier services for confidential data and **SHOULD** consider separating names from other data for transmission. When using mail services, registries **SHOULD** consider using double envelopes, with the confidential information in a separate envelope marked "confidential," including a contact telephone number, and enclosed in the mailing envelope. Registries **SHOULD** consider using tear-free envelopes marked "confidential."
- Precautions **MUST** be taken for both the physical and electronic security of confidential data.
- Computer use of confidential data **MUST** be controlled by electronic and, if possible, physical measures to enhance the security of the data, including the use of a separate room, use of passwords, automatic logging of all attempts to enter the system, and different levels of access to the data.
- Training and demonstrations of the computer system **SHOULD** be performed with separate fictitious or anonymous datasets.
- Consideration **MUST** be given to obtaining expert advice on security against unauthorized remote electronic access if it is impossible to use isolated data processing systems.
- Measures **MUST** be taken to ensure the physical security of confidential data stored on paper, microfilm, microfiche, etc.
- A policy **MUST** be developed for the safe disposal of confidential data. If a private document destruction company is used, the central registry **MUST** have documented procedures for disposal of confidential data and the security measures used by the company's employees.

2.2.8.2. Standards for Personnel Policies and Procedures

Central Registry Staff Members/Employees:

The central registry staff (including students, volunteers, and contractual workers) **MUST** sign, as part of their employment agreement, a declaration that they will not release confidential information to unauthorized persons. This declaration **SHOULD** remain in effect after cessation of employment. The director **SHOULD** maintain a list of staff members indicating the nature and extent of their access to registry data.

- The training of all central registry staff **MUST** include a comprehensive session concerning the confidentiality of data.

- Failure to observe the confidentiality policies **MUST** result in firm disciplinary action, including the potential for termination of employment. Some circumstances **MAY** warrant legal action against central registry staff members who fail to comply with the registry's confidentiality policies. Depending on the jurisdiction, there also **MAY** be criminal penalties for failure to maintain the required confidentiality.

- Central registry staff **SHOULD** re-sign the confidentiality agreement at least annually and could be timed with employees' annual performance evaluations.

Non-registry staff:

- Non-registry staff, especially medical investigators, **MAY** request access to confidential registry data. Such requests **MUST** be in writing. All non-registry staff that request access to these records **MUST**, at a minimum, agree to adhere to the same confidentiality safeguards practiced by central registry staff (see Section 4.1.1.).

- Most requests **MAY** be adequately addressed without the release of confidential information. Whenever possible, it is preferable to respond to requests without the use of confidential information.

2.2.8.3. Standards for Policies and Procedures for Release of Registry Data

Release of central cancer registry data for clinical purposes, for research, and for health care planning is important to the utility of the registry, and the registry **MUST** develop procedures for data release that ensure the maintenance of confidentiality. See Section 4.1.1. for a detailed discussion of confidentiality issues in research, reporting, and release of registry data.

- For the purpose of complete case ascertainment, the central registry **MAY** exchange confidential data with other central registries if reciprocal case-sharing agreements (see Appendix F) that include confidentiality provisions are implemented.

- The central registry **MAY** permit the release of confidential data to treating hospitals in their own or other states and provinces/territories for the purpose of patient follow-up.

- It is recommended that plans be made for the possible cessation of central registry activity to maintain the subsequent utility of the database while safeguarding the confidentiality of its data.

2.2.9. Death Clearance

Death clearance is an essential step in achieving complete population-based reporting (see *Procedure Guidelines for Cancer Registries Series V: Resolving Death Clearance Issues 2003*). It serves as a check on the completeness of reporting from other sources and often identifies tumor records that should have been reported from those sources but were not. It also identifies patients known only to the physician. Furthermore, tumor records that remain as death certificate only (DCO) cases, after follow-back, **MUST** be included as incident cases by the registry.

Death clearance for this purpose means identification of all deaths with cancer mentioned as a cause of death that are not accounted for in the registry's files. This section does not address death clearance for the purposes of obtaining follow-up on tumor records already registered.

2.2.9.1. Standards

Death clearance for the purposes of case identification **SHOULD** be performed when the death files are complete for the calendar year being cleared, and with enough time for follow-back to be completed and the results incorporated into the central registry's database before the registry publishes cancer incidence rates for the calendar year. Timing must be planned carefully. The goals are to link every cancer from the time period against every death from that period, avoiding unnecessary follow-back but distributing the follow-back workload across a reasonable time.

In practice, death clearance usually is performed more than once for tumor records in a given time period. The death file for a given year may not be completed soon enough to meet the central registry's needs, either because of coding delays at the vital statistics office or because not all deaths of state/province/territory residents occurring in other jurisdictions have been incorporated (states and provinces/territories exchange death records on residents from other locales through the transcript exchange program). The central registry's files also may be incomplete at the time of initial linkage. Early linkages **MAY** be performed with incomplete death or registry files. Additional linkage or linkages then **MUST** be performed when the registry considers its case file to be complete and the vital statistics office considers the death file complete for the year.

The central registry also **SHOULD:**

- Include a tumor linkage comparison in its death clearance (i.e., verify that, for patients in both the registry file and the death file, the cancers are of the same primary sites). If there are discrepancies, follow-back as necessary to determine if the patients had additional reportable tumors that should be registered. The rules in Canada and the United States may differ regarding the handling of discrepant diagnoses.

  All registries apply the World Health Organization (WHO) rules (manually and/or with an automated mortality classification system) to classify and select the underlying cause of death. Across jurisdictions, the number of codes kept in the vital statistics database may differ. At a minimum, the ICD-10 code for the underlying cause of death will be tabulated, with some agencies maintaining a multiple causes of death file that would include codes for the other causes of death recorded on each certificate. Varying outcomes of tumor linkage may occur if one registry receives, from the vital statistics agency for death clearance, just the underlying cause of death code, and another registry receives the underlying cause of death code and multiple causes of death file.

- Employ standard coding for DCO cases in their files.

- Analyze the results of death clearance, monitor them regularly, and use the information as feedback in the quality control cycle to improve casefinding and completeness of reporting from hospitals and other sources.

2.2.9.2. Required Components

The following components generally will be required for the performance of death clearance:

- The central registry **SHOULD** establish a formal agreement with the state, provincial, or national vital statistics office covering access to computer records and paper files, subsequent use of death record information, and costs.

- The central registry's computer system **MUST** have the ability to perform record linkage between the death files and tumor records and identify matches, non-matches, and potential matches with cancer as a cause of death.

- The central registry **MUST** have staff adequate in number and trained in casefinding and abstracting to perform follow-back. An estimate for staffing is that one full-time equivalent (FTE) per 18,000 annual cases, on average, should be sufficient to attain a DCO rate lower than 2 percent.

- The central registry **SHOULD** have a system for tracking progress and results of follow-back. This system preferably **SHOULD** be automated, but **MAY** be manual.

2.2.10. Training in Casefinding and Multiple Primary Determination

To ensure that the personnel actually performing case ascertainment and abstracting are aware of the reporting rules and methods, it is important to make training available. SEER rules to determine multiple primaries are the *de facto* standard in the United States and the NAACCR standard for both central and hospital-based registries. The SEER website includes training modules for casefinding and multiple primaries (http://training.seer.cancer.gov/).

2.2.10.1. Standards

Before data collection for the central registry begins, the registry **SHOULD** provide training, in the following areas, to all personnel who will be responsible for tumor identification and abstracting:

- Criteria for case reportability.

- Rules for multiple primary determinations.

Training **SHOULD** be provided to central registry staff and to staff in all reporting facilities where the staff may be identifying tumor records for the registry. Specific training is important when non-CTRs will be identifying tumor records for the central registry, but CTRs also will require specific training for the central registry reporting requirements.

Training **MAY** be offered at professional association meetings or at workshops scheduled by the central registry. Professional publications and central registry newsletter articles also **MAY** be used to deal with reporting problems.

See Section 3.2.4. for other training standards.

2.2.11. Monitoring Completeness of Reporting and Ensuring Compliance by All Reporting Facilities and Practitioners

Monitoring the completeness of casefinding for reporting facilities is a required component of the central registry's quality control operations. Even when the reporting facilities are performing the casefinding, it ultimately is the central registry's responsibility to verify that the facilities are reporting all appropriate tumors and to take corrective action when problems are discovered.

2.2.11.1. Standards

The central registry **SHOULD** monitor the processing of the casefinding sources on a regular basis. Frequent monitoring enables the registry to quickly identify problems and take corrective action. Facility-specific management reports used to monitor the status of reporting should be shared with the facility (see *NAACCR Cancer Registry Management Reports: Design and Implementation*).

The central registry **SHOULD** prepare and review various management reports such as the following to monitor the status of reporting:

- Completeness of reporting for each facility, each county, and the entire coverage area.
- Status of screening of the casefinding sources, such as each type of pathology report (i.e., surgical specimens, cytologies, autopsies, bone marrows, etc.), disease and operations indices, and radiation treatment logs for each facility.
- Status of death clearance processing.
- Counts for primary site tumors, for applicable facilities, and for the entire coverage area that are diagnosed and/or treated in an outpatient setting so that potential non-hospital underreporting is identified.
- Report of the percent of histologically confirmed tumors for each reporting facility may identify time periods where some casefinding sources were not reviewed.

When the number of reported tumors deviates widely from the number expected, the central registry **SHOULD** undertake the necessary procedures to determine the possible reasons. Tumor reporting may be late or incomplete, or the numbers may accurately reflect changes in the occurrence or distribution of cancer. A hospital's census may be down, patients may have shifted to another hospital or clinic, or expected population growth may not have occurred.

If the state/provincial/territorial reporting law provides for a means of enforcing the reporting by facilities and practitioners, the central registry **MUST** undertake the necessary procedures to obtain complete reporting from all facilities.

2.2.12. Casefinding Audits

Although observed-to-expected ratios and incidence-to-mortality ratios can provide some estimates of the level of completeness of registration, they reflect how the registry performs as compared to the previous history. Cancer incidence and/or the diagnostic practices in a registry catchment area may or may not be the same as in previous years.

The design of an audit will depend on the definition of "cancer", the reporting practices of the institutions in the area, reporting requirements and policies, and ascertainment methods used by the registry.

Central cancer registries **SHOULD** perform an independent review of casefinding sources in reporting facilities to determine facility reporting completeness.

2.2.12.1. Standards for Types of Audits

More than one type of audit **SHOULD** be used to assess completeness. Generally, each reporting facility should be routinely audited at least once every 3 years. Audits should be conducted when there is a documented decline in case reports from a facility (i.e., less than 90 percent of the previous year's case submission) in the data, evidence of other problems in reporting data, a change in reporting requirements, or as part of special studies. A rotating schedule **MAY** be set up for performing various types of audits. Audits **MAY** include, but not necessarily be limited to:

- Comparison of (an) independent method(s) of case ascertainment with tumors routinely reported, generating an estimate of percent completeness.

- Special studies to analyze the effect of including or excluding certain possible sources of cancer case identification on the completeness of case ascertainment (e.g., study to assess the impact of ignoring radiology logs, gynecological cytologies, etc.).

- Surveys of medical practitioners who might diagnose a reportable tumor outside of the usual sources of case identification (e.g., dermatologists who read their own slides, out-of-state pathology laboratories that process specimens from the registry's area).

- Other audit designs will be appropriate, based on the definition of "cancer," the reporting regulations, medical practice and referral patterns, and the geography of different states and provinces/territories.

2.2.13. Patient Follow-Up

Registries intending to evaluate survival and/or quality of life **MUST** follow all registered patients for life (often, carcinomas *in situ* of the cervix uteri and basal and squamous skin cancers, if registered, are not followed). Methods of obtaining follow-up will vary due to local considerations, such as the number of tumor records being followed by hospital cancer programs and the availability of databases against which the registry files can be linked.

Follow-up methods are classified as active or passive. Active follow-up includes contact(s) on an individual made with a primary source (i.e., individual or physician) or secondary source (i.e., online access to individual information) to update information on the individual. Passive follow-up updates information on the individual by use of linkage(s) with external databases. Central registries usually will need to employ a combination of complementary methods to achieve acceptable levels of success and avoid bias in the lost-to-follow-up group.

2.2.13.1. Standards

The choice of methods or sources for obtaining patient follow-up **SHOULD** be driven by:

- The availability of the method or source to the central registry.
- The effectiveness of the method or source.

Patients that are listed in any database as "do not contact patient" **SHOULD** be excluded from all follow-up activities that include any type of patient contact, but **SHOULD** remain in passive follow-up procedures and selective active follow-up processes such as sending letters to physicians.

Passive follow-up sources include, but are not limited to:

- Department of Motor Vehicles (DMV) files of licensed drivers.
- U.S. Centers for Medicare and Medicaid Services (CMS).
- State/provincial/territorial death files.
- U.S. Social Security Epidemiological Vital Status Data.
- U.S. Social Security Administration Death Master File.
- U.S. Election and voter registration files.
- Canadian Mortality Database.
- U. S. National Death Index (NDI).
- U.S. HMO or other health plan files with service and billing dates.
- Hospital discharge data.

When the central cancer registry and the hospital-based cancer registries are both performing follow-up activities, efforts should be coordinated so that information sources are not contacted repeatedly for the same data. Commonly used sources for active follow-up include, but are not limited to:

- Hospitals.
- Local/family physicians.
- Specialist physicians.
- Nursing homes.
- Telephone books.

- The Internet.
  - Telephone books.
  - Reverse directories.
  - Genealogy.
  - Social security number search.
  - Newspaper archives.

Use of each source **SHOULD** be evaluated with the following criteria:

- Is the source available to the central registry?

- Are appropriate linkage variables available in both the case file and the external file so that linkage is possible?

- Can the central registry's computer system perform the required linkage?

- Will the central registry maintain control over the confidentiality of its case files in any linkage activity?

- Is the method appropriate to the population being followed? For example, U.S. Medicare files contain information primarily on those ages 65 and over.

- Will the method contribute to the overall success of the follow-up effort or compensate for a bias in other methods used?

- Will the method provide timely follow-up? For example, motor vehicle department files may contain information on license renewal that may only occur every 5 years, or voter information may only be useful in election years.

2.3. OUTCOME MEASURES

2.3.1. Percent Death Certificate Only

The NAACCR method for calculating DCO cases is a multi-step process.

Step 1 is the matching of death records for a specific year against all records in the central cancer registry and identifying those records that do not match.

Step 2 is the elimination of non-reportable cases, such as:

- Deaths not caused by cancer but coded as a cancer death.

- Out-of-jurisdiction residents.

- Cancers diagnosed before the central cancer registry reference date.

Step 3 is the resolution of potential DCOs. This means that the remaining unmatched cases must be cleared according to the central cancer registry's death clearance protocol. Cases that are not resolved at the time the DCO rate is calculated are true DCO cases.

Step 4 is the calculation of the DCO rate.

$$\frac{\text{(\# of true DCOs for the year)}}{\text{(Total \# of cancer cases for the year)}} \times 100 = \text{DCO rate}$$

The percentage of DCO cases traditionally has been used to measure registry completeness. In long-standing central registries with very complete coverage, the percentage of DCO cases probably is more efficient at measuring the quality and quantity of follow-back activities. A more useful measure might be the proportion of cases initially identified through death certificates that would otherwise have been unreported, regardless of their eventual type of reporting source, but this is not a measure for which there is any consensus on codes or any history of collection. Central registries continue to use percent DCO because it is simple and identifies registries that clearly are incomplete, although it does not discriminate well among relatively complete registries.

For new central registries, the first year of death certificate follow-back will be the most difficult because of the number of prevalent cases on the death file (i.e., the number of patients dying of a cancer diagnosed prior to the registry's reference date).

2.3.1.1. Standards

NAACCR has established criteria for recognizing population-based cancer registries that achieve excellence and is awarding gold and silver certificates for those registries that meet pre-established criteria. The NAACCR standard for DCO is less than 3 percent for gold and less than 5 percent for silver.

The contractual standard for SEER registries is a 1.5 percent DCO rate. Values greater than 1.5 percent require analysis and explanation. If the DCO percentage rate is 0, death clearance has not been performed. DCO percentage rates of more than 3 probably are a result of underreporting from other sources, or from incomplete follow-back, or both.

2.3.2. Observed and Expected Case Counts

Incomplete ascertainment of tumor records can result in artificially low incidence rates and can lead to incorrect conclusions about the cancer burden in the population. There are a number of ways central registry staff can determine the level of data completeness in the cancer registry: calculating the percentage of cases identified by DCO; analyzing collected data to be sure they follow known patterns (e.g., incidence > mortality); and, most importantly, conducting special studies or audits. Additionally, the comparison of the expected number of tumor records for a given population with the observed number of unduplicated tumor records submitted to the registry over a specified time period is very useful in determining whether standards of case ascertainment are being met and whether the data collected by the registry are complete enough for analysis.

2.3.2.1. Methodology for Calculating Observed and Expected Cases

Many methods **MAY** be used to calculate expected numbers of cases, from the simple to the very sophisticated. It is preferable that estimates be based on actual incidence data for the population at risk or, if those data are not available, on incidence data for a population similar in racial composition. For the most accurate estimate of expected numbers, some method of adjusting for time trends **MAY** be included, although this adds to the complexity of the calculations.

The method that NAACCR uses to measure completeness of case ascertainment is the incidence-to-mortality rate ratio. Previously, the use of mortality rates was not useful, but the interpretation of incidence-to-mortality rate ratio has become refined. The use of this method makes the following assumptions:

- Cancer death rates are complete.
- The ratio of SEER Incidence to U.S. Mortality rates is 80% and is similar within race-sex site groups (20% allowance for variation in case fatality).

For a complete list of assumptions and the calculation method, see Appendix G.

All calculations and analyses addressed in this section and in Section 2.3.3. assume that duplicate records for persons and tumors have been eliminated, that each tumor record is counted only once, and that all patient and tumor information has been consolidated.

2.3.2.2. Standards

- The central registry **SHOULD** compare observed and expected numbers at regular intervals during the year.
- If the size of the population is large enough to yield stable numbers, expected case counts **SHOULD** be compared to observed counts by county and/or region of the coverage area, by race/ethnicity if minorities make up an important part of the population, and by cancer site. Sites comprising the greater proportion of cancers reported to the central registry **SHOULD** include breast, colon and rectum, lung, and prostate.
- The expected number of cases **SHOULD** be evaluated and revised annually based on actual numbers of cases and other considerations, such as known trends toward increasing or decreasing rates of cancer of specific sites or changes in the population due to in- or out-migration.
- Interpretation of the comparison of observed and expected counts requires a thorough knowledge of the underlying population. There **MAY** be specific reasons other than problems in data collection as to why observed numbers are higher or lower than expected.
- Calculating and interpreting the comparison of observed and expected counts **SHOULD** not supplant other quality control activities, particularly casefinding audits.

2.3.3. Other Analyses

Experience has shown that certain patterns occur in cancer data. Non-conformance with one or more of these patterns may indicate incorrect data. The central registry **SHOULD** assign a qualified person to evaluate data and use his or her judgment to determine whether or not any deviations from these standards or norms are accurate.

2.3.3.1. Standards

Data **SHOULD** be analyzed for the following patterns:

- Incidence rates and frequencies **SHOULD** be greater than mortality rates and frequencies. If mortality exceeds incidence for cancer of any site, the data for that site **MUST** be verified.

- Lung, liver, and pancreas are typical sites for DCO cases. Investigation is required if there are no DCO cases for these sites.

- Rates of cancer of the corpus uteri are higher for whites than for African-Americans. Generally, rates for cancers of the cervix uteri are higher for African-Americans and Hispanics than for whites. The exception seems to be rural whites in Kentucky.

- Significant numbers of melanomas occur only in the white population.

- In cases where the age distribution of the population at risk is similar to the national norm, childhood cancers (ages 0-14) account for 1 percent of the total number of cases, and cancers in persons over 80 years of age account for 10 to 15 percent of the total. As the population ages, it is expected that cancer in persons over 80 years of age will increase.

- Hispanics have lower rates of all cancers except those of the cervix uteri, esophagus, stomach, and pancreas.

- African-Americans have higher rates of prostate cancer than any other race.

- Microscopically confirmed cases account for approximately 93 percent of all cases in the United States. About 5 percent are clinically diagnosed, and about 1 to 1.5 percent are DCO cases. In Canada, the percent of cases microscopically confirmed is between 80 to 95 percent. Percentages may vary due to differences in the reporting sources used by provincial/territorial/central cancer registries to ascertain cases.

- The primary site of the cancer is unknown for about 5 percent of all cases.

2.3.4. Timeliness of Central Registry Reporting

Timely reporting of tumor information is an important goal for a central registry. Epidemiology, cancer control, and clinical users benefit from speedy access to the most current information. However, completeness and accuracy of data also are essential goals. Reports based on incomplete or inaccurate data can misinform scientists and the public about the true picture of cancer in the central registry's area.

The speed with which central registry data can be collected, processed, analyzed, and reported depends on many factors, some of which are within the registry's control and others that are not. Historically, abstracting began 6 months after the cancer diagnosis because treatment was usually complete within 6 months. Now, treatment can extend well beyond 6 months for some cancers, but the demand for current cancer information requires more timely data collection. Efficient data collection methods, computer and software training, telecommunications, and well-trained staff all can influence the timeliness of reporting of tumor records from facilities, within limits. Many facilities are capable of concurrent reporting and can complete abstracts in "real time." Electronic path reporting has expedited case identification and the abstracting process for some central cancer registries.

Transmission of cases from a reporting facility to the central cancer registry also impacts the timeliness of reporting. Many central cancer registries have their own standards for data transmission. Some central registries require the facility to transmit weekly or monthly, and other central registries require a facility to transmit data for every 100 cases abstracted.

Once tumor records have been received by the central registry, a wide variety of activities take place, as outlined in Chapters 3, 4, and 5. All of these processing steps take time, and some—notably death clearance, sharing of tumor records with other central registries, and establishment of population denominators—impose external delays on the registry. Central registries need sufficient staff and processes to ensure timely abstracting.

2.3.4.1. Standards

- *American College of Surgeons:* Cases **MUST** be abstracted within 6 months of date of first contact (COC Cancer Program Standards 3.3).

- *Surveillance, Epidemiology, and End Results Program:* The registry is under contract to provide complete counts of new cases for a calendar year within 22 months after the calendar year ends.

- *Centers for Disease Control and Prevention/National Program of Cancer Registries:* Within 12 months of the close of the diagnosis year, 90 percent of expected, unduplicated cases are available to be counted as incident cases at the central cancer registry; and, within 24 months of the close of the diagnosis year, 95 percent of expected, unduplicated cases are available to be counted as incident cases at the central cancer registry.

- *North American Association of Central Cancer Registries:* Within 23 months of the close of a diagnosis year, the registry **SHOULD** contain at least 95 percent of the expected cases of reportable cancer occurring in residents during that year.

- *Canadian Cancer Registry:* As of March 1, 2003, the annual Canadian Cancer Registry Call for Data (new and updated cases) deadline is 14 months turn-around time from year-end (i.e., 2001 data due March 1, 2003).

2.3.5. Casefinding Audit Results

Casefinding audits are studies involving independent reascertainment of tumor records, usually in a sample of facilities and, within each facility, a sample of time periods. Tumor records identified during the audit are enumerated and matched against the central registry's files. Unmatched cases are followed back to verify their reportability, and the percent of cases actually missed that should have been reported is calculated.

Studies are designed for a variety of purposes and with varying degrees of statistical rigor. Most studies focus on hospital reporting and thus provide an estimate of the completeness of reporting for hospitals only, not a true central registry completeness estimate. The following sources are problematic to review in a systematic way and usually have not been incorporated into audit protocols:

- Physician offices.

- Clinics and outpatient facilities, including radiation therapy centers and surgery treatment centers.

- Freestanding and out-of-state pathology laboratories.

- Hospices.

- Facilities outside the coverage area of a central registry but treating residents from the registry's area.

Well-designed protocols with careful sampling plans and formal analysis plans are important when calculating an estimate of the central registry's completeness that will be made public or used to assess registry completeness. If the goal is to identify possible ascertainment problems in facilities and to take corrective action, more informal methods **MAY** be appropriate; however, there are other advantages to a formal well-documented protocol and written findings. It will allow repetition of the study at a later time or in another area or group of facilities, and findings can be compared over time and across samples if the same study design is used and results are well documented.

2.3.5.1. Standards

Standards have not been established for the design of casefinding studies or the statistical analysis of the results. However, it is important that a statistician or epidemiologist familiar with central cancer registries as well as sampling methods design such studies. In 2003, NAACCR conducted a Best Practices Workshop on Casefinding Audits. It was determined that casefinding audits were an important function of central registries to:

- Evaluate completeness of case ascertainment for an individual reporting facility and/or for the central cancer registry.
- Evaluate data reliability.
- Use outcomes to identify training issues.
- Identify strengths and deficiencies in reporting facility casefinding procedures.
- Establish estimated case counts for reporting facilities.
- Identify specific underreported primary sites.
- Evaluate timeliness of case submission.
- Measure change in casefinding and data submission processes.

To indicate possible baseline values in studies of this type, completeness of casefinding studies carried out by the SEER Program and by NPCR are presented below:

> In 2002, the SEER Program conducted casefinding studies of reportable cases diagnosed in 2000 in the four SEER expansion registries. The cancer sites audited were: breast, bladder, bronchus/lung, colon/rectum, and prostate. All of the five primary sites audited were selected because they are among the most frequent sites reported by cancer registries, accounting for 60 percent of the cases in the SEER database over the past 5 years. Prostate was chosen because it also often is diagnosed and treated in non-hospital settings and, therefore, more often is missed. Pathology and cytology reports were the source documents for the audit. Limiting this audit to the most ordinary sites and source documents achieved the goal of performing the audit as efficiently and cost-effectively as possible.
>
> As reported by the registries, there were 100,962 new incidence cases for the hospitals audited. The auditors reviewed 7,110 cases, 7 percent of eligible cases. The audit identified a total of 417 missed and late cases out of 7,110 cases reviewed (5.9 percent missed). The

missed/late case rates were set up to be self-weighting. The individual registry missed-case rates ranged from a high of 11.5 percent to a low of 1.0 percent.

The 2004 casefinding audit included all 14 SEER registries, and all cancer sites were audited. The audit took place between June and October 2004. Results are not yet available for the 2004 study.

NPCR performs case completeness and data quality audits in central registries to assess the level of completeness and data quality. The NPCR has an audit protocol and work plan written specifically for the central registry that outlines all procedures to be performed. The central registries are to be audited once in each 5-year grant period.

2.3.6. Follow-Up Success Rates

Different formulas are used to calculate the percent successful follow-up. They vary by whether deceased individuals are included in the numerator and/or denominator and whether the month of follow-up is considered or only the year of follow-up. Any standard established **MUST** specify the formula to be used. The NAACCR Best Practice Workshop held in 2003 recommended that central cancer registries follow the SEER method to calculate follow-up. SEER conducts an annual follow-up calculation for its registries. The Best Practice recommendation was that central cancer registries evaluate follow-up rates at least quarterly. Some central cancer registries found that a monthly evaluation of follow-up was beneficial.

For the population-based registry's purpose of calculating patient survival based on accumulated follow-up data, it is crucial that the percent of cases successfully followed be as high as possible and that the cases lost to follow-up are an unbiased group.

2.3.6.1. Standards

Two national organizations, SEER and the ACoS, have established standards for follow-up rates for the participants.

- *SEER:* The SEER Program includes a standard for follow-up success rates in the scope of work for contracts with its participating registries. The requirement is for a success rate of at least 90 percent, preferably 95 percent or greater overall, and there are separate requirements by age grouping. The SEER formula for calculating successful follow-up, applied separately to invasive and *in situ* cancers (excluding cervix *in situ*), is as follows:

  Assume that Y is the last year of data submitted. The percent of patients diagnosed during the years prior to and who have current follow-up is defined as:

  $P = 100(D+A)/T$

  D is the number dead prior to January 1, Y +1, A is the number follow-up dates on or after January 1, (Y+1) (includes alive and dead), and T is the total number of patients being followed. P can be calculated for individual years of diagnosis up through Y-1 and for all years combined prior to Y.

Age-specific requirements are:	Age < 20	at least 90 percent but must not be below 80 percent
	Age 20-64	at least 90 percent but must not be below 80 percent
	Age 65+	at least 95 percent but must not be below 90 percent
	All ages	at least 95 percent but must not be below 90 percent

SEER does not require follow-up of *in situ* cancers of the cervix uteri.

- *ACoS Commission on Cancer:* See Commission on Cancer Program Standards 2004, Standards 3.4 (an 80 percent follow-up rate is maintained for all analytic patients from the cancer registry reference date) and 3.5 (a 90 percent follow-up rate is maintained for all analytic patients diagnosed within the last 5 years, or from the cancer registry reference date, whichever is shorter). Long-term follow-up is essential to evaluate outcomes of cancer care. Accurate follow-up data enable facilities to compare outcomes with regional, state, or national statistics. Follow-up information is obtained at least annually for all living analytic patients included in the cancer registry database.

The registry **SHOULD** apply the calculations to subgroups of patients to evaluate for bias. For example, calculation of follow-up rates by sex for three age groups, those under 15, 15 to 64, and those 65 and over, **MAY** show that, although the overall rate is very high, the registry is not successfully following its pediatric cancers, especially among females. An analysis by ethnic group or geographic area might identify other groups with poor follow-up.

Chapter 3: Data Quality

3.1. STRUCTURAL REQUIREMENTS

3.1.1. General Requirements

Every feature of the central registry's operations can impact data quality. Examples include:

- Laws and regulations under which the central registry operates.
- Relationships with hospitals, outpatient facilities, and physicians in the registry's coverage area.
- Data collection system design and capabilities.
- Qualifications and training of central registry staff.
- Review of data for analysis and reporting.

Quality control encompasses the personnel and activities that focus on the assessment and improvement of data quality.

3.1.1.1. Standards

The registry **MUST** have a quality assurance program with specified activities integrated into basic central registry operations. Definitions for the quality assurance program **SHOULD** include:

- Assignment of a qualified individual to perform quality control activities.
- Schedule for routine edits and reports.
- Steps to be taken when specified conditions are not met.

The central registry **SHOULD** carefully document each of these activities. Documentation should include procedural changes as well as any non-routine dataset evaluation(s) that are undertaken.

The central registry's budget **MUST** specify, and adequately fund, quality control staff and activities.

Based on quality control activity results, procedures **SHOULD** identify how further action will be taken for those areas requiring improvement. Guidelines to monitor the status of follow-up completion also should be provided.

3.1.2. Staffing Guidelines for Data Quality

Central registry staffing **MUST** consist of personnel who have adequate skills to conduct registry business in a timely, competent manner. Registry staff should be skilled in the following areas:

- Knowledge of quality control activities.
- Data evaluation and analysis (including statistics and sampling).

- Training and professional development.
- Organizational and communication skills.

3.1.2.1. Standards

Adequate quality control activities in the central registry **MUST** include:

- *Certified Tumor Registrar(s) (CTR):* One or more CTRs **MUST** be involved in monitoring abstract review, training staff who abstract or edit data (central registry employees and staff at reporting facilities), and conducting quality control activities. CTRs provide expertise in the diagnosis and treatment of tumors, casefinding procedures, and follow-up.
- *Abstractors and Coders:* Whether the central registry employs abstractors or coders in the office (for abstracting or routine editing) or in the field, they **MUST** know the data definitions and coding instructions used by the central registry. They need to be well trained in abstracting tumor data from patient records.
- *Quality Control:* The central registry **SHOULD** identify one person to manage and maintain quality control activities. This person often will be responsible for training central registry and hospital staff who collect registry data. This person **SHOULD** have primary responsibility for the interpretation of quality control audit results.
- *Computer Expertise:* The central registry **MUST** have knowledgeable information technology (IT) staff available to assist in the design and implementation of edits and special studies.
- *A Statistical Analyst:* An individual who knows and understands cancer surveillance **MUST** design and evaluate output for routine data analysis and special studies. Expertise is needed in sampling techniques and the application of appropriate statistical measures. The person **MUST** know and understand the criteria for undertaking remedial action. The person **MUST** be familiar with statistics, evaluation tools, and/or epidemiology.

3.1.3. Procedure Manuals, Coding Manuals, and Other Documentation

To establish standards, maintain continuity, and document changes over time, the central registry **MUST** maintain complete documentation that reflects both current and historical practices. The documentation **SHOULD** incorporate all aspects of the central registry's operations including its definitions and methods. Documentation is most often found in procedure manuals, coding manuals, and other manuals specific to registry operations.

3.1.3.1. Standards

The central registry **MUST** provide adequate staff and time to prepare and maintain high-quality, up-to-date documentation or manuals.

The registry **MUST** document:

- Dataset and data definitions.
- Codes.

- Coding rule interpretations and procedures.

- Decisions or recommendations of its medical advisors.

The central registry **MUST** have a mechanism for updating and maintaining currency of documentation. To promote data comparability the registry **MUST** incorporate or reference material utilized from any standard setters (e.g., SEER, ACoS, NPCR, and NAACCR).

Documentation **MUST** be provided to all central registry employees involved in data collection, management, and analysis, including employees of the hospitals and facilities that report data to the registry. Appropriate sections of the documentation that explain definitions and methods **SHOULD** be provided to investigators and users of the data.

Documentation can be in the form of printed material, including data dictionaries, coding manuals, and procedure manuals. Online electronic documentation is available and standardized for some resources. The central registry's documentation **MAY** be in printed form, online, or in a combination of media as long as it meets the needs of the local reporting facilities.

3.1.4. Edits and Data Processing Capabilities for Data Quality

A computer program may better perform certain repetitive manual processes. Over the years, cancer registry software has been developed to address an increasing number of registry tasks, enabling staff to focus on activities requiring human judgment, analysis, or interaction. In most cases where technology use has increased, data quality improves.

Central registry computer software systems **MUST** provide a repository for data and the tools to generate incidence reports, research data, or other registry end products. It also is a major focal point for quality control processes. One basic function of central registry software is maintenance of data integrity. Careful and effective data management and the implementation of adequate system security accomplish this task. These functions are covered in Chapter 5 of this document (Data Management). The present section covers design characteristics of the computer system that directly relate to quality control activities of the central registry. Routine quality control functions that **SHOULD** be built into a central registry's computer system include:

- *Edits:* Data edits are logical rules, typically embodied in a computer algorithm, that evaluate to "true," "false," or "maybe," for any value(s) of (a) data item(s). Central registry edits are applied to all records to check for item validity, internal consistency, and inter-record consistency. Data edits may involve a single field, multiple fields in a single record, multiple fields in different records within one database, or multiple fields in multiple databases (see Sections 5.1.4. and 5.8.).

- *Process Controls:* Statistical process control involves the prospective monitoring of rationally aggregated results of inspection. Process controls can involve errors in abstracts (or batches) that are detected (e.g., edit rejection rates) as well as other aspects of central registry data and operation that do not necessarily represent errors, but that should exhibit stability over time or across regions (e.g., percent unknown primaries). Process control design requires statistical expertise, including specification of an appropriate probability model, selection of a sampling plan and rational subgroups, selection of appropriate control charting procedures, and specification of control limits.

  Retained information from edit procedures **SHOULD** be analyzed on a regular basis to identify area(s) for improvement (i.e., data sources, coders, item code structure, or clarity of instructions in the manuals). The computer system **SHOULD** contain flags set to reflect the nature and disposition of

edit failures and include analytic routines for evaluating their contents. The data are summarized across time for individual data sources or item codes. Items **SHOULD** include the date each tumor record was accessioned into the registry and the date the tumor record was updated so that delays between case reporting and accession can be evaluated.

- *Capabilities for Special Studies:* The system **SHOULD** be able to draw appropriate samples, enable efficient data entry for tumor records from the field, produce automated comparisons of original and reabstracted or recoded data, and analyze results to support audits.

3.1.4.1. Standardized Edits

Data edited differently may vary systematically, lending to non-comparability. Edits need to be standardized across all registries for the following reasons:

- The utility of local data is compromised when data categorization is not comparable.
- A standard edit contributes to comparable data.
- Errors in primary editing steps cannot be fixed by subsequent edits.

Standards for edits are discussed in Section 5.8.1. of this document and are included in the electronic NAACCR edits metafile that can be downloaded from the NAACCR website (www.naaccr.org).

3.1.4.2. Required Components

The following components generally will be required for automated quality control procedures:

- *Computer Edits:* The central registry **MUST** have a system of computerized data edits with the following characteristics (see Sections 5.1.4. and 5.8.):
  - Standard program code or algorithm wherever possible.
  - Single-field, multi-field, multi-record, and multi-database edits as appropriate.
  - Flexibility for change.
  - Production of reports and error messages that are meaningful to those correcting errors and to everyone that interprets data.
  - Documentation and/or tables about the logic and performance, which are available and understandable to those who either correct errors or use the data.
  - Provisions for edit output that **MAY** be returned to individual facilities for resolution.

- *Process Controls:* The central registry **SHOULD** provide process controls. The data items necessary to identify and store quality measures and the analytic routines for systematically evaluating them **SHOULD** be built into the computer system.

  Retained information from edit procedures **SHOULD** be analyzed on a regular basis to identify area(s) for improvement (i.e., data sources, coders, item code structure, or clarity of instructions in the manuals). The computer system **SHOULD** contain flags set to reflect the nature and disposition of edit failures and include analytic routines for evaluating their contents. The data are summarized across time for individual data sources or item codes. Items **SHOULD** include the date each tumor record was accessioned into the registry and the date the tumor record was updated so that delays between case reporting and accession can be evaluated.

- *Audits:* The central registry system **SHOULD** allow drawing of samples for quality control studies by any desired characteristic.

- *Staff:* The central registry **MUST** have staff trained in abstracting and coding to track and correct edit failures (see Section 3.1.2.).

3.1.4.3. Standards for Data Entry, Data Meaning, Data Representation, Datasets, and Record Layout

3.1.4.3.1. Standardization of Data Entry

Accepted output is facilitated by the standardization of as many of the required steps for data collection and processing as possible. Standardization of the following registry software application features may improve data comparability:

- Prompts.

- Coding choice lists.

- Online help.

- Edits: single-field, multi-field, multi-record, or multi-database.

- Error messages.

Auto-coding is convenient but can be risky, especially for histology variables for which modifiers to a root word change the histology code.

Central registries will vary in the extent of control they have over developing standardization. Some registries obtain data collected by hospitals that use a variety of software applications. However, central registries **SHOULD** take the following steps to encourage standardization:

- Adopt existing data standards, including those in NAACCR's various standard and operational documents for cancer registries (see *NAACCR Procedure Guidelines for Cancer Registries*).

- Encourage mechanisms for the definition and publication of additional standards. These include communication with other central registries; work with NAACCR committees; and, communication with standard-setting organizations.

3.1.4.3.2. Standardization of Code Definitions

Trend analysis depends on a historical continuity in data definitions. In some cases when categories are discontinued, continuity may be preserved by maintaining the collection of the old categories while collection of the new categories begins. When additional detail is desired, ensure that standard categories are feasible when data definitions are combined and/or collapsed.

3.1.4.3.3. Standard Datasets

Central registries **SHOULD** collect data items to meet appropriate regulations (e.g. state, provincial/ territorial, federal). NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary,* provides data collection requirements of each standard-setting organization. Other data items

collected by a central registry **MAY** be identified by local constituencies for specific cancer control purposes, such as:

- Patient care evaluation.
- Descriptive epidemiology and surveillance.
- Research.
- Incidence.
- Outcomes and survival.

3.1.4.3.4. Standardization of Data Exchange Format

Standardization of the electronic format for data exchange improves the quality of merged files. This includes specification of:

- Data and data translations codes.
- Item sequence and record layout.
- Electronic media specifications.

NAACCR's recommended exchange format is presented in the *Standards for Cancer Registries Volume I: Data Exchange Standards and Record Description,* and *Volume II: Data Standards and Data Dictionary.*

3.1.4.4. Standards for Frequency and Timing of Data Edits

Edits **SHOULD** be run at the reporting source prior to central registry submission, which facilitates immediate verification/review of edit failures. This improves the success of obtaining accurate clarification, minimizes permanent information loss, and increases the value of the data.

Item, internal consistency, and inter-record edits **SHOULD** be applied routinely before new records are added to the database. Serious edit failures **SHOULD** be withheld from incorporation into the analytic database until they are resolved.

Continuous analysis of edit failures **SHOULD** be performed. Changes in staff, reporting facilities, vendors, new procedures or other data-collection conditions that are not stabilized require special attention (see Section 5.8.).

Information on EDITS is located on the NAACCR website Registration Standards page (www.naaccr.org).

3.1.4.5. Standards for Record Consolidation

Record consolidation is an important function of central cancer registries. It ensures that all submitted tumor records are counted only once. When records are not consolidated, over-counting of cancer incidence occurs. The NAACCR Record Consolidation Committee published record consolidation guidelines in the *Central Cancer Registry Record Consolidation: Principles and Processes* documents and published two reports, the

Report of the Record Consolidation Committee, 1999 (available at www.naaccr.org) and *Creation of a Record Consolidation Test File: Report to the NAACCR Board* (Springfield, IL, 2003).

3.2. PROCESS STANDARDS

3.2.1. Standards for Data Codes

Any central registry that collects a data item that has a national standard **SHOULD** use standard codes. It is very difficult to combine or compare data with other registries when different codes are employed. Central registries that use a different set of codes for an item **SHOULD:**

- Completely map codes to standard codes. Central registry codes **MAY** provide more detail, but **SHOULD NOT** provide less detail.
- Export data only after they have been fully converted to standard codes.
- Receive and process data from other registries in the standard codes.

3.2.2. Standards for Data Text

To perform quality control review of coded data, abstracted text summaries from the medical record **SHOULD** be reviewed. Text information **SHOULD** be included in the registry's dataset in computerized form along with the data codes to facilitate quality control. See NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary* for text field definitions and a recommended abbreviations list.

Text information **SHOULD** be transmitted along with codes when tumor records are shared with other registries (see Section 2.2.5.)

3.2.3. Standards for Data Edits

Standard data fields **SHOULD** be edited using the appropriate edit standards for that field. The EDITS metafile contains standard edits for each of the standard-setting organizations. These edits **SHOULD** be used at several levels, but at a minimum, before loading case information in a central registry and prior to release of data. Central registries **SHOULD** require reporting facilities to use the EDITS metafile prior to data submission.

Computer systems under development **SHOULD** be designed with the expectation of incorporating the EDITS metafile as a standard.

3.2.4. Training for Improved Data Quality

Training is an essential component for a population-based registry to ensure data collection is accurate, consistent, and complete (see Section 2.2.10.).

3.2.4.1. Required Components

Training **MUST** be provided to the central registry staff involved in data collection and quality control and to the staff of facilities that are reporting data to the registry. Training activities in the following areas are recommended:

- *Reporting Requirements:* Instruction on reporting requirements, including frequency of reporting, mechanism of reporting, and required data items. Documentation **MUST** be provided that defines the reporting requirements.

- *Data Collection:* Instruction on reportable neoplasms, casefinding procedures, abstracting requirements, ICD-O coding, staging, and, where appropriate, treatment coding **MUST** be provided. The instruction **MUST** be based on the standardized reference manuals that the central registry officially adopts.

- *Quality Control:* Instruction in visual and computer edits and feedback regarding edit results **SHOULD** be provided to the data collection staff and other staff from reporting facilities.

- *Data Processing:* Instruction regarding the use of computer software **SHOULD** be provided.

3.2.4.2. Standards for Training Methods

A variety of methods **MAY** be utilized, including:

- Satellite and land-based video conferences with beginning and advanced training and educational workshops.

- Formal programs with beginner and advanced training classes, workshops, educational programs, and symposia, plus regularly scheduled in-service training.

- Audits to identify areas that need additional training.

- Feedback to data collectors on the types and patterns of errors identified during quality control activities.

- Site visits to evaluate and train at a data collection site or central registry.

- NAACCR CDs on core central registry analysis.

- Training and educational media.

- Web-based training modules.

Faculty **SHOULD** include physicians, CTRs, epidemiologists, statisticians, and computer experts.

The central registry **SHOULD** obtain approval of its workshops for formal continuing education credits for CTRs. Contact NCRA for more information (www.ncra-usa.org).

The central registry **SHOULD** use standardized training materials provided by standard-setting organizations. COC, NAACCR, NPCR, and SEER provide training and education resources on their websites (see Appendix E).

3.2.5. Quality Control Activities

Although it is appropriate and necessary to design a quality control program to fit the needs of a particular central registry and its users, certain quality control activities will be universally applicable, such as:

- *Process Control:* Statistical process control involves the prospective monitoring of rationally aggregated results of inspection. Process controls can involve where errors in abstracts (or batches) that are detected (e.g., edit rejection rates) as well as other aspects of central registry data and operation that do not necessarily represent errors, but that should exhibit stability over time or across regions (e.g., percent unknown primaries). Process control design requires statistical expertise, including specification of an appropriate probability model, selection of a sampling plan and rational subgroups, selection of appropriate control charting procedures, and specification of control limits.

- *Special Assessments:* Central registries **SHOULD** perform special assessments to evaluate registry-specific issues (e.g., data item inconsistencies on changes in reporting sources) and to address special requests for review of specific data. Special assessments that can be standardized **SHOULD** be executed on a routine basis to enhance data quality.

- *Reabstracting Audits:* Reabstracting audits describe the process of independently reabstracting tumor records from the source patient records, coding the data, and comparing the abstracted and coded data to the data already in the registry. This type of study historically has been used in central registries, and the methods are well developed.

- *Recoding Audits:* Recoding audits involve independently reassigning codes to abstracted text information but not reviewing the source documents. This type of study is conducted frequently, and is very useful in training new coders; it is easier and less expensive to perform than reabstracting, but the method cannot detect problems with abstracting.

- *Reliability Studies:* Reliability studies are designed to test participants' understanding and adherence to coding rules and practices. This is the only study that can evaluate the overall performance of coders and abstractors. The participants code from identical source documents under controlled conditions. When the coding phase of the study is complete, the coders and abstractors can work with experts to reconcile answers. The final results can be statistically represented by comparing the results to accuracy goals for each data item.

3.2.5.1. Standards for Process Controls

Process controls represent an additional level of sophistication, in which the aggregated results of inspection are tracked, usually over time, and used to determine objectively whether or not a process is "within normal limits." Design of statistical process controls requires the specification of a sampling plan, selection of rational subgroups, computation of control limits, selection of a charting strategy (if control charts will be used), and specification of frequency of updates. These issues as well as actions to be taken **SHOULD** be fully documented. Measures of central registry quality that should benefit from formal development of process controls include, but are not limited to, the following:

- Visual review rejection rates.

- Duplicate entry/recoding/reabstracting rejection rates.

- Edit check failure rates—overall and/or failure on the most important data items.

- Missing data and use of unknown or ill-defined codes for data items considered critical to analysis by the central registry.

- Number of tumor records submitted.
- Lag time in reporting.
- Percent DCO.
- Reabstracting agreement rates.

Automated support for process controls is strongly recommended. For example, the computer can assist in the acquisition, management, and charting of process control data, and these functions can be built into central registry software systems (see Section 5.6.).

3.2.5.2. Standards for Special Assessments

Central registries **SHOULD** periodically plan and execute casefinding audits to assess overall completeness of reporting and reabstracting audits to assess overall data reliability (see Sections 2.2.12. and 2.3.5. for discussions of casefinding audits). Reabstracting and recoding studies have a long history in central registries. The methodologies are well defined, and comparison data may be available (see Section 3.3.1.). Additional assessments **MAY** be undertaken to address specific tumors, problem areas, or feasibility of proposed changes. All special assessments **SHOULD** be planned and executed according to a formal, written protocol including the following:

- Introduction and rationale.
- Statement of purpose.
- Sampling plan, including sample size considerations, stratifications, and randomization.
- Eligibility criteria and study population.
- Procedures to be followed for study execution.
- Analysis plan, including data management, statistical analysis, and summary statistics to be computed.

Completed studies **SHOULD** be analyzed and the results communicated to management, data suppliers, and data users. Central registries **SHOULD** address training needs indicated by results.

3.2.6. Dissemination of Quality Control Activity Results

Identifying and correcting data errors is required to maintain quality data. In addition to correcting errors, it is essential that feedback be given to the data abstractor so that the quality of data will be maintained and recurring errors eliminated.

3.2.6.1. Standards

To reduce the number of data errors and avoid recurring problems, feedback **MUST** be provided in a timely manner.

When abstracts are corrected or changed at the central registry, information about the changes **SHOULD** be returned to the abstractor for review. Discrepancy reports or error reports from edits also **MAY** be returned.

The central registry **SHOULD** provide results of recoding audits, casefinding audits, and reabstracting audits with analysis of discrepancies and recommendations for improvement to abstractors. Feedback on findings of audit studies and interpretation of the results **SHOULD** be given to all who participate in a study as well as the pool of individuals or organizations represented by the study participants.

The feedback **SHOULD** identify problems and recommend actions that could be undertaken to correct problems and improve data quality. Feedback may be given through telephone calls or one-on-one meetings. Summary audit study results also **SHOULD** be made available to data users to assist in the interpretation of the data.

The central registry **SHOULD** incorporate the results of quality control activities as feedback to other aspects of registry functioning. For example, the central registry **SHOULD:**

- Interpret the results of quality monitoring, and incorporate the conclusions when revising training materials, documentation, or item definition as needed.

- Provide useful evaluative data, so that data users have an adequate context for interpreting their results.

3.3. OUTCOME MEASURES

3.3.1. Reabstracting and Recoding Audits

Reabstracting audits and recoding audits are often used to retrospectively assess accuracy (agreement with source medical records), validity (produce desired results), and reproducibility (agreement among data collectors) of registry data. Audits are studies on a sample of cases and **MUST** be done in accordance with a study protocol that states the study objectives, describes the sampling scheme, and outlines plans for the analysis. These studies have a long history in cancer registries, and the methodologies are well developed.

Three sampling designs that are applicable to both casefinding and reabstracting are:

- *Random Sample:* A sample of size "n" from the population chosen in such a way that every set of "n" individuals has equal chance to be in the sample actually selected. A random number table, found as a reference in statistics handbooks, may be used to randomly assign numbers to each facility.

- *Stratified Sample:* A sample in which the population first is divided into groups of similar individuals, called strata, and then a simple random sample is chosen from each stratum and combined to form the full sample. For example, hospitals are grouped by geographic location and then randomly selected from each group.

- *Multi-Stage Sample Design:* A sample drawn in stages using probability sampling methods. This method is more appropriate for states with large numbers of facilities, or for multi-state audits. For example, hospitals are grouped by geographic location. The total number of groups needed is selected. The desired number of groups is selected using a random number table, and from the groups selected, the desired number of hospitals is selected using a random number table.

The objective of a reabstracting study is to characterize the level of agreement between data in the registry and data reabstracted and recoded from source records (the hospital medical records for most cases) by expert auditors. For each reabstracted data item, the auditor's codes are compared to the original codes to identify discrepancies. If the codes do not match, the discrepancy is classified as to severity according to major and minor discrepancy definitions set up in advance for the specific study (see Appendix H for sample major-minor definitions used by SEER). Such studies require an arbitration or reconciliation mechanism to determine which of the discrepant answers is correct for purpose of the study.

Recoding audits help to characterize the level of agreement within data records already in the registry. Expert auditors use the text contained in the abstract to recode a sample of actual case abstracts in the registry database. As in a reabstracting study, for each recoded case, codes for each data item are compared for discrepancies with those assigned by the expert.

3.3.1.1. Study Results

The registry can learn a variety of things from reabstracting and recoding audits, including:

- Overall and item-specific agreement rates for the sample of cases studied, which **SHOULD** be expressed in terms of severity (see Appendix H).
- Types of tumor records in which discrepancies occur more frequently.
- Sources of variation (e.g., misinterpretation of source document information, information not available at initial abstracting, misinterpretation of coding rules, inadequate or erroneous computer consolidation of data between records). However, when it is not possible to identify the source of variation, additional data collection may be needed.
- Effect of misclassifications on data analysis and use (e.g., are tumors more frequently over-staged or under-staged?).
- Data quality with respect to other factors such as the age of the registry, who collects the data (hospital registrars versus non-registrars versus central registry), training and skills of the registrars collecting the data, and difficulty of abstracting and coding the specific data items.

Where indicated, this information **SHOULD** be used to identify training needs and to modify registry processes and procedures to ensure future improvement in data quality.

3.3.1.2. General Standards

Target rates for data quality **SHOULD** be established and the performance of the central registry and individual reporting facilities should be measured using the target rates. Target agreement rates will vary from one data item to another, depending on the impact that data item has on incidence, rates, the complexity and detail of the coding scheme, and the quality of medical record information upon which coded information is based.

3.3.1.3. Standards for Reabstracting Studies

There are no national standards for agreement rates from reabstracting studies, but some central registries have set standards for their reporting facilities. NAACCR has not set standards for reabstracting. SEER set goals for the 2000 and 2001 reabstracting studies. These goals are compared to the actual scores achieved

during reabstracting with the intent of establishing benchmarks for reabstracting agreement rates. For a complete discussion of reabstracting studies, see the CDC/NPCR, and NAACCR Educational CD, Audits: Casefinding and Reabstracting.

A SEER registry's performance in meeting or exceeding the goals established by SEER is measured using star graphs. The error rate for a specific item is defined as the number of errors divided by the number of possibilities for making the error (i.e., the number of cases) within one registry.

The stars are assigned by a mathematical calculation using the registry's error rate for that data item and the SEER goal for that data item. If the SEER goal was 95 percent, a registry would receive five stars if they met or exceeded the 95 percent goal, four stars if they achieved 94.9-85.5 percent accuracy, three stars if they scored between 85.4 and 76 percent, two stars if they scored between 75.9 and 66.5 percent, and one star if they scored under 66.5 percent.

Central cancer registries **SHOULD** check the standards of their national program.

3.3.1.4. Standards for Recoding Studies

Recoding studies usually are based on tumor abstract source documents and therefore remove abstracting differences as a possible source of code variation. Consequently, higher agreement rates are expected from recoding studies than from reabstracting studies.

Recoding studies do not measure the accuracy of the coding with respect to the medical record; they measure the accuracy of coding as function of the quality of the text justification submitted with the abstract. Poor performance on a recoding audit indicates a need for training on how to write informative text, in addition to training on how to code medical information.

3.3.2. Abstracting and Coding Reliability Studies

In contrast to reabstracting and recoding audits described above in which data already in the registry are compared with those collected by an expert auditor in cancer registration, reliability studies involve the abstracting and coding of a set of actual cases by abstractors or coders. Reliability studies measure abstractors' and coders' compliance with established coding rules and standards. These studies include a reconciliation process that provides a measure of agreement between the abstractors and coders.

The reliability study measures the quality of the abstracting/coding process in terms of reproducibility under special circumstances. Results from this study method help identify ambiguity or inadequacy of existing data definitions and rules, and areas that require further registrar education and training. This method also is useful for testing whether new codes should be implemented as defined, and the degree to which there is likely to be consistency in coding.

Two primary advantages of the reliability study are: (1) ease of comparing individual coders or groups of coders to some standard, and (2) relative simplicity and adaptability of the approach.

3.3.2.1. Standards

Kappa statistics measure agreement between reviewers. In quality control studies, the kappa statistic is a measurement to assess the proportion of agreement beyond chance among two or more reviewers on specific data items. The maximum value of the kappa statistic is +1 if there is exact and complete agreement between the reviewers, and a minimum of -1 if there is not. For most targets, values greater than 0.75 represent excellent agreement beyond chance. Values below 0.40 represent poor agreement beyond chance. Values

between 0.40 and 0.75 represent fair-to-good agreement beyond chance (Fleiss, J.L. [1981] *Statistical Methods for Rates and Proportions*, Second edition, John Wiley & Sons, New York).

3.3.3. Unknown Values

The proportion of tumors with unknown values for various data items can be an indicator of data quality. Unknown values can result from problems with:

- Data collection system or access to necessary source documents.
- Item and code values that are defined.
- Misapplication of coding rules.

However, unknown values also can accurately reflect a limited workup or ambiguity in the medical record. A high proportion of unknown values for a data item may indicate that the item cannot be collected as defined, and that it may be appropriate to drop the item from the dataset. Modification of the definitions may decrease the proportion of unknown codes. The proportion of unknown values usually varies by primary site.

3.3.3.1. Standards

For a specific data item related to a specific primary site, the percent coded unknown **SHOULD** be evaluated according to how analysis will be affected. Will incidence rates be affected, or will survival rates? Will misleading conclusions from the data be possible because of the high percent of unknown values? Depending on the analysis being performed, the percent unknown may be more or less problematic. For example, will the percentage of cases of melanoma with unknown race result in the rate of melanoma for all races combined being higher than the rate for whites? The NAACCR Registry Certification Committee has established minimum standards for percent unknown for four variables (see Table 1).

Table 1. NAACCR Criteria and Standards for Gold/Silver Certification

Criterion	Gold Standard	Gold Error Tolerance	Silver Standard	Silver Error Tolerance
1. Completeness	≥ 95%	-1.0	≥ 90%	-1.0
2. Passing Edits	100%	0	≥ 97%	-0.4
3. DCOs	≤ 3%	0.4	≤ 5%	0.4
4. Timeliness	Within 23 months		Within 23 months	
5. Duplicate Records	≤ 1/1,000	0.4	≤ 2/1,000	0.4
6. Missing Data Fields – Sex, Age, County	≤ 2%	0.4	≤ 3%	0.4
7. Race	≤ 3%	0.4	≤ 5%	0.4

Data completeness benchmarks from the SEER Program are available for more variables, and these are summarized in Table 2.

Table 2. Actual Percent Unknown for Selected Data Items – SEER Program

Data Item	SEER Public Use Data File 1973-2000 (2000 cases only), N = 174,622
Race = 99	1.26
Birthplace = 999	51.49
Marital Status = 9	5.71
Sequence Number = 99	0.00
Primary Site = C80.9	1.91
Histologic Type = 8000 or 8001	1.72
Diagnostic Confirmation = 9	1.43
Surgery = 09	1.06
Radiation = 8 or 9	2.52
Summary Stage*	
Breast	1.71 (N = 31,701)
Colon and Rectum	4.67 (N = 19,167)
Lung	7.83 (N = 20,310)

* Data calculated using SEER Historic Stage.
Source: SEER Program, 200l.

Chapter 4: Data Analysis and Reporting

4.1. STRUCTURAL REQUIREMENTS

4.1.1. Confidentiality Policies and Procedures: Issues in Research, Reporting, and Release of Registry Data

Confidentiality is the cancer registry's responsibility to the patients in the database and is of paramount concern to all cancer registries. There may be no greater threat to the operation and maintenance of a cancer registry than an actual or perceived breach of confidentiality. In fact, an actual or perceived breach of confidentiality in one registry may threaten all registries.

This section reviews the elements of a comprehensive confidentiality policy that relates to research uses, reporting, and release of cancer data. See Section 2.1.1.5. for standards for confidentiality provisions of laws and regulations, Section 2.2.8. for a discussion of confidentiality issues in data collection and management, Section 4.2.4. for a discussion of confidentiality issues and public use data files, and Appendix I for the NAACCR Policy Statement 99-01: Confidentiality.

Maintaining patient confidentiality while collecting and using high-quality data presents significant challenges. The Inventory of Best Practices Assurance of Confidentiality and Security was developed as an inventory of best practices (see Appendix J). Registries **SHOULD** use the Inventory to prioritize their needs and action steps to improve security and data confidentiality protection processes (see the *NAACCR 2002 Workshop Report on Data Security and Confidentiality*, available at www.naaccr.org).

The Privacy Rule of the U.S. Health Insurance Portability and Accountability Act governs the use and disclosure of some health-related information. This federal law clearly defines the "covered entities" to which the Privacy Rule applies. They are: (1) a health plan, (2) a health care clearinghouse, or (3) a health care provider.

Because U.S. central cancer registries do not perform any of these functions, they are not covered entities and the HIPAA Privacy Rule does not govern their activities nor the information that they hold, including the release of registry data. A document of HIPAA frequently asked questions (FAQ) has been developed for registries and can be found on the NAACCR website: www.naaccr.org. The FAQ document is updated as necessary to reflect ongoing interpretations and revision to these rules.

4.1.1.1. Definition of Confidential Data

Although the tumor reporting laws and regulations under which the central registry operates may define only patient-specific data as confidential, registries **SHOULD** consider any information that specifically identifies a health care professional or an institution as confidential. Information that characterizes the caseload of a specific institution or health care professional also **SHOULD** be considered proprietary and confidential.

Other information may be used to identify individuals or institutions through indirect means. For example:

- A report may inadvertently provide enough non-confidential information to identify a specific individual. Consider a report that indicates that a prostate cancer was diagnosed in a 65 year-old African American male in a geographic area whose residents are primarily of Asian ancestry. Even though no confidential information is released, this information might allow someone with knowledge of the geographic area to identify the patient.

- Characterizing cases diagnosed in a geographic region whose health care is provided by a single physician or institution may inadvertently provide confidential information about the caseload of the health care professional or facility.

- Combinations of variables such as postal code or census tract plus birth date and sex may be sufficient to specifically identify an individual.

- Linkage of external files with non-confidential registry data (e.g., registry data with identifiers deleted), whether authorized or not, may enable re-identification of individuals.

4.1.1.2. Standards for Laws and Regulations Governing Confidentiality

Laws and regulations pertaining to confidentiality of tumor data vary by geopolitical location. The central registry **SHOULD** contact legal counsel to determine which rules govern the registry's area of coverage. The relevant laws may include those stipulating governmental access to documents, covering privacy, covering medical records, and preventing release of confidential data for any legal proceedings. Cancer registries operating within provincial/territorial/state/federal governments or agencies will be subject to laws and regulations pertaining to the government's collection, use, and release of information.

4.1.1.3. Standards for Policies and Procedures for Release of Confidential Data

- Confidential information about data subjects or data suppliers **MUST NOT** be released for purposes other than those specified by the central registry.

- Confidential information **MAY** be released to health care providers and institutions directly involved in the care of the patient, for example:
  - A hospital cancer registrar requests a list of all prostate cancer patients who have been treated at his or her facility.
  - A physician requests a list of patients he or she has treated for breast cancer.

- Central cancer registries **SHOULD** abide with their specific law or regulations that may have specific procedures for release of an individual's data to that individual.

- Confidential information **MUST NOT** under any circumstances be published or made available to the general public.

- Inquiries from the press/media **SHOULD** be referred to the delegated authority that can fully respond to these communications. For example, press requests often have to be referred to a public information spokesperson prior to a referral directly to the central cancer registry.

- Measures **MUST** be taken to eliminate the possibility that individuals might be identifiable from tables containing cells with very small figures/counts (see the example provided in Section 4.1.1.1.).

- Central registries **MUST** provide a document describing their procedures and criteria for release of registry data to researchers who request access to data.

4.1.1.3.1. Inappropriate Uses of Confidential Information

If the central registry is located within a governmental agency such as a health department, the registry **MUST** develop clear policies regarding access to data by other sections or programs of the department. Access by other programs could jeopardize confidentiality and may be inappropriate.

Confidential cancer registry data **MUST NEVER** be made available for uses such as the following:

- Businesses that are trying to market a product to cancer patients.
- Health care institutions that are trying to recruit new patients.
- Insurance companies that are trying to determine the medical status of a patient.
- Next-of-kin of reported patients.

4.1.1.4. Standards for Suppressing Non-Confidential Data for Summary Statistics

Reports of summary statistics generally do not raise confidentiality concerns. However, confidential information can be conveyed inadvertently through summary statistics. To avoid this situation, the central cancer registry **SHOULD** institute a policy to suppress the publication of summary statistics in instances when data are being presented for geographic areas with small populations. For example, some registries suppress the reporting of statistical data when there are fewer than six (this number varies) cases reported in a single cell of a table if the cell of the table represents a combination of variables, such as sub-state or sub-provincial geographic area, race, age, and sex, that could inadvertently identify individuals. Some jurisdictions use denominator rules basing them on the size/count of the population. However, for straightforward breakdowns by age, sex, and large geographic areas, cells with 0, 1, or a few cases normally need not be suppressed.

Confidentiality concerns include the capability of identifying a patient from the data file as well as the potential to gain new information about a patient on the file or to re-identify a patient through linkage of the registry file with other electronic files. NAACCR has developed the Record Uniqueness Program to test data files for the potential of patient identifiability. This tool evaluates the data file for identifiably and potential re-identifiably, because most of the confidential data items are not released to researchers or on public use files. The Record Uniqueness Program is available for download on the NAACCR website (www.naaccr.org). It contains complete instructions for use and interpretation.

4.1.1.5. Standards for Use of Registry Data for Research

4.1.1.5.1. Release of Confidential Data to Scientific Investigators

Requests for central cancer registry data for research often can be satisfied through provision of a public use data file of non-confidential data (see Section 4.2.4.). When non-confidential data are not sufficient to answer the question, the central registry **MUST** determine who is and is not qualified to use cancer registry data for research purposes. The central cancer registry may suffer if it allows its data to be used for inappropriate purposes. The central registry **MUST** develop an application for researchers to apply for use of confidential facts. In addition, the registry **MUST** develop a set of guidelines to govern the accessibility of cancer registry data to independent scientific investigators. Registry data **SHOULD** be made available for scientific research only after the following criteria have been met.

- Requests for registry data to be used for research **MUST** be in writing and include a suitable detailed outline of the proposed research and a justification of any need for confidential data. The central registry is responsible for ensuring that researchers do not receive more data than are needed to answer the research question.

- Appropriate central registry staff **MUST** review the written research plan. Requests for data **MUST** meet the registry's guidelines on confidentiality. The central registry **MUST** determine that the research needs could not adequately be addressed with non-confidential information.

- The central registry **MUST** have access to an IRB (U.S.) or ethics committee (Canada).

- An appropriate IRB or ethics committee **SHOULD** approve the proposed research. The investigator **SHOULD** provide evidence that all appropriate IRBs or ethics committees have approved the research.

- The Principal Investigator **MUST** sign a written agreement to adhere to all confidentiality policies. Written agreements **MUST** include provisions for use of the information and for its return or destruction at the end of the study.

- The scientific objectives of the study **SHOULD** be peer reviewed to ensure scientific validity.

- The registry **MUST** obtain evidence that researchers using cancer registry data will adhere to the central registry's guidelines on confidentiality.

NAACCR developed an IRB and established IRB guidelines to review all projects that are NAACCR sponsored or that use NAACCR data files that were prepared from the aggregation of registries' data through the annual call for data or through special studies. The NAACCR IRB does not review studies that do not fall directly under the purview of the NAACCR IRB (see www.naaccr.org).

4.1.1.5.2. Review of Research Results

Once the central registry has granted an investigator access to confidential information for purposes of scientific research, the registry **MUST** ensure that confidential information is not, under any circumstances, published or displayed in reports that summarize the research results. The central registry **SHOULD** retain the right to review any reports prior to their dissemination to ensure that confidentiality has been respected.

4.1.1.5.3. Patient Contact for Participation in Epidemiologic Studies

Central cancer registries can identify cancer patients as potential subjects for the epidemiologic studies. In these instances, the investigators **MUST** meet all the criteria outlined above. Philosophies differ as to whether physician permission is needed prior to patient contact. Many patient advocacy groups maintain that only a patient has the right to decide study participation and his/her physician does not have the right to make that choice on the patient's behalf. Consequently, in many current epidemiologic studies, the physician is contacted to inform him/her that the patient will be contacted to participate in a study and to ask whether there are any contraindications to patient contact (e.g., patient too ill, patient unaware of diagnosis, etc.). Many investigators feel that this procedure protects the physician from any risk of adverse action on the part of the patient. Other investigators still insist on physician permission before contacting the patient. Furthermore, local IRBs also may insist on physician permission as a condition of study approval.

4.1.2. Population Data

Producing estimates of the number of persons in the population at risk covered by the central registry, stratified by year, age, sex, race, and geographic units, is a fundamental function of a population-based registry. The jurisdiction under which the registry operates may apply various constraints on population counts that are to be used. For example, a central cancer registry in a health department may be required to use official population estimates approved by its local government and officially approved race or ethnic categories that are used for all other government programs in that geographic area.

4.1.2.1. General Requirements

The amount of detail the central registry will need to know about the population will vary, depending on the type of rates that are to be calculated. Crude rates can be calculated with an estimate of the size of the total population living within the registry's coverage area. However, crude rates are not useful for comparative analyses, because age is strongly related to the risk of cancer. Knowledge of the age distribution of the population is required to calculate both age-specific and age-adjusted incidence rates. Often, incidence rates are calculated for specific population sectors, such as sex and race, which requires population counts for each of these factors.

4.1.2.2. Standards for Sources of Population Estimates

The central registry **MUST** identify the most appropriate sources of available population data for its area. The U.S. Census Bureau is the most common source of population data in the United States. The Census Bureau conducts decennial censuses. A Canadian census is conducted every 5 years by Statistics Canada. Both organizations regularly produce estimates for censal, postcensal, intercensal, and projected populations (see Sections 4.2.1.2. and 4.2.2.3.5.).

State/provincial, territorial, and local governmental agencies often are a good source for additional information about the size and characteristics of a population in the central registry's area. Some agencies or jurisdictions employ demographers that can be a source of expertise to the registry.

The jurisdiction of the central registry may require the use of official population estimates or official race and ethnic categories. The registry **MUST** develop relationships with appropriate agencies and become aware of such requirements.

4.1.2.3. Standards for Ethnic, Racial, and Other Population Groups

Cancer rates vary by ethnic and racial groups in the United States. For this reason, it is useful to calculate incidence rates separately for ethnic and racial groups within the central registry's coverage area. Of primary concern when calculating ethnic and race-specific rates is the comparability of definitions between the numerator (i.e., tumor records) and the denominator (i.e., population estimates). Specifically, the methods that are used to define a person's race or ethnicity in the numerator of the rate **SHOULD** be as comparable as possible to those used in the denominator. Unfortunately, it can be difficult to obtain appropriate estimates of the size of the population for individual years by age, race, ethnicity, and geography. When calculating rates by ethnicity and race, the registry **MUST** carefully document the methods by which race and ethnicity were assigned, both in the numerators and the denominators.

For example, attempts to identify individuals of Hispanic/Latino ethnicity have been based on numerous methods, including self designation, surname, country of birth, and use of the Spanish language. However, estimates on the size of the Latino population from the Census are based on self identification. Some groups use various approaches to enhance these counts based on knowledge of reported undercounts of the

population in question. In some instances, the method of Latino identification in the numerator and the choice of denominators could have an effect on the accuracy of cancer incidence rates.

NAACCR members have addressed the need to enhance cancer information for race and ethnic populations. One result of this effort is the development and application of a standard approach, the NAACCR Hispanic Identification Algorithm (NHIA), to enhance the identification of Hispanic/Latino persons with cancer. Employing this standardized approach makes it possible to combine cancer statistics more reliably (see *Report of the NAACCR Expert Panel on Hispanic Identification 2003*, available at www.naaccr.org) from multiple registries and to conduct meaningful comparisons among them.

In 2004, the NAACCR Asian/Pacific Islander Identification Panel convened to develop a standardized approach to enhance the identification of Asian/Pacific Islander populations.

Another example involves obtaining appropriate population estimates for U.S. Native Americans. Population estimates for Native groups are often available from both tribal and non-tribal sources. When using these data, the central registry **MUST** be careful to distinguish between a complete tribal census, which may enumerate all members of a tribe regardless of geographic area of residence, and an enumeration of tribal members who live within a defined geographic area. For reporting purposes, the central cancer registry is most often interested in the population that resides within a defined geographic area.

SEER has developed new guidelines to reduce the lack of consistency in interpolating races from the variables such as birthplace or geographic homogeneity. Race definitions and classifications in the *SEER Program Code Manual* are used by the Census Bureau and adhere to the October 30, 1997 Federal Register Notice entitled, "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity" issued by the Office of Management and Budget.

Other populations that may warrant special consideration with regard to denominator ascertainment include active duty military personnel, institutionalized individuals (such as prisoners and hospital patients), part-time residents, undocumented workers, and homeless or other non-permanent residents.

4.1.2.4. Standards for Interpretation of Population Estimates

It is the responsibility of the central registry staff to understand how the population estimates were derived, their limitations, and any potential impact on cancer rates. The registry staff **MUST** consult with local experts, especially demographers and members or representatives of special populations, to assure that the registry is collecting racial and ethnic data in a manner that is consistent with population data. Furthermore, the central registry **MUST** work with these experts to assure that the data are reported in as accurate and sensitive a manner as is possible.

4.1.3. Staffing Guidelines for Data Analysis and Reporting

The appropriate analysis, interpretation, use, and dissemination of cancer data are primary functions of the central cancer registry. The registry **MUST** identify staff members and consultants who are qualified to conduct and interpret appropriate analyses of registry data.

4.1.3.1. Standards for Number and Type of Staff

The central registry **MUST** have access to expertise to conduct appropriate analyses and interpret results. This includes experts from the fields of oncology, pathology, public health, epidemiology, statistics, and demography, and also may include computer programmers. The experts may be full-time or part-time, and they may be members of the registry staff or consultants.

Data analysis staff and consultants **MUST** work closely with the central registry's quality control and data management staff to ensure quality data are produced and available for analysis. When appropriate, registry staff **SHOULD** conduct orientation sessions for expert consultants to ensure that they have adequate knowledge of registry operations and procedures.

When it is not possible for a central cancer registry to retain a staff member for the sole purpose of data analysis and interpretation, the registry may wish to develop the analysis skills of abstractors or other staff members so that they may assist consultants in the preparation of reports. Special training programs in epidemiology and statistics are available to meet these needs, such as those conducted by NAACCR (e.g., the Cancer Surveillance Institute I [CSI I] and Cancer Surveillance Institute II [CSI II], the Toolkit, and the Advanced Course; for more information visit the NAACCR website at www.naaccr.org).

Each central cancer registry **SHOULD** designate one or more staff members to serve as a liaison between the public and the central registry. By centralizing the responsibility for these interactions, the registry cuts down on possible duplications of effort. This practice also minimizes the opportunity for misunderstandings that occur when information is obtained from multiple sources.

4.1.3.2. Standards for Continuing Education

Staff involved in data analysis and reporting **SHOULD** be offered opportunities for and encouraged to pursue continuing education so that they remain informed about analysis methods and trends in cancer data.

4.1.3.2.1. Continuing Education

Continuing education **SHOULD** be provided to data analysis staff to assure that they have up-to-date knowledge about trends in cancer incidence, diagnosis, management, treatment, outcomes, and survival; statistical and epidemiological methods; demographic trends and methods; computer capabilities and other technologies; and cancer registries.

4.1.3.2.2. Access to Professional Literature, Online Services, and Other Activities

Data analysis staff **MUST** be supplied with appropriate references and literature to provide ongoing continuing education and to answer questions that arise. Current pertinent reference books and journals **MUST** be immediately available. The central registry **SHOULD** provide access to online services and bulletin board services so that staff have rapid access to the most current information.

4.1.3.2.3. Professional Associations and User Groups

Central registry staff **MUST** be encouraged and funded to participate in local and national professional associations and user groups. The registry budget **MUST** include funds for participation by one or more persons at scheduled meetings. The registry **MUST** fund data analysis staff to attend scientific meetings, special symposia, conferences, and courses that may occur from time to time.

4.2. PROCESS STANDARDS

4.2.1. Analysis Categories and Recoded Groups

Many data items in a cancer registry are collected using precise code categories so that these data will be useful for many purposes. Primary site, histologic type, age, race and ethnicity, and extent of disease all are examples. When the precision is not a factor for certain analyses, the registry **SHOULD** use standardized groupings of detailed codes.

The selection of standard categories for analysis and presentation **MAY** depend on the choice and/or availability of comparison data. Although conventional standards do exist, the choice of methods depends on many factors, including the number of tumor records available for study, the availability of comparison data, and the needs of the investigator. For example, central cancer registries that want to compare their incidence data with those of the SEER Program will need to conform to the methods by which SEER data were derived. Some investigators may need to develop special categories of data that are not routinely published. For example, the incidence rates for specific histologic types of cancer are not always published in routine reports, and investigators may have difficulty obtaining comparison data on them. Nonetheless, the cancer registry **SHOULD** be flexible to accommodate these investigators on an *ad hoc* basis.

The SEER*Stat statistical software provides a convenient mechanism for the analysis of SEER and other cancer-related databases. It is a powerful computer-based tool to view individual cancer records and to produce descriptive statistics for studying cancer in a population (see www.seer.cancer.gov).

The SEER*Prep software converts ASCII text data fields to the SEER*Stat database format, allowing registries to analyze cancer data using SEER*Stat. SEER*Prep performs two main functions: (1) it converts text data to the specific binary format required by SEER*Stat, and (2) it creates the SEER*Stat data dictionary (see www.seer.cancer.gov).

4.2.1.1. Standards for Grouping by Primary Site and Histologic Type

Tumor records are commonly grouped by a combination of primary site and histologic type. A standard grouping used by the SEER Program is presented in Tables 3 and 4. Table 3 is a recoding scheme for tumors coded in ICD-O-2, and Table 4 is a recoding scheme for tumors coded in ICD-O-3. Each table provides for two levels of detail-specific sites and grouped sites. The primary categorization is by site, but some histologic types are given categories. For example, extranodal lymphomas are reported with lymphomas in this scheme rather than with their primary sites. The SEER Program makes the recode available on request as a computer program that assigns each tumor to its appropriate recoded group, coded using the values in the third column.

Registries **SHOULD** use the SEER recoding scheme of cancer site categories for routine analyses.

Another important standard is the grouping used by WHO in its *Cancer Incidence in Five Continents*. This grouping is based on the ICD-9 classification system rather than ICD-O. Registries **SHOULD** use this set of categories for international comparisons, especially when ICD-O categories are not available (the SEER Program will provide, on request, a conversion program and documentation converting ICD-O-2 to ICD-9).

The etiology of pediatric cancers is different from adults in that cell type is more important than the organ site, and thus there is a different set of cancer categories. The standard is the International Classification of Childhood Cancer 1996, shown in Table 5. It is based on ICD-O-2, but also includes some non-malignant diagnoses and some categories from SNOMED for non-neoplastic conditions. Registries **SHOULD** use this set of categories for comparing data on pediatric cancers.

For cancer mortality data, the diagnoses are classified using ICD rather than ICD-O. The analysis categories used by the National Center for Health Statistics (NCHS) in its mortality statistics do not correspond to the categories used by registries for cancer incidence. SEER provides the recode shown in Table 6 for ICD cancer mortality diagnosis categories comparable to the incidence categories in Tables 3 and 4. Registries **SHOULD** use this recode when cancer incidence and mortality are being compared for specific sites.

Table 3. Standard Site Analysis Categories With ICD-O-2 Codes SEER Site Recode ICD-O-2 (1/27/2003) Definition

Site Group	ICD-O-2 Site	ICD-O-2 Histology (Type)	Recode
Oral Cavity and Pharynx			
Lip	C000-C009	Excluding 9590-9989, and sometimes 9050-9055, 9140+	20010
Tongue	C019-C029		20020
Salivary Gland	C079-C089		20030
Floor of Mouth	C040-C049		20040
Gum and Other Mouth	C030-C039, C050-C059, C060-C069		20050
Nasopharynx	C110-C119		20060
Tonsil	C090-C099		20070
Oropharynx	C100-C109		20080
Hypopharynx	C129, C130-C139		20090
Other Oral Cavity and Pharynx	C140, C142-C148		20100
Digestive System			
Esophagus	C150-C159	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21010
Stomach	C160-C169		21020
Small Intestine	C170-C179		21030
Colon and Rectum			
Colon excluding Rectum			
Cecum	C180	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21041
Appendix	C181		21042
Ascending Colon	C182		21043
Hepatic Flexure	C183		21044
Transverse Colon	C184		21045
Splenic Flexure	C185		21046
Descending Colon	C186		21047
Sigmoid Colon	C187		21048
Large Intestine, NOS	C188-C189, C260		21049
Rectum and Rectosigmoid Junction			
Rectosigmoid Junction	C199	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21051
Rectum	C209		21052
Anus, Anal Canal and Anorectum	C210-C212, C218		21060
Liver and Intrahepatic Bile Duct			
Liver	C220	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21071
Intrahepatic Bile Duct	C221		21072
Gallbladder	C239		21080

Site Group	ICD-O-2 Site	ICD-O-2 Histology (Type)	Recode
Other Biliary	C240-C249		21090
Pancreas	C250-C259		21100
Retroperitoneum	C480		21110
Peritoneum, Omentum, and Mesentery	C481-C482		21120
Other Digestive Organs	C268-C269, C488		21130
Respiratory System			
Nose, Nasal Cavity, and Middle Ear	C300-C301, C310-C319	Excluding 9590-9989, and sometimes 9050-9055, 9140+	22010
Larynx	C320-C329		22020
Lung and Bronchus	C340-C349		22030
Pleura	C384		22050
Trachea, Mediastinum, and Other Respiratory Organs	C339, C381-C383, C388, C390, C398, C399		22060
Bones and Joints	C400-C419	Excluding 9590-9989, and sometimes 9050-9055, 9140+	23000
Soft Tissue Including Heart	C380, C470-C479, C490-C499	Excluding 9590-9989, and sometimes 9050-9055, 9140+	24000
Skin Excluding Basal and Squamous			
Melanoma of the Skin	C440-C449	8720-8790	25010
Other Non-Epithelial Skin	C440-C449	Excluding 8000-8004, 8010-8045, 8050-8082, 8090-8110, 8720-8790, 9590-9989, and sometimes 9050-9055, 9140+	25020
Breast	C500-C509	Excluding 9590-9989, and sometimes 9050-9055, 9140+	26000
Female Genital System			
Cervix Uteri	C530-C539	Excluding 9590-9989, and sometimes 9050-9055, 9140+	27010
Corpus and Uterus, NOS			
Corpus Uteri	C540-C549	Excluding 9590-9989, and sometimes 9050-9055, 9140+	27020
Uterus, NOS	C559		27030
Ovary	C569		27040
Vagina	C529		27050
Vulva	C510-C519		27060
Other Female Genital Organs	C570-C589		27070
Male Genital System			
Prostate	C619	Excluding 9590-9989, and sometimes 9050-9055, 9140+	28010
Testis	C620-C629		28020

Site Group	ICD-O-2 Site	ICD-O-2 Histology (Type)	Recode
Penis	C600-C609		28030
Other Male Genital Organs	C630-C639		28040
Urinary System			
Urinary Bladder	C670-C679	Excluding 9590-9989, and sometimes 9050-9055, 9140+	29010
Kidney and Renal Pelvis	C649, C659		29020
Ureter	C669		29030
Other Urinary Organs	C680-C689		29040
Eye and Orbit	C690-C699	Excluding 9590-9989, and sometimes 9050-9055, 9140+	30000
Brain and Other Nervous System			
Brain	C710-C719	Excluding 9530-9539, 9590-9989, and sometimes 9050-9055, 9140+	31010
Cranial Nerves Other Nervous System	C710-C719	9530-9539	31040
	C700-C709, C720-C729	Excluding 9590-9989, and sometimes 9050-9055, 9140+	
Endocrine System			
Thyroid	C739	Excluding 9590-9989, and sometimes 9050-9055, 9140+	32010
Other Endocrine Including Thymus	C379, C740-C749, C750-C759		32020
Lymphoma			
Hodgkin Lymphoma			
Hodgkin - Nodal	C024, C098-C099, C111, C142, C379, C422, C770-C779	9650-9667	33011
Hodgkin - Extranodal	All other sites		33012
Non-Hodgkin Lymphoma			
NHL - Nodal	C024, C098-C099, C111, C142, C379, C422, C770-C779	9590-9595, 9670-9677, 9680-9688, 9690-9698, 9700-9717, 9823, 9827	33041
NHL - Extranodal	All sites except C024, C098-C099, C111, C142, C379, C422, C770-C779	9590-9595, 9670-9677, 9680-9688, 9690-9698, 9700-9717	33042

Site Group	ICD-O-2 Site	ICD-O-2 Histology (Type)	Recode
	All sites except C024, C098-C099, C111, C142, C379, C420-C422, C424, C770-C779	9823, 9827	
Myeloma		9731-9732	34000
Leukemia			
Lymphocytic Leukemia			
Acute Lymphocytic Leukemia		9821, 9826, 9828	35011
Chronic Lymphocytic Leukemia	C420, C421, C424	9823	35012
Other Lymphocytic Leukemia		9820, 9822, 9824-9825, 9850, 9940-9941	35013
Myeloid and Monocytic Leukemia			
Acute Myeloid Leukemia		9840-9841, 9861, 9866, 9867, 9871-9874, 9910	35021
Acute Monocytic Leukemia		9891	35031
Chronic Myeloid Leukemia		9863, 9868	35022
Other Myeloid/Monocytic Leukemia		9860, 9862, 9864, 9880, 9890, 9892-9894, 9930	35023
Other Leukemia			
Other Acute Leukemia		9801, 9931, 9932	35041
Aleukemic, Subleukemic, and NOS		9800, 9802-9804, 9830, 9842, 9870, 9900	35043
	C420, C421, C424	9827	
Mesothelioma*		9050-9055	36010
Kaposi Sarcoma*		9140	36020
Miscellaneous		9720-9723, 9740-9741, 9760-9768, 9950, 9960-9962, 9970, 9980-9984, 9989	37000
	C760-C768, C809	Excluding 9590-9989, and sometimes 9050-9055, 9140+	
	C420-C424		
	C770-C779		
Invalid	Site or histology code not within valid range or site code not found in this table.		99999

* The Site Recode variable can be created with or without Mesothelioma (9050-9055) and Kaposi Sarcoma (9140) as separate groupings. The table above documents both possibilities.

Source: SEER 2003.

Table 4. Standard Site Analysis Categories With ICD-O-3 Codes

SEER Site Recode ICD-O-3 (1/27/2003) Definition

Site Group	ICD-O-3 Site	ICD-O-3 Histology (Type)	Recode
Oral Cavity and Pharynx			
Lip	C000-C009	Excluding 9590-9989, and sometimes 9050-9055, 9140+	20010
Tongue	C019-C029		20020
Salivary Gland	C079-C089		20030
Floor of Mouth	C040-C049		20040
Gum and Other Mouth	C030-C039, C050-C059, C060-C069		20050
Nasopharynx	C110-C119		20060
Tonsil	C090-C099		20070
Oropharynx	C100-C109		20080
Hypopharynx	C129, C130-C139		20090
Other Oral Cavity and Pharynx	C140, C142-C148		20100
Digestive System			
Esophagus	C150-C159	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21010
Stomach	C160-C169		21020
Small Intestine	C170-C179		21030
Colon and Rectum			
Colon excluding Rectum			
Cecum	C180	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21041
Appendix	C181		21042
Ascending Colon	C182		21043
Hepatic Flexure	C183		21044
Transverse Colon	C184		21045
Splenic Flexure	C185		21046
Descending Colon	C186		21047
Sigmoid Colon	C187		21048
Large Intestine, NOS	C188-C189, C260		21049
Rectum and Rectosigmoid Junction			
Rectosigmoid Junction	C199	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21051
Rectum	C209		21052
Anus, Anal Canal, and Anorectum	C210-C212, C218		21060
Liver and Intrahepatic Bile Duct			

Site Group	ICD-O-3 Site	ICD-O-3 Histology (Type)	Recode
Liver	C220	Excluding 9590-9989, and sometimes 9050-9055, 9140+	21071
Intrahepatic Bile Duct	C221		21072
Gallbladder	C239		21080
Other Biliary	C240-C249		21090
Pancreas	C250-C259		21100
Retroperitoneum	C480		21110
Peritoneum, Omentum, and Mesentery	C481-C482		21120
Other Digestive Organs	C268-C269, C488		21130
Respiratory System			
Nose, Nasal Cavity, and Middle Ear	C300-C301, C310-C319	Excluding 9590-9989, and sometimes 9050-9055, 9140+	22010
Larynx	C320-C329		22020
Lung and Bronchus	C340-C349		22030
Pleura	C384		22050
Trachea, Mediastinum, and Other Respiratory Organs	C339, C381-C383, C388, C390, C398, C399		22060
Bones and Joints	C400-C419	Excluding 9590-9989, and sometimes 9050-9055, 9140+	23000
Soft Tissue Including Heart	C380, C470-C479, C490-C499	Excluding 9590-9989, and sometimes 9050-9055, 9140+	24000
Skin Excluding Basal and Squamous			
Melanoma of the Skin	C440-C449	8720-8790	25010
Other Non-Epithelial Skin	C440-C449	Excluding 8000-8005, 8010-8045, 8050-8084, 8090-8110, 8720-8790, 9590-9989, and sometimes 9050-9055, 9140+	25020
Breast	C500-C509	Excluding 9590-9989, and sometimes 9050-9055, 9140+	26000
Female Genital System			
Cervix Uteri	C530-C539	Excluding 9590-9989, and sometimes 9050-9055, 9140+	27010
Corpus and Uterus, NOS			
Corpus Uteri	C540-C549	Excluding 9590-9989, and sometimes 9050-9055, 9140+	27020
Uterus, NOS	C559		27030
Ovary	C569		27040
Vagina	C529		27050
Vulva	C510-C519		27060
Other Female Genital Organs	C570-C589		27070
Male Genital System			

Site Group	ICD-O-3 Site	ICD-O-3 Histology (Type)	Recode
Prostate	C619	Excluding 9590-9989, and sometimes 9050-9055, 9140+	28010
Testis	C620-C629		28020
Penis	C600-C609		28030
Other Male Genital Organs	C630-C639		28040
Urinary System			
Urinary Bladder	C670-C679	Excluding 9590-9989, and sometimes 9050-9055, 9140+	29010
Kidney and Renal Pelvis	C649, C659		29020
Ureter	C669		29030
Other Urinary Organs	C680-C689		29040
Eye and Orbit	C690-C699	Excluding 9590-9989, and sometimes 9050-9055, 9140+	30000
Brain and Other Nervous System			
Brain	C710-C719	Excluding 9530-9539, 9590-9989, and sometimes 9050-9055, 9140+	31010
Cranial Nerves Other Nervous System	C710-C719	9530-9539	31040
	C700-C709, C720-C729	Excluding 9590-9989, and sometimes 9050-9055, 9140+	
Endocrine System			
Thyroid	C739	Excluding 9590-9989, and sometimes 9050-9055, 9140+	32010
Other Endocrine Including Thymus	C379, C740-C749, C750-C759		32020
Lymphoma			
Hodgkin Lymphoma			
Hodgkin - Nodal	C024, C098-C099, C111, C142, C379, C422, C770-C779	9650-9667	33011
Hodgkin - Extranodal	All other sites		33012
Non-Hodgkin Lymphoma			
NHL - Nodal	C024, C098,C099, C111,C142, C379,C422, C770-C779	9590-9596, 9670-9671, 9673, 9675, 9678-9680, 9684, 9687, 9689-9691, 9695, 9698-9702, 9705, 9708-9709, 9714-9719, 9727-9729, 9823, 9827	33041
NHL - Extranodal	All sites except C024, C098-C099, C111, C142, C379, C422, C770-C779	9590-9596, 9670-9671, 9673, 9675, 9678-9680, 9684, 9687, 9689-9691, 9695, 9698-9702, 9705, 9708-9709, 9714-9719, 9727-9729	33042
	All sites except C024, C098-C099, C111, C142, C379, C420-C422, C424, C770-C779	9823, 9827	
Myeloma		9731-9732, 9734	34000

Site Group	ICD-O-3 Site	ICD-O-3 Histology (Type)	Recode
Leukemia			
Lymphocytic Leukemia			
Acute Lymphocytic Leukemia		9826,9835-9837	35011
Chronic Lymphocytic Leukemia	C420, C421, C424	9823	35012
Other Lymphocytic Leukemia		9820, 9832-9834, 9940	35013
Myeloid and Monocytic Leukemia			
Acute Myeloid Leukemia		9840, 9861, 9866, 9867, 9871-9874, 9895-9897, 9910, 9920	35021
Acute Monocytic Leukemia		9891	35031
Chronic Myeloid Leukemia		9863, 9875, 9876, 9945, 9946	35022
Other Myeloid/Monocytic Leukemia		9860, 9930	35023
Other Leukemia			
Other Acute Leukemia		9801, 9805, 9931	35041
Aleukemic, Subleukemic, and NOS		9733, 9742, 9800, 9831, 9870, 9948, 9963, 9964	35043
	C420, C421, C424	9827	
Mesothelioma*		9050-9055	36010
Kaposi Sarcoma*		9140	36020
Miscellaneous		9740-9741, 9750-9758, 9760-9769, 9950, 9960-9962, 9970, 9975, 9980, 9982-9987, 9989	37000
	C760-C768, C809	Excluding 9590-9989, and sometimes 9050-9055, 9140+	
	C420-C424		
	C770-C779		
Invalid	Site or histology code not within valid range or site code not found in this table.		99999

* The Site Recode variable can be created with or without Mesothelioma (9050-9055) and Kaposi Sarcoma (9140) as separate groupings. The table above documents both possibilities.
Source: SEER 2003.

Table 5. Standard Site/Histology Analysis Categories for Pediatric Cancers With ICD-O-2 Codes

It should be noted that in this table, the ICD-O-2 site-code refers to the site of the **primary** tumor. The presence of a behavior code /6 implies that the histological diagnosis (M-code) is based on the biopsy of a metastasis; nevertheless, the associated site code refers to the known (or suspected) primary site from which the metastasis derived.

ICCC From IARC Technical Report No. 29

Site Group	ICD-O-2 Histology (Type)	ICD-O-2 Site	Recode
I Leukemia			
(a) Lymphoid Leukemia			
Acute Lymphocytic Leukemia (ALL)	9821, 9828	C000-C809	012
Lymphoid Excluding ALL	9820, 9822-9827, 9850	C000-C809	011
(b) Acute Non-Lymphocytic Leukemia	9840, 9841, 9861, 9864, 9866, 9867, 9871-9874, 9891, 9894, 9910	C000-C809	013
(c) Chronic Myeloid Leukemia	9863, 9868	C000-C809	015
(d) Other Specified Leukemia	9830, 9842, 9860, 9862, 9870,9875-9890, 9892, 9893, 9900, 9930-9941	C000-C809	016
(e) Unspecified Leukemia	9800-9804	C000-C809	017
II Lymphoma and Reticuloendothelial Neoplasms			
(a) Hodgkin Lymphoma	9650-9667	C000-C809	021
(b) Non-Hodgkin Lymphoma	9591-9595, 9670-9686, 9688, 9690-9717, 9723	C000-C809	022
(c) Burkitt Lymphoma	9687	C000-C809	023
(d) Miscellaneous Lymphoreticular Neoplasms	9720, 9731-9764	C000-C809	024
(e) Unspecified Lymphoma	9590	C000-C809	025
III CNS and Miscellaneous Intracranial and Intraspinal Neoplasms			
(a) Ependymoma	9383, 9390-9394	C000-C809	031
(b) Astrocytoma	9380	C723	032
	9381, 9400-9441*	C000-C809*	032
(c) Primitive Neuroectodermal Tumors	9470-9473*	C000-C809*	033
(d) Other Gliomas	9380	C700-C722, C724-C729	034
	9382, 9384	C000-C809	034
	9442-9460, 9481	C000-C809	034
(e) Miscellaneous Intracranial and Intraspinal Neoplasms	8270-8281, 8300, 9350-9362, 9480, 9505, 9530-9539*	C000-C809*	035

Site Group	ICD-O-2 Histology (Type)	ICD-O-2 Site	Recode
(f) Unspecified Intracranial and Intraspinal Neoplasms	8000-8004	C700-C729, C751-C753	036
IV Sympathetic Nervous System Tumors			
(a) Neuroblastoma and Ganglioneuroblastoma	9490, 9500*	C000-C809*	041
(b) Other Sympathetic Nervous System Tumors	8680, 8693-8710, 9501-9504, 9520-9523	C000-C809	042
V Retinoblastoma	9510-9512	C000-C809	051
VI Renal tumors			
(a) Wilms Tumor, Rhabdoid, and Clear Cell Sarcoma	8963	C649, C809	061
	8960, 8964	C000-C809	061
(b) Renal Carcinoma	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8143, 8155, 8190-8201, 8210, 8211, 8221-8231,8240, 8241, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8401, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8573	C649	062
	8312	C000-C809	062
(c) Unspecified Malignant Renal Tumors	8000-8004	C649	063
VII Hepatic Tumors			
(a) Hepatoblastoma	8970	C000-C809	071
(b) Hepatic Carcinoma	8010-8041, 8050-8075, 8082, 8120-8122, 8140, 8141, 8143, 8155, 8190-8201, 8210, 8211, 8230, 8231,8240, 8241, 8244-8246, 8260-8263, 8310, 8320, 8323, 8401, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8573	C220, C221	072
	8160-8180	C000-C809	072
(c) Unspecified Malignant Hepatic Tumors	8000-8004	C220, C221	073
VIII Malignant Bone Tumors			
(a) Osteosarcoma	9180-9200	C000-C809	081
(b) Chrondosarcoma	9220-9230	C000-C809	082
	9231, 9240	C400-C419	082
(c) Ewing Sarcoma	9260	C400-C419, C809	083
	9363, 9364	C400-C419	083
(d) Other Specified Malignant Bone Tumors	8812, 9250, 9261-9330, 9370	C000-C809	084†

Site Group	ICD-O-2 Histology (Type)	ICD-O-2 Site	Recode
(e) Unspecified Malignant Bone Tumors	8000-8004, 8800, 8801, 8803, 8804	C400-C419	085†
IX Soft-Tissue Sarcomas			
(a) Rhabdomyosarcoma and Embryonal Sarcoma	8900-8920, 8991	C000-C809	091
(b) Fibrosarcoma, Neurofibrosarcoma, and Other Fibromatous Neoplasms	8810, 8811, 8813-8833, 9540-9561	C000-C809	092
(c) Kaposi Sarcoma	9140	C000-C809	093
(d) Other Specified Soft-Tissue Sarcomas	8840-8896, 8982, 8990, 9040-9044, 9120-9134, 9150-9170, 9251, 9581	C000-C809	094
	8963	C000-C639, C659-C768	094
	9231, 9240, 9363, 9364	C000-C399, C440-C809	094
	9260	C000-C399, C470-C768	094
(e) Unspecified Soft-Tissue Sarcomas	8800-8804	C000-C399, C440-C809	095
X Germ-Cell, Trophoblastic, and Other Gonadal Neoplasms			
(a) Intracranial and Intraspinal Germ-Cell Tumors	9060-9102	C700-C729, C751-C753	101
(b) Other and Unspecified Non-Gonadal Germ-Cell Tumors	9060-9102	C000-C559, C570-C619, C630-C699, C739-C750, C754-C809	102
(c) Gonadal Germ-Cell Tumors	9060-9102	C569, C620-C629	103
(d) Gonadal Carcinomas	8010-8041, 8050-8075, 8120-8122, 8130-8141, 8143, 8155, 8190-8201, 8210, 8211, 8221-8241, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8573	C569, C620-C629	104
	8380, 8381, 8441-8473	C000-C809	104
(e) Other and Unspecified Malignant Gonadal Tumors	8590-8670, 9000	C000-C809	105
	8000-8004	C569, C620-C629	105

Site Group	ICD-O-2 Histology (Type)	ICD-O-2 Site	Recode
XI Carcinomas and Other Malignant Epithelial Neoplasms			
(a) Adrenocortical Carcinoma	8370-8375	C000-C809	111
(b) Thyroid Carcinoma	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8155, 8190, 8200, 8201, 8211, 8230, 8231, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8430, 8440, 8480, 8481, 8500-8573	C739	112
	8330-8350	C000-C809	112
(c) Nasopharyngeal Carcinoma	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8155, 8190, 8200, 8201, 8211, 8230, 8231, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8430, 8440, 8480, 8481, 8504, 8510, 8550, 8560-8573	C110-C119	113
(d) Malignant Melanoma	8720-8780	C000-C809	114
(e) Skin Carcinoma	8010-8041, 8050-8075, 8082, 8090-8110, 8140, 8143, 8147, 8190, 8200, 8240, 8246, 8247, 8260, 8310, 8320, 8323, 8390-8420, 8430, 8480, 8542, 8560, 8570-8573, 8940	C440-C449	115
(f) Other and Unspecified Carcinomas	8010-8082, 8120-8155, 8190-8263, 8290, 8310, 8314-8323, 8430-8440, 8480-8580, 8940, 8941	C000-C109, C129-C218, C239-C399, C480-C488, C500-C559, C570-C619, C630-C639, C659-C729, C750-C809	116
XII Other and Unspecified Malignant Neoplasms			
(a) Other Specified Malignant Tumors	8930, 8933, 8950, 8951, 8971-8981, 9020, 9050-9053, 9110, 9580	C000-C809	121

Site Group	ICD-O-2 Histology (Type)	ICD-O-2 Site	Recode
(b) Other Unspecified Malignant Tumors	8000-8004	C000-C218, C239-C399, C420-C559, C570-C619, C630-C639, C659-C699, C739-C750, C754-C809	122
Not Classified by ICCC			999

[*] ICD-O-2 histology and site definitions differ between IARC Technical Report No. 29 and the SEER modification.
[†] Recode value differs between IARC Technical Report No. 29 and the SEER modification.
Source: SEER 2003.

Table 6. Standard Site Analysis Categories for Mortality Data (ICD-9 and ICD-10) SEER Cause of Death Recode 1969+ (3/25/2004)

For ICD-8 (1968-1978), All Malignant Cancers is defined as 140-207. Individual ICD-8 cancer codes are converted to ICD-9 prior to creating this variable.

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9** (1979-1998)*	**ICD-10** (1999+)*
All Malignant Cancers	140-208, 238.6	C00-C97
Oral Cavity and Pharynx		
Lip	140	C00
Tongue	141	C01-C02
Salivary Gland	142	C07-C08
Floor of Mouth	144	C04
Gum and Other Mouth	143, 145	C03, C05-C06
Nasopharynx	147	C11
Tonsil	146.0-146.2	C09
Oropharynx	146.3-146.9	C10
Hypopharynx	148	C12-C13
Other Oral Cavity and Pharynx	149	C14
Digestive System		
Esophagus	150	C15
Stomach	151	C16
Small Intestine	152	C17
Colon and Rectum		
Colon Excluding Rectum	153, 159.0	C18, C26.0
Rectum and Rectosigmoid Junction	154.0-154.1	C19-C20
Anus, Anal Canal, and Anorectum	154.2-154.3, 154.8	C21

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9** (1979-1998)*	**ICD-10** (1999+)*
Liver and Intrahepatic Bile Duct		
Liver	155.0, 155.2	C22.0, C22.2-C22.4, C22.7, C22.9
Intrahepatic Bile Duct	155.1	C22.1
Gallbladder	156.0	C23
Other Biliary	156.1-156.2, 156.8-156.9	C24
Pancreas	157	C25
Retroperitoneum	158.0	C48.0
Peritoneum, Omentum, and Mesentery	158.8-158.9	C45.1+, C48.1-C48.2
Other Digestive Organs	159.8-159.9	C26.8-C26.9, C48.8
Respiratory System		
Nose, Nasal Cavity, and Middle Ear	160	C30-C31
Larynx	161	C32
Lung and Bronchus	162.2-162.5, 162.8-162.9	C34
Pleura	163	C38.4, C45.0+
Trachea, Mediastinum, and Other Respiratory Organs	162.0, 164.2-164.3, 164.8-164.9, 165	C33, C38.1-C38.3, C38.8, C39
Bones and Joints	170	C40-C41
Soft Tissue including Heart‡	164.1, 171	C47, C49, C38.0, C45.2+
Skin Excluding Basal and Squamous		
Melanoma of the Skin	172	C43
Other Non-Epithelial Skin	173	C44, C46+
Breast	174-175	C50
Female Genital System		
Cervix Uteri	180	C53
Corpus and Uterus, NOS		
Corpus Uteri	182	C54
Uterus, NOS	179	C55
Ovary	183.0	C56
Vagina	184.0	C52
Vulva	184.1-184.4	C51

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9** (1979-1998)*	**ICD-10** (1999+)*
Other Female Genital Organs	181, 183.2-183.5, 183.8-183.9, 184.8-184.9	C57-C58
Male Genital System		
Prostate	185	C61
Testis	186	C62
Penis	187.1-187.4	C60
Other Male Genital Organs	187.5-187.9	C63
Urinary System		
Urinary Bladder	188	C67
Kidney and Renal Pelvis	189.0-189.1	C64-C65
Ureter	189.2	C66
Other Urinary Organs	189.3-189.4, 189.8-189.9	C68
Eye and Orbit	190	C69
Brain and Other Nervous System	191, 192	C70, C71, C72
Endocrine System		
Thyroid	193	C73
Other Endocrine Including Thymus‡	164.0, 194	C37, C74-C75
Lymphoma		
Hodgkin Lymphoma	201	C81
Non-Hodgkin Lymphoma	200, 202.0-202.2, 202.8-202.9	C82-C85, C96.3
Myeloma	203.0, 238.6	C90.0, C90.2
Leukemia		
Lymphocytic Leukemia		
Acute Lymphocytic Leukemia	204.0	C91.0
Chronic Lymphocytic Leukemia	204.1	C91.1
Other Lymphocytic Leukemia	202.4, 204.2, 204.8-204.9	C91.2-C91.4, C91.7, C91.9
Myeloid and Monocytic Leukemia		
Acute myeloid	205.0, 207.0, 207.2	C92.0, C92.4-C92.5, C94.0, C94.2
Acute Monocytic Leukemia	206.0	C93.0

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9** (1979-1998)*	**ICD-10** (1999+)*
Chronic Myeloid Leukemia	205.1	C92.1
Other Myeloid/Monocytic Leukemia	205.2-205.3, 205.8-205.9, 206.1-206.2, 206.8-206.9	C92.2-C92.3, C92.7, C92.9, C93.1-C93.2, C93.7, C93.9
Other Leukemia		
Other Acute Leukemia	208.0	C94.4, C94.5, C95.0
Aleukemic, Subleukemic, and NOS	203.1, 207.1, 207.8, 208.1-208.2, 208.8-208.9	C90.1, C91.5, C94.1, C94.3, C94.7, C95.1, C95.2, C95.7, C95.9
Mesothelioma(ICD-10 only)†	N/A	C45+
Kaposi Sarcoma (ICD-10 only)†	N/A	C46+
Miscellaneous Malignant Cancer	159.1, 195-199, 202.3, 202.5-202.6, 203.8	C26.1, C45.7+, C45.9+, C76-C80, C88, C96.0-C96.2, C96.7, C96.9, C97

Non-Cancer Causes of Death			
Non-Cancer Causes of Death	**ICD-8** (1968-1978) *	**ICD-9** (1979-1998) *	**ICD-10** (1999+)*
In situ, Benign, or Unknown Behavior Neoplasm	208-239	210-237, 238.0-238.5, 238.7-238.9, 239	D00-D48
Tuberculosis	010-018	010-018	A15-A19
Syphilis	090-097	090-097	A50-A53
Human Immunodeficiency Virus (HIV) (1987+)	N/A	042-044	B20-B24
Septicemia	38	38	A40-A41
Other Infectious and Parasitic Diseases	001-009, 020-037, 039-043, 045-065, 067-076, 078-089, 098-130.1, 130.3-136	001-009, 020-037, 039-041, 045-088, 098-139	A00-A09, A20-A39, A42-A49, A54-B19, B25-B99
Diabetes Mellitus	250	250	E10-E14
Alzheimer's (ICD-9 and 10 only)	N/A	331.0	G30
Diseases of Heart	390-398, 402, 404, 410-429	390-398, 402, 404, 410-429	I00-I09, I11, I13, I20-I51
Hypertension Without Heart Disease	400-401, 403	401, 403	I10, I12
Cerebrovascular Diseases	430-438	430-438	I60-I69
Atherosclerosis	440	440	I70
Aortic Aneurysm and Dissection	441	441	I71
Other Diseases of Arteries, Arterioles, Capillaries	442-448	442-448	I72-I78
Pneumonia and Influenza	470-474, 480-486	480-487	J10-J18
Chronic Obstructive Pulmonary Disease and Allied Cond	490-493, 519.3	490-496	J40-J47
Stomach and Duodenal Ulcers	531-533	531-533	K25-K28
Chronic Liver Disease and Cirrhosis	571	571	K70, K73-K74

Non-Cancer Causes of Death			
Non-Cancer Causes of Death	ICD-8 (1968-1978) *	ICD-9 (1979-1998) *	ICD-10 (1999+)*
Nephritis, Nephrotic Syndrome, and Nephrosis	580-584, 593.0-593.3, 593.5	580-589	N00-N07, N17-N19, N25-N27
Complications of Pregnancy, Childbirth, Puerperium	630-678	630-676	A34, O00-O95, O98-O99
Congenital Anomalies	740-759	740-759	Q00-Q99
Certain Conditions Originating in Perinatal Period	760-779	760-779	P00-P96
Symptoms, Signs, and Ill-Defined Conditions	780-796	780-799	R00-R99
Accidents and Adverse Effects	800-949§	800-949§	V01-X59, Y85-Y86
Suicide and Self-Inflicted Injury	950-959§	950-959§	X60-X84, Y87.0
Homicide and Legal Intervention	960-978§	960-978§	X85-Y09, Y35, Y87.1, Y89.0

* All ICD codes are tested for validity prior to generating this variable. Those deemed invalid are classified as Unknown/missing/invalid COD. Those deemed valid but not meeting the definition of any above grouping are classified as Other Cause of Death.
† This variable can be created with or without Mesothelioma (C45) and Kaposi Sarcoma (C46) as separate groupings. The table above documents both possibilities. Note this is only possible with ICD-10.
‡ ICD-8 code 192.5 is coded to Other Endocrine including Thymus for age at death < 20 years and Soft Tissue including Heart for age at death 20+ years.
§ External causes of injury and poisoning.
Source: SEER 2003.

4.2.1.2. Standards for Age Categories

The age distribution of cancer patients is most often summarized in 5- or 10-year age groups. The registry **SHOULD** use the recommended 5-year age groups beginning with the category 0, and continuing through ages 85 and older (i.e., 0, 1-4, 5-9, 10-14, …75-79, 80-84, 85+). These are the standard groups used for population denominators. Pediatric cancers are defined as those occurring under age 15 and under age 20. For some pediatric cancers, single-year age groups are desired when incidence rates change dramatically within the 5-year interval. Some registries use "75+" years as the oldest age category, but there is increasing interest in cancer in older age groups, and it is important to provide data for the oldest groups.

If a particular analysis does not use 5-year age groups (e.g., when the number of cases is small), the registry **SHOULD** choose age groups that allow for appropriate comparisons with data for the population at risk.

4.2.1.3. Standards for Time Period Categories

One year is the shortest interval used to present cancer incidence statistics. However, one of the primary concerns in determining how best to summarize data by time period is the number of tumor records that are available for analysis. Thus the choice of time period intervals normally is based on the length of time the registry has been in existence and the size of the population covered. Analyses usually are based on calendar year of diagnosis. Cancer surveillance periods generally are based on 1-5 years.

Central cancer registries that have covered large populations may well have sufficient data to evaluate time trends in cancer statistics on a year-by-year basis. In contrast, registries with a small population base will have insufficient data to present stable reliable statistics in such detail. Three- or 5-year averages are useful tools to reduce random variation in statistics created from small numbers. Time trends should not be analyzed for periods of less than 5 years, and a 10-year period is more acceptable. Statistical approach depends on the number of years in the interval. For example, short-term trends estimated annual percent change (EAPC) or linear regression are commonly used methods. For longer-term trends such as 20 years or more, joint-point analysis is a useful method.

Registry staff **SHOULD** consult with an experienced epidemiologist, biostatistician, or demographer to determine how best to present temporal trends in cancer statistics.

4.2.1.4. Canadian Standards for Geographic Area Categories

The Canadian standard is the Standard Geographic Classification (SGC). The SGC relates to the patient's usual, permanent place of residence at the time of diagnosis for a particular tumor. The code includes the province/territory (2 digits), census division (2 digits), and census subdivision (3 digits). The census division identifies the county and the census subdivision identifies the municipality. There also is a census tract that is a smaller geographic unit than municipality and is found only in large urban communities. The census tract is comprised of 3-digit Census Metropolitan Area/Census Agglomeration code (CMA/CA), followed by a census tract code that is unique only within CMA/CA. Canadian data normally are tabulated by province and territory of residence as well as for Canada as a whole. The 7-digit SGC code allows the 13 jurisdictions to be tabulated individually, or as part of one of six regions, by using the first digit alone, as shown in Table 7.

Table 7. SGC Codes for Canadian Provinces and Territories

Region	Code	Province/Territory
Atlantic	10	Newfoundland and Labrador*
	11	Prince Edward Island
	12	Nova Scotia
	13	New Brunswick
Quebec	24	Quebec
Ontario	35	Ontario
Prairies	46	Manitoba
	47	Saskatchewan
	48	Alberta
Pacific	59	British Columbia
North	60	Yukon
	61	Northwest Territories
	62	Nunavut

*The boundaries, names, codes, and status of the standard geographic areas reflect those in effect on January 1, 2001, with the exception of the name change of the province of Newfoundland and Labrador (previously Newfoundland) which became effective on December 6, 2001.
Source: Statistics Canada, Standards Division 2001.

4.2.1.5. Standards for Treatment Categories

The information provided in this section is according to SEER rules. For routine reports, first course of treatment generally is reported in categories that group the modalities given. The SEER recode is presented in Table 8 (SEER 2002). It groups cases into 19 categories showing combinations of the major modalities of surgery, radiation, chemotherapy, and hormonal therapy; grouping all other treatments into an “other” category. Each grouping is assigned two codes, one 4-digit and one 2-digit, identical in meaning. The 2-digit code assigns consecutive numbers to the recoded groupings.

Table 8. Standard Treatment Analysis Categories

	Treatment Modality									Group Code
Category Name	**S 1983-1997**	**S 1998-2002**	**RNS**	**R**	**RCNS**	**C**	**H**	**BRM**	**O**	**REC B**
Surgery Only	10-99	10-90	0	0, 7-9	0, 7-9	0, 7-9	0, 7-9	0-9	0-9	01
Radiation Only	00-09	00,99	1-9	1-5 or 1		0, 7-9	0, 7-9	0-9	0-9	02
Chemotherapy Only	00-09	00,99	1-9	0, 7-9	0, 7-9	1-3	0, 7-9	0-9	0-9	03
Hormonal Therapy Only	00-09	00,99	1-9	0, 7-9	0, 7-9	0, 7-9	1-3	0-9	0-9	04
Surgery and Radiation	10-99	10-90	0	1-5 or 1		0, 7-9	0, 7-9	0-9	0-9	05
Surgery and Chemotherapy	10-99	10-90	0	0, 7-9	0, 7-9	1-3	0, 7-9	0-9	0-9	06
Surgery and Hormonal Therapy	10-99	10-90	0	0, 7-9	0, 7-9	0, 7-9	1-3	0-9	0-9	07
Radiation and Chemotherapy	00-09	00,99	1-9	1-5 or 1		1-3	0, 7-9	0-9	0-9	08
Radiation and Hormonal Therapy	00-09	00,99	1-9	1-5 or 1		0, 7-9	1-3	0-9	0-9	09
Chemotherapy and Hormonal Therapy	00-09	00,99	1-9	0, 7-9	0, 7-9	1-3	1-3	0-9	0-9	10
Surgery, Radiation, and Chemotherapy	10-99	10-90	0	1-5 or 1		1-3	0, 7-9	0-9	0-9	11
Surgery, Radiation, and Hormonal Therapy	10-99	10-90	0	1-5 or 1		0, 7-9	1-3	0-9	0-9	12
Surgery, Chemotherapy, and Hormonal Therapy	10-99	10-90	0	0, 7-9	0, 7-9	1-3	1-3	0-9	0-9	13
Radiation, Chemotherapy, and Hormonal Therapy	00-09	00,99	1-9	1-5 or 1		1-3	1-3	0-9	0-9	14

	Treatment Modality									Group Code
Category Name	**S 1983-1997**	**S 1998-2002**	**RNS**	**R**	**RCNS**	**C**	**H**	**BRM**	**O**	**REC B**
Surgery, Radiation, Chemotherapy, and Hormonal Therapy	10-99	10-90	0	1-5 or 1		1-3	1-3	0-9	0-9	15
Other Treatment	00-09 00-09	00,99 00,99	1-9 1-9	0, 7-9 0, 7-9	0, 7-9 0, 7-9	0, 7-9 0, 7-9	0, 7-9 0, 7-9	1 0, 7-9	0-9 1-6	16
No Treatment	00-09	00,99	1-9	0, 7, 9	0, 7,9	0, 7,9	0,7,9	0,7,9	0,7,9	17
Treatment Unknown	00-09 00-09	00,99 00,99	8,9 1-9	8,9 0, 7-9	8,9 0, 7-9	8,9 0, 7-9	8,9 0, 7-9	8,9 0, 7-9	8,9 0, 7-9	18
Invalid Treatment Code										99

SEER has removed Recode A from the table.
S = Surgery; RNS = Reason no surgery; R = Radiation: RCNS = Radiation to brain and central nervous system;
C = Chemotherapy; H = Hormonal therapy; BRM = Biological response modifiers; O =Other therapy.
Source: SEER 1993b.

According to the SEER rules, at least one modality must be none. For modalities other than surgery, none is indicated by a 0 or 7 (patient refused). Before 1988, no surgery is specified by 09 in site-specific surgery and 2, 6 in reason for no cancer-directed surgery. After 1988, no surgery is specified by 00-07 in site-specific surgery and 1-7 in reason for no cancer-directed surgery.

In the second category of treatment unknown, at least one modality must specify recommended, unknown if performed.

4.2.1.6. Standards for Grouping by Stage of Disease

Collaborative stage was implemented January 1, 2004. The collaborative stage schema incorporates all of the fields from the SEER 10-digit Extent of Disease (EOD) (in a modified form) plus several additional fields (see *NAACCR Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*). When collaborative stage data items are coded, a computer algorithm derives the AJCC Sixth Edition Tumor, Nodes, Metastasis stage; SEER Summary Stage 1977; and SEER Summary Stage 2000. The derived collaborative staging field selected for analysis **SHOULD** be based on the purpose of the study.

4.2.2. Statistical Methods

It is important to consider each of the methods outlined below in the context of three key elements of epidemiologic inquiry: person, place, and time. Analyses usually are based on tumor records (i.e.,

independent primary cancers, of which a person can have more than one; however, some analyses focus on persons rather than tumors.

- *Person:* Reports of cancer data **SHOULD** document the demographic characteristics of the tumors represented in the report. At a minimum, these characteristics should include sex, age, and race/ethnicity (U.S.). A person may be represented by more than one tumor (i.e., more than one primary cancer) in the registry's files.

- *Place:* Reports of cancer data **MUST** specify the geographic area of coverage for the tumors represented in the report. Typically, the area of coverage follows political boundaries such as provinces, states, counties, and cities, or census entities (see Section 4.1.1.4. for a discussion on confidentiality and data for small areas).

- *Time:* Reports **MUST** clearly state the relevant time period of study. Cancer statistics usually are reported in calendar years annually, based on the diagnosis year and not the year the case was reported.

4.2.2.1. Standards for Counts

The most basic unit of measure for cancer registry data is the simple enumeration of tumors. Knowledge of the number of tumor records can be of great use for health planning purposes where it is important to measure the burden of cancer on existing health care resources and to assess the need for additional resources. However, simple counts of tumors are of limited value as a measure of disease risk, for which incidence rates are preferable (see Section 2.1.2.2.).

4.2.2.2. Standards for Proportions

4.2.2.2.1. Simple Proportions

Simple proportions are useful for describing basic characteristics of registry data. Examples include:

- Percent distribution of tumors by stage of disease at diagnosis.

- Proportion of tumors with a histologically confirmed diagnosis.

- Proportion of tumors that received a given treatment modality.

4.2.2.2.2. Percent Distribution by Site

A percentage distribution by site is useful for showing at a glance which cancers account for the majority of tumor records. Usually, cancers of the breast, lung, colorectum, and prostate together will account for more than one-half of all cancers, with each accounting for 12 to 15 percent of all cancers. This is a useful distribution to present so that non-population-based registries can assess whether their data represent a skewed distribution of tumors.

4.2.2.2.3. Proportional Incidence

As outlined in Section 4.2.2.3., incidence rates are the measure of choice for expressing disease risk; however, appropriate population estimates are not always available to serve as the denominators for rate calculations. In these instances, the proportional incidence ratio (PIR) may serve as a useful way to compare risk of disease in

two populations. This measure compares the relative importance of a specific cancer in relation to all cancers in two groups in a specified time period.

The PIR is calculated using the proportional distribution within a defined group (e.g., whites) to estimate the expected proportion in another group (e.g., Japanese). The observed proportion then is compared to the expected proportion as an estimate of risk. Specifically, the proportion of all tumors accounted for by a specific site is calculated for each age and sex group in the comparison population (e.g., whites). These proportions then are applied to the number of all cancers in each age and sex group in the comparison population (e.g., Japanese) to estimate the number of expected tumors of that type by age and sex. Expected numbers are summed across age and sex groups to obtain an "age-adjusted" expected number of tumors. The ratio of the observed tumors compared to the expected tumors yields the PIR. The PIR generally is multiplied by 100; a PIR of greater than 100 indicates that the observed proportion was greater than the expected proportion, and usually indicates an increased disease risk.

4.2.2.3. Standards for Incidence Rates

4.2.2.3.1. Standardization

A set of techniques is used to remove as much as possible the effects of difference in age or other confounding variables, when comparing two or more populations. The common method uses weighted averaging of rates specific for age, sex, or some other potential confounding variables(s) according to some specified distribution of these variables.

- *Direct Method:* The specific rates in a study population are averaged, using the distribution of a specified standard population as weights. The directly standardized rate represents what the crude rate would have been in the study population if that population had the same distribution as the standard population with respect to the variables(s) for which the adjustment or standardization was carried out.

- *Indirect Method:* This is used to compare study populations for which the specific rates are either statistically unstable or unknown. The specific rates in the standard population are averaged, using as weights the distribution of the study population. The ratio of the crude rate for the study population to the weighted average so obtained is the standardized mortality ratio (SMR). The indirectly standardized rate itself is the product of the SMR and the crude rate for the standard population.

- *Standardized Incidence Ratios:* The ratio of the number of events observed in the study group or population to the number that would be expected if the study population had the same specific rates as the standard population, multiplied by 100.

- *Standardized Mortality Ratio:* The ratio of the number of deaths observed in the study group or population to the number that would be expected if the study population had the same specific rates as the standard population, multiplied by 100.

4.2.2.3.2. Incidence Rates

Incidence rates are more useful measures of disease risk than proportional rates. Incidence rates express the number of new tumors diagnosed in a population with respect to the size of the population and the time period under study. Specific incidence rates can include:

- *Crude Incidence Rate:* The simplest incidence rate, obtained by dividing the number of new tumors by the size of the population at risk of developing cancer during the study period. The crude rate does not take into account the age distribution of the population; therefore, crude rates are not suitable for comparison across place and time.

- *Age-Specific Incidence Rate:* The age-specific incidence rate is the incidence rate for a defined age group.

- *Age-Adjusted Incidence Rate:* The age-adjusted incidence rate is a rate that adjusts for the differential impact of age on cancer risk (i.e., older persons higher risk than younger persons) and this is useful in comparison of cancer rates among different locations, populations or other factors. Usually, standardization for age is carried out through the direct method.

4.2.2.3.3. Case Selection Criteria

When selecting cases for incidence rate calculations:

- Include only resident cases first diagnosed during the selected time period.

- If a cancer is a DCO, count resident cases reported as incident at the date of death.

- Include cases discovered at autopsy for residents only.

- Include invasive cases only (with the exception of urinary bladder cancer) in calculation of rates of malignant neoplasms.

4.2.2.3.4. Denominators for Rate Calculation

One of the most important steps in calculating incidence or mortality rates is to obtain appropriate population estimates to serve as the denominator for the rate calculation. These estimates represent the population at risk. For a central cancer registry, these estimates would represent the population that resides within the registry's designated coverage area. For incidence rates, the population estimates should correspond to the population that resides within the registry's capture area for the time period during which the newly diagnosed tumors were identified in the population at risk (see Section 4.1.2. for a general discussion of population estimates).

4.2.2.3.5. Standard Population

The choice of an appropriate standard population is an issue in the calculation of age-adjusted rates. The choice of data for comparison may dictate the choice for standard population.

- *U.S. Standard:* The age structure of the U.S. population has changed considerably from the 1970 U.S. Standard population. This led to the adoption of the year 2000 standard for computing age-adjusted rates. Many national agencies, such as the NCHS, adopted the 2000 U.S. standard, effective for 1999 and later diagnoses, deaths, or other health statistics. The 1970 and 2000 U.S. Standard populations are shown in Table 9.

- *Canadian Standard:* Canada's 1991 and 1996 population is used to standardize rates for routine comparisons within Canada. The 1991 and 1996 populations are shown in Table 9. The standard selected for NAACCR publications follows the recommendation of Statistics Canada.

- *World Standard:* Another common comparison population, and the one used in WHO's *Cancer Incidence in Five Continents* (Waterhouse J, Miur C, Correa P, Powell J (eds). *Cancer Incidence in Five Continents, Volume III.* Lyon, France: International Agency for Research on Cancer, IARC Scientific Publications No. 15, 1976) is the world standard used by the IARC, also shown in Table 9. This is useful for international comparisons. There also is a World 2000 Standard that is not used for cancer.

Table 9. Standard Populations

Age Group	Numbers in Group				
	1970 U.S. Standard	2000 U.S. Standard	1991 Canadian Population	1996 Canadian Population	World Standard Population
All Ages	1,000,000	1,000,000	1,000,000	1,000,000	100,000
< 5	84,416	69,135	69,465	66,235	12,000
5-9	98,204	72,533	69,464	67,985	10,000
10-14	102,304	73,032	68,034	67,716	9,000
15-19	93,845	72,169	68,495	67,841	9,000
20-24	80,561	66,478	77,016	67,761	8,000
25-29	66,320	64,529	89,944	72,914	8,000
30-34	56,249	71,044	92,400	87,030	6,000
35-39	54,656	80,762	83,388	88,510	6,000
40-44	58,958	81,851	76,063	80,055	6,000
45-49	59,622	72,118	59,536	71,847	6,000
50-54	54,643	62,716	47,649	55,812	5,000
55-59	49,077	48,454	44,041	44,869	4,000
60-64	42,403	38,793	42,326	40,705	4,000
65-69	34,406	34.264	38,570	37,858	3,000
70-74	26,789	31,773	29,660	32,589	2,000
75-79	18,871	26,999	22,127	23,232	1,000
80-84	11,241	17,842	13,575	15,424	500
85+	7,435	15,508	10,237	11,617	500

4.2.2.3.6. Guidelines for Incidence Rate Calculations

When calculating incidence rates for the registry as a whole or for any geographic area within the registry's area of coverage, the registry **SHOULD:**

- Eliminate cases with unknown age, sex, or geographic area of residence from all calculations. They **SHOULD** be excluded from rate calculations where appropriate, and the report **SHOULD** show the number of cases that were excluded because of unknown data (see Section 3.3.3. for a discussion on unknown values).

- Evaluate variability in rates and select the most appropriate method to present the rates. Show the standard errors, suppress rates based on small numbers, or otherwise footnote the results based on small numbers of cases.

4.2.2.3.7. Units of Measure

Cancer incidence rates **SHOULD** be expressed per 100,000 population per unit of time. Some rare cancers (childhood cancers, for example) are expressed per 1,000,000 population per unit of time.

4.2.2.4. Standards for Death Rates

Death rates most often are reported by local health agencies or bureaus of vital statistics based on information reported through death registration. However, because of their expertise and focus on cancer and need for confidentiality with incidence rate calculations, central cancer registries need to calculate cancer death rates as well. Cancer death rates **SHOULD** be based on the underlying cause of death as reported through the death registration process.

As with incidence rates, death rates can be expressed as crude, age-specific, or age-adjusted. The methods outlined above for incidence rates also are applicable to death rates (the same denominators should be used for mortality as for incidence for the identical time period). The population estimates used **MUST** correspond to the same time period during which the deaths of interest occurred.

The accuracy of death rates as a measure of cancer occurrence has been shown to vary by type of cancer. For this reason, caution **SHOULD** be exercised in the use and interpretation of cancer death rates.

4.2.2.5. Standards for Survival Analysis

Survival analysis entails measuring the length of time between two events. Most frequently for cancer registries, the initial event is the date of cancer diagnosis, and the second event is a subsequent outcome, such as death. Survival rates can be used as an index of the quality of, not only early diagnosis, but also care following a diagnosis of cancer.

- Determine purpose of the study.

- Select cases based on the purpose of the study.
  - All inclusions and exclusions **MUST** be accounted for.

- Follow-up **MUST** be at least 90 percent complete for the patient group selected.

4.2.2.5.1. Data Requirements

The following data items are the minimal requirements for calculating survival rates:

- *Date of Diagnosis.*

- *Date of Last Contact:* The date of last contact represents the calendar time at which information was last obtained on the subject. If the patient is deceased, the date of last contact is the date of death. The accurate ascertainment of the date of last contact for all cancer patients is a key factor in the validity of survival analysis (when survival to recurrence of cancer is being calculated, it is the date of recurrence that is used as the subsequent outcome).

- *Vital Status:* Vital status describes the last known condition of the subject. This item indicates whether the subject was alive or dead at the date of last contact. Some methods of survival analysis require knowledge of the cause of death. When survival to recurrence is being calculated, the patient's recurrence status is used instead of vital status.

4.2.2.5.2. Standard Methods

Four standard methods of survival analysis are described below.

- *Observed Survival Rate:* The observed survival rate is calculated by the life-table (actuarial) method. This method provides an estimate of the probability of an individual surviving to the end of a specified time interval, given that the person was alive at the beginning of this interval.

- *Relative Survival Rate:* The relative survival rate also is calculated by the life-table (actuarial) method. This method adjusts the observed survival rate to account for other causes of death that would be expected if the study subjects experienced the same mortality rates as the general population of similar age, race, sex, and calendar period of observation. By adjusting for other causes of death, this method attempts to estimate the effect of the cancer alone on survival. What this method does, in fact, is measure the excess mortality that the cohort experiences in comparison to the general population. The accuracy of this method is a function of how the study subjects differ from the general population. If the only difference is the fact of cancer, then this method works well. One notable exception is lung cancer, in which the cancer cohort also is at excess risk of death from heart disease compared to the general population due to a large number with a history of smoking.

- *Kaplan-Meier:* The Kaplan-Meier Method, also known as the product limit method, is a special case of the standard life table technique used for survival analysis. Kaplan-Meier is computationally similar to the standard life-table method, but the intervals of survival time are defined differently for the two methods. In the Kaplan-Meier Method, a calculation (of the observed survival rate) is done every time a patient dies rather than during a specific regular interval, such as a year or a month. Thus, it results in a more exact description of the pattern of survival. The graphic display of survival rates derived from Kaplan-Meier is particularly useful for determining the median survival time and for comparing the survival experiences of two or more groups of patients. Because multiple calculations are required, the Kaplan-Meier Method generally is used when the number of patients is small, generally 25 to 30, as usually is the case in clinical trials. Statistics texts should be consulted for more details.

- *Cox Proportional Hazards Model:* The Cox Proportional Hazards Model allows for the comparison of survival rates between two or more groups, with simultaneous adjustment for potentially confounding variables.

4.2.2.5.3. Interpretation

Survival from cancer is determined by many factors, including the patient's age, stage of disease at diagnosis, histologic type of cancer, treatment, and the presence of other illnesses. Comparison of survival rates among institutions or geographic areas **MUST** be interpreted carefully, especially if the respective patient populations differ with regard to prognostic factors.

Calculation, interpretation, and reporting of survival rates **SHOULD** be undertaken only under the supervision of a qualified biostatistician or epidemiologist with expertise in survival analysis and after the registry has employed standard approaches to identify completely all deaths among the registered cancer cases (i.e., proactive follow-up of cancer cases).

4.2.3. Reports

The dissemination of data is an important function of the central cancer registry. Registry data may appear routinely in a standard format or may be prepared on an *ad hoc* basis in response to specific inquiries. The reputation and usefulness of a central cancer registry often is judged by the accuracy, timeliness, and clarity of its reports.

In designing reports, it may be useful to compare one registry's experience with similar data from other cancer registries. Similarly, it may be helpful to design reports that are comparable within a registration system.

Registries **SHOULD** obtain copies of reports and newsletters from other registries to use as models when developing their own publications. Most cancer registries are pleased to include other registries in the routine distribution of their reports and newsletters.

For a discussion of data management considerations in the design and production of reports, see Section 5.6.

4.2.3.1. Standards for Type and Frequency of Reports

4.2.3.1.1. Summary of Central Registry Data

Central cancer registries **SHOULD** assemble a comprehensive summary of the cancer burden (incidence and mortality) within their area of coverage. At a minimum, the report **MUST** tabulate tumors by primary site, sex, race, age group, and sub-regions of the area.

In addition, these reports **SHOULD** provide population-based incidence and/or death rates, tabulated by site groups, age, and sex. If available, survival rates **MAY** be presented in these reports. Where possible, incidence, death, and survival rates **SHOULD** be displayed by ethnicity, race, and stage. If the registry has been in existence for a sufficiently long time period, and if the number of cases permits, the report **SHOULD** include temporal trends in cancer incidence, death, and survival rates. Some registries may elect to provide similar information by sub-geographic area.

Summaries of central cancer registry data **SHOULD** be published annually.

4.2.3.1.2. Reports to Hospitals and Other Facilities

A central registry **SHOULD** provide a facility-specific summary to all reporting facilities within its jurisdiction, reflecting all cases for which the facility is the reporting source, including non-residents, non-analytic cases, and any other cases reported by the facility. At a minimum, these reports **SHOULD** tabulate the facility's tumor records by type of cancer, age, sex, and race using the standard groups described in Section 4.2.1. It is extremely useful to provide data that allow facilities to compare their own tumor records with summary, non-confidential data for the central registry's entire coverage area.

Facilities participating in the ACoS Approvals Program are required to present data, when available, in their annual reports comparing their facility's experience to a larger population. The central registry can meet this need by providing reports including tables and graphs showing frequencies, percent distributions, and, if available, survival data by primary site, stage of disease at diagnosis, and age at diagnosis. The central registry's data generally will not be as timely as the facility's, so comparison data from earlier years **MAY** be used. The most recent comparison data **SHOULD** be used.

Hospital and institutional summaries often include a list of the cancer patients seen at the facility. These lists **MAY** include patient name, age, stage of disease at diagnosis, histologic type, primary site, and date of last follow-up. It is helpful to provide patient lists sorted alphabetically, by the facility's accession number, and cancer type. However, confidentiality **MUST** be guaranteed.

At a minimum, hospital and institutional summaries **SHOULD** be provided annually. However, some central cancer registries generate these reports quarterly or semi-annually. Also, the frequency of these reports **MAY** depend on the facility's caseload, so that facilities with a large number of tumor records receive the reports more frequently than facilities with a small number of cancer patients.

In addition to reports as described above, the registry **MAY** consider providing patient follow-up information to hospitals, such as results of death clearance and other follow-up activities. This can be of great value to hospital cancer registries in reducing follow-up workload.

The registry **SHOULD** consider producing the reports on electronic media in addition to or replacing hardcopy reports.

4.2.3.1.3. Reports to Physicians

Reports to individual physicians **MAY** include descriptive statistics for their specialty (e.g., melanoma for dermatologists). Physicians may make special data requests or request follow-up information; these reports **SHOULD** be generated upon request (see Section 4.2.3.1.6.).

4.2.3.1.4. Newsletters

Newsletters are useful tools for the dissemination of registry information to members of the medical community and the general public. Newsletter articles may focus on registry activities or provide a useful vehicle for disseminating data. Some registries focus a single issue of their newsletter on data for a specific type of cancer.

The publication of newsletters, as well as the frequency of publication, will vary by registry, often depending on resources and available staff time. Typically, newsletters are produced quarterly or semi-annually.

4.2.3.1.5. Joint Publications

Some central cancer registries in the United States and Canada issue joint publications with survivor groups, groups with special cancer interests, or their cancer society. *Canadian Cancer Statistics* and publications from the Colorado and North Carolina registries are three examples of joint publications with the local cancer societies.

4.2.3.1.6. Requests for Information

Requests for information, whether from the medical community, press, governmental agencies, legislators, or the general public, **SHOULD** be addressed in a timely manner. The registry **SHOULD** keep a central cumulative log of all requests for information and **SHOULD** keep a file of responses to all requests. (See Section 4.1.1.).

Caution **MUST** be exercised when using confidential information with data gathered from other registries (through data exchange agreements) and from Vital Statistics. The confidentiality guidelines of all agencies **MUST** be taken into account.

4.2.3.1.7. Occasional or Special Topic Reports

The registry **SHOULD** produce focused reports as needed on topics of special interest, for example, in-depth analyses of specific cancer sites, geographic areas, or cancer disparities.

4.2.3.2. Standards for Narrative Text

4.2.3.2.1. General Considerations

An important component of any report is the narrative text that accompanies the presentation of the data. As outlined in Section 4.2.3.2.2., the narrative guides the reader by documenting methods used to produce the report, highlighting important findings, and interpreting the results.

4.2.3.2.2. Documentation

One of the primary functions of the narrative is to document the methods by which the data were collected, compiled, and analyzed.

- The report **SHOULD** include an overview of the registry's data collection methods.
- The narrative **SHOULD** specify the classification systems used to collect, code, and tabulate the data (e.g., ICD-O-3 for tumor diagnoses, and ICD-10 for mortality diagnoses).
- The report **MUST** clearly identify any recodes used and the statistical methodology that was used to conduct the analysis and prepare the report. References to more detailed descriptions of methods **SHOULD** be cited when the methodology cannot be fully described in the report.
- The report **MUST** identify the geographic area of coverage of the central cancer registry, as well as any specific geographic areas on which the report may focus.
- The report **MUST** clearly state the time period for which cases are tabulated.

- The narrative **MUST** document the source of the population counts that were used to calculate the rates when incidence and/or mortality rates are presented. If age-adjusted rates are included, the report **MUST** indicate the choice of standard population. A separate table of the relevant population counts, including the distribution of the standard population, **SHOULD** be provided.

4.2.3.2.3. Highlighting and Interpreting the Results

An explanatory narrative may be used to provide a more complete description of data, (i.e., what is outstanding, different, or notable). Consideration **MUST** be given to the audience reading the material to prevent misinterpretation of the text and the data.

Changes in data collection procedures or changes in disease classification **MUST** be documented because they may lead to a misinterpretation of the data. Similarly, changes in diagnostic methods or procedures may affect the numbers of tumors diagnosed or their classification into cancer site groups.

The reader **MUST** be cautioned against drawing definitive conclusions when the measures are based on small numbers.

4.2.3.2.4. Quality Indicators

Data quality can be an important contributor to the data interpretation and should be considered before conclusions are drawn. The report **SHOULD** address what is known about the completeness and accuracy of the data in the report. For incidence statistics, this **SHOULD** include information used in NAACCR Registry Certification:

- Completeness of case ascertainment.
- Accuracy (i.e., passing EDITS).
- Death certificate only cases.
- Timeliness.
- Duplicate reports.
- Completeness of key data variables (sex, age, county, race).

Other registry data uses also **SHOULD** involve a quality assessment of the variables used in the analysis before the analysis is conducted to evaluate whether the data are sufficiently complete and accurate to use in the analysis.

4.2.3.3. Standards for Displaying the Data

4.2.3.3.1. Tables

Numerical data often are displayed in tabular format. Tables **MUST** stand alone; that is, they **MUST** be fully comprehensible if separated apart from the narrative text. Descriptive titles, headings, and footnotes are used to explain the contents of the table. If data from a source other than the registry are used, a reference to the source **MUST** be noted.

4.2.3.3.2. Graphs and Charts

The graphical presentation of data often is more intuitively appealing than a table full of numbers (use tables when precision is important, use graphs when a more general idea or picture is desired). However, 3-dimensional charts or graphs **SHOULD NOT** be used when presenting bivariate data, because the depth of lines or bars can be misleading. If the results of a combination of three variables are displayed simultaneously, then 3-dimensional charts are appropriate. Some of the most common types of graphs are listed below.

- *Line Graphs:* Line graphs are constructed by plotting the values for two variables on an x-y axis, and then connecting the points. Line graphs are most often used to display time trends in age-adjusted incidence rates. When choosing the scale of the y-axis for presenting time trends, a decision needs to be made whether the absolute change or the rate of change is of more interest. Rates of change can only be shown on a logarithmic scale.

- *Bar Graphs and Histograms:* Bar graphs and histograms use horizontal or vertical bars to represent categorical data.

- *Pie Charts:* Pie charts can be used to display percent of the total, (e.g. site-specific stage groupings). To construct a pie chart, a circle is divided into segments, like slices of a pie, to represent various contributions to the whole.

4.2.3.3.3. Maps

Maps can be an effective method to display data. Maps can be used to compare summary statistics and rates for different geographic areas or to plot locations of specific cases as might be required in cancer cluster analyses. Software packages have made sophisticated complex mapping techniques available to every registry at relatively low cost. Polar coordinates for registry cases can be obtained automatically as part of a geocoding process (see Section 5.5.2.7). Selecting the appropriate and statistically valid mapping techniques, scales, colors, and other aspects of maps all require a great deal of thought and training to prevent unwarranted conclusions, breaches of confidentiality, or public alarm. For example, highlighting the county with the highest rate of a cancer in red on a map might be misleading to the public and scientifically indefensible if the county's rate is not significantly different from the next five ranked counties. Problems of small numbers and confidentiality apply to maps just as they do to other presentations of data (see the NAACCR document *Using Geographic Information Systems Technology in the Collection, Analysis and Presentation of Cancer Registry Data: A Handbook of Best Practices*).

4.2.3.3.4. Titles

Titles should identify:

- What the entries in the tables, charts, or maps are (e.g., number of cases, percents, rates, ratios, etc.).

- How the data are subdivided (e.g., by race, sex, age, histology, etc.).

- Who is included (e.g., all races, both sexes, etc.).

- Where the data are from (e.g., the SEER Program, Utah, Memorial Hospital, etc.).

- Time period covered (e.g., 1985-89, etc.).

The preferred order of elements in titles is: (1) what and how classified, (2) who, (3) where, and (4) when.

4.2.3.4. Standards for Review of Reports

Registries **MUST** follow written rules, protocols, and procedures for release of information. The central cancer registry **MUST** designate staff members to review all routine reports and responses to requests for information before the information is released to assure that confidentiality of the data is protected. In addition, participating facilities/organizations **SHOULD** request a courtesy review of the publication prior to release (see Section 4.1.1.).

All questions regarding the quality of the data **MUST** be brought to the attention of the quality control staff and **SHOULD** be resolved before data are released.

All questions regarding the appropriate interpretation of registry data **MUST** be brought to the attention of appropriate staff and **SHOULD** be resolved before the data are released.

Because of the possible ramifications for the registry, its participating facilities, and its parent organizations, the Registry Director or designee **MUST** review and approve all information released to the news media. The Registry Director or designee **SHOULD** inform the appropriate supervisors, stakeholders and data providers before release so that they will be able to answer any subsequent questions from the press or the community.

4.2.4. Electronic Publication and Distribution of Registry Data

In addition to preparing written reports on paper, some registries publish data in electronic form. Often, requests for registry data for research can be satisfied through provision of a data file developed for public use. Epidemiologists, biostatisticians, public health officers, and students all can benefit greatly from the ability to formulate and run their own queries of the data. Public use files can be provided to universities, medical schools, health departments, physicians, epidemiologists, voluntary cancer societies, and science journalists. The registry can proactively provide such files to potential users or respond to requests. For a user with the capacity to manipulate the electronic file, there are great advantages of flexibility.

Electronic distribution of registry data places additional obligations on the registry to protect confidentiality by restricting inappropriate uses. This normally can be handled with signed confidentiality agreements to all recipients of the data. The NAACCR Record Uniqueness Program **SHOULD** be used to evaluate all released data files for the potential risk of confidentiality breaches (www.naaccr.org).

Examples of the Data Confidentiality Agreement for NAACCR Researchers and the SEER Public Use File Agreement are presented in Appendix I.

A potential concern with public use files is preventing scientifically inappropriate analyses and conclusions. The registry may have additional burden of training the users of their data files.

4.2.4.1. Types of Electronic Publication

The most versatile method of electronic publishing is the “Public Use File,” generally an electronic file with one record per tumor with identifiers deleted. Some data items that might lead indirectly to a patient’s identity, especially for small areas, also might be omitted. Collapsed categories may be used, or values for standard recodes can be a part of each record.

Another approach combines the data file with analysis software used on Web query systems. The software can have built-in recodes to appropriate analysis groups; can suppress statistically insignificant or

meaningless results, or can suppress cells with small numbers. The potential user base is broadened when the analysis software is provided and is user friendly. An example is CINA+ Online.

CINA + Online, an online query system, was developed as a publicly available data source. It provides access to incidence data on all SEER major and minor cancer sites (including pediatric groups) for North America, the United States and Canada, with individual state- or province-specific data available. The online system is a flexible interactive query system that offers a choice of custom-designed tables, charts (multi-line graphs, pie charts, or bar graphs), and maps.

4.2.4.2. Distribution Methods

Public use data should be available on CDs or an Internet access client server environment. NAACCR provides an annual statistical monograph of cancer incidence in the U.S. and Canada (CINA); an online query system of cancer incidence data (CINA + Online); and a data file for NAACCR groups to conduct cancer surveillance research (CINA Deluxe), all products designed to meet the needs of a variety of potential users.

4.2.4.3. Standards

NAACCR encourages registries to broaden their user base and enhance the accessibility and utility of cancer registry data through electronic publishing.

Registries providing public use data files **MUST** implement specific policies and procedures to protect the strict confidentiality of the data and prevent unauthorized linkages with external files. See Appendix K for an example of an agreement that a user must sign to obtain a NAACCR public use file.

Chapter 5: Data Management

5.1. STRUCTURAL REQUIREMENTS

For cancer registries, the advancement in computer software and hardware has increased efficiency of data collection and improved data quality, standardization, and accessibility. It also has facilitated collaborative pooling of data. Computers have enhanced our ability to more fully use the rich resource of cancer registry data.

The potential of these advances has not been fully realized. Important gains have been made in maximizing the cost-effectiveness of registry operations and the speed and accuracy with which the registry can be used to answer important scientific, clinical, and policy questions. Computers have enabled registry staff to perform more work with the same or fewer resources, as they have been integrated into many aspects of registry structure and operations. As we continue to move into the 21st century, cancer registries face restrictions in resources at a time when the population is aging, causing the number of reportable cases to continue to grow. Cancer registries must employ appropriate applications of computer technology.

Chapter 5 describes specific functional requirements, system design considerations, software and hardware requirements, and other features that are important to fulfilling the functions of a central cancer registry and that any central registry **SHOULD** be able to perform. The words "computer system" or "system" in this section generally refer to the complete system, including the hardware and software (i.e., the equipment and programs). This chapter will not recommend specific software or hardware. The technology will not remain static, and many future advances will be useful to central registries. Thus, it is the goal of this chapter to outline a set of general functional requirements that each central registry **SHOULD** meet, and to encourage every registry to include these functions and to go well beyond them where possible. This chapter specifically addresses central registries at state and provincial/territorial levels, and those central registries at a regional level within a larger central registry system. Requirements for systems at a national level may vary somewhat from those stated here, and these differences are not addressed.

This section does not address general-purpose computer tools such as word processing, accounting, spreadsheets, or desktop publishing, although the central registry will require a wide variety of computer resources beyond those that are addressed in this section.

5.1.1. Overview of Major System Functions

The utility of a cancer registry system **SHOULD** be measured by the ability of a given hardware and software combination to effectively accomplish those tasks assigned to a central registry. A central registry **SHOULD** be designed not only to collect accurate, error-free data, but also to provide appropriate reports, statistics, and data files for researchers, collaborative projects, or national surveillance programs. A registry data processing system **SHOULD:**

- Have the capacity to handle the central registry caseload.
- Provide multiple modes of data interface, including data entry.
- Support appropriate linkage of patient data with hospital and other data.
- Ensure data integrity, completeness, and accuracy.

- Produce standard reports.
- Provide tools for *ad hoc* analyses, lists, and reports.
- Communicate with regional/national data sharing efforts.
- Incorporate appropriate security.
- Be cost-effective and affordable.
- Be dynamic (i.e., easily and inexpensively changed over time).
- Have adequate performance that supports timely data entry, analysis, and reporting.

Registry operations, data management, and data quality rely heavily on software vendor capability and capacity. Software updates should be provided promptly and with pertinent instruction to maximize data capture, completeness, and accuracy. Registries are encouraged to maintain open communications with software vendors to ensure that adequate training and support are available.

5.1.2. Importance of Standards

For reasons of efficiency and comparability, it is important for central registries to adopt existing standards where they exist, and to actually use existing resources in their systems. Idiosyncratic systems are more costly to maintain and enforce hidden costs and sometimes contribute to a problem of incompatible data.

5.1.3. Standards for Functional Requirements

The major functions of a central registry system are listed below.

5.1.3.1. Support for All Registry Activities

The central registry's computer system **MUST** be able to support the efficient and effective execution of all of the tasks in Chapters 2, 3, and 4, including routine operations, analyses, reports, quality monitoring, communications with facilities and providers, etc.

5.1.3.2. Computerized Data Collection

Abstractors employed by the central registry and those in reporting facilities **SHOULD** use computer-based data collection software for abstracting tumor data from source documents. The software **SHOULD** include features such as standard edits (see the discussions of data processing, data quality, and standard edits in Section 3.1.4; adherence to standards in Section 5.1.4.; and EDITS in Section 5.8.).

5.1.3.3. Electronic Transmission

The central registry **SHOULD** require or encourage submission of data, including codes and text, in standardized electronic form, by means of a network, modem, diskette/CD or other electronic media. The central registry **SHOULD** encourage the use of NAACCR's data exchange standard for such transmissions. (See *NAACCR Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

5.1.3.4. New Case File Processing

Files containing new tumor records often are received from hospital-based registries, central registry abstractors, or other sources in machine-readable form. The system **MUST** edit the tumor records, determine their relationship to tumor records already in the database, apply the new tumor records to the database, and retain information on the quantity and quality of data for management reports and quality control. The process described touches on linkage (i.e., determining the relationship to tumors already in the database).

5.1.3.5. Tumor Record Data Maintenance

This function involves updating data on tumor records in the database. The system **MUST** receive changes to tumor records from multiple sources and provide the means (interactively or batch) to edit the data and apply changes to the proper tumor records (see Section 5.9.).

5.1.3.6. Person Versus Tumor

The system **MUST** allow viewing of the data and generation of reports using either the person or the tumor as the basic unit.

5.1.3.7. Reporting

The ultimate goal of a central registry is to use the data for useful information. Chapter 4 outlines reporting requirements in detail. The database management system **MUST** have an adequate subsystem for retrieving files that can be exported into SAS, Excel, SPSS, SEER*Stat, or other analytic software tool. The system **SHOULD** have the capacity to produce both standardized and *ad hoc* reports providing data for administrative management (i.e., registry workload, operations, etc.) in addition to analytical purposes.

5.1.3.8. File Extraction

The computer system **MUST** be capable of producing flat-file subsets of the database for analysis, quality control, data submission, follow-up, or other uses.

5.1.3.9. Quality Control

This function includes tracking the progress of tumor record processing and providing support for all of the quality control activities discussed in Chapter 3. The system **MUST** be equipped to monitor the sources, amounts, types, and quality of tumor record data received and provide management information about how well the source data are captured and transmitted.

5.1.3.10. Online Inquiry

The system **SHOULD** allow retrieval of tumor record data for computer terminal display through specific database keys and user-specified search criteria.

5.1.3.11. Record Linkage

Matching registry data with outside sources is an important method for ascertaining cases and obtaining follow-up on registered cases. A flexible method, or at least the ability to create external files for linkage to death certificate files, drivers' license data files, or other files, **MUST** be included (see Section 5.11.).

5.1.3.12. Follow-Back

The database management system (DBMS) **SHOULD** support management of the death clearance follow-back process and related tasks (see Section 2.2.9.). However, there may be some variation on how registries manage the death clearance follow-back. This might include a separate database that eventually will be used to link DCO) cases back to the master file.

5.1.3.13. Parameter Maintenance

The system **SHOULD** provide easy updating of table variables and denominator data.

5.1.3.14. Administration

Database administration tasks such as backup, disaster recovery, and disk maintenance **MUST** be provided, either by the facility or a third party. Registries should communicate with their IT team to determine optimal solutions for individual facilities.

5.1.3.15. Security

The system **MUST** ensure the integrity of the data and programs and protect the confidentiality of patient, facility, and provider data. A password-protected log-in to the system is highly encouraged as the first line of access to patient level data. Registries **SHOULD** frequently communicate with hospital IT staff to ensure that multi-level security features (firewalls, virus protection, etc.) are in place and operating normally at all times. (See Section 5.1.5.2.3.).

5.1.3.16. Data Sharing

The system **MUST** be able to share all data with other central registries, federal surveillance programs, NAACCR, and other calls for data, such as the National Cancer Data Base. The registry **MUST** use NAACCR's data exchange standard whenever possible.

5.1.3.17. Communications

The system **SHOULD** provide telecommunications capabilities for the import and export of files and interfacing with e-mail programs and Internet providers.

5.1.4. Adherence to Standards

5.1.4.1. NAACCR Data Standards

The system **SHOULD** meet all of the standards specified in *NAACCR Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*, including the required data items collected and their codes and formats (see Section 3.2.1.).

5.1.4.2. Standard Edits

The central registry **SHOULD** use standard data edits (see Section 5.8).

5.1.4.3. Data Exchange Standard

The central registry's system **MUST** be able to read and write files adhering to the most current version of NAACCR's data exchange standards as specified in NAACCR's *Standards for Cancer Registries Volume I: Data Exchange Standards and Record Description.*

5.1.4.4. Analysis Standards

The system **SHOULD** provide the capability to produce analyses using all of the standards described in Chapter 4, including:

- Use of standard analysis categories.
- Application of standard statistical methods.
- Provision for use of multiple population standards.
- Production of standard reports.

5.1.5. Standards for Other System Design Considerations

The following issues **MUST** be considered carefully when choosing or designing a registry system.

5.1.5.1. Performance Requirements

The central registry **MUST** specify performance requirements based on factors of volume, timing, processing requirements, and the anticipated number of simultaneous users. The specific requirements will vary by registry. Generally, the growth in case completeness and case reporting and the required reportable data items are predictable. It is possible to anticipate the amount of disk space the database will require and the amount of computing power required to handle the anticipated number of transactions and reporting load. Interactive response rates will diminish as the database gets bigger and as more users are added to the system. Interactive response times are difficult to estimate, but general performance estimates are available in trade journals and vendor advertising.

5.1.5.2. Internal Control Requirements

Control policies and procedures **MUST** be implemented that provide for accuracy, security, and maintenance of data confidentiality.

5.1.5.2.1. Accuracy of the Information

The central registry will be totally responsible for the accuracy of its information, from entry through subsequent processing, permanent maintenance, and finally to reports. Accuracy of the information can be maintained by:

- Providing extensive editing capability.
- Describing a series of procedures to be followed by the central registry staff to assist with data entry and to ensure that errors detected during the editing process are corrected and that the data are resubmitted for processing.

- Restricting the users who are authorized to access the database to make changes to the data.

5.1.5.2.2. Confidentiality

The registry's computer system **MUST** contain a series of internal procedures to ensure that:

- Access to automated information is restricted to authorized persons.
- Control is maintained over all documents that contain sensitive information to ensure that these documents are available only to authorized persons.
- Requests for information that require personal identifiers are screened to verify that the requestor is authorized to have the requested information (see Sections 2.2.8. and 4.1.1. for detailed discussions of handling confidential data).

The *NAACCR 2002 Workshop Report, Data Security and Confidentiality*, is located on the NAACCR website (www.naaccr.org).

5.1.5.2.3. Security

The registry **MUST** ensure the security of all of the elements of its system, through procedures such as the following:

- Equipment located at the central registry, and possibly elsewhere, **MUST** be protected from theft and from accidental or deliberate damage or misuse.
- Once programs are completed and in routine use, they **MUST** be protected against tampering. Program maintenance **MUST** be carefully controlled.
- Data **MUST** be protected against inappropriate destruction, modification, or dissemination, whether inadvertent or deliberate.
- Annual review of all confidentiality and security operations using the NAACCR Inventory of Confidentiality and Security **SHOULD** be conducted (see Appendix J).
- Procedures for backup, archival, and disaster recovery for both data and programs **MUST** be implemented.
- When staff resign or are terminated, the registry **MUST** change passwords or other security procedures to protect against sabotage.

5.1.5.2.4. Autonomy

Experience has shown that efficiency, responsiveness, quality, and security are enhanced when the registry has control over its own data management system, including the hardware, software, and personnel. The registry **SHOULD** have control over the selection of and use of all hardware, software, and personnel. When resources are shared with other programs or offices, the registry **SHOULD** have control over the priorities and activities such that performance of registry functions is not compromised.

5.1.5.2.5. Funding

The central registry's budget **MUST** provide specified and adequate funding for data management equipment, software, and personnel, including adequate funds for:

- Maintenance of equipment.
- Upgrades of equipment and software for improved performance.
- Implementation of new standards as they become available.
- Implementation of new technologies and software that will enhance the efficiency, effectiveness, and security of the registry and its data.

5.2. HARDWARE REQUIREMENTS

A recommendation for specific hardware is not useful, due to the many local factors that need to be addressed; the speed at which computer technology changes, causing recommendations to become obsolete quickly; and the fact that making specific recommendations is not part of the NAACCR mission. Prior to considering hardware options, the central registry's planners **SHOULD** consider all of the points listed in Section 5.1.1., as well as the following questions:

- What type of operating system would best fit the central registry's situation (i.e., multi-user, single-user, network, etc.)?
- What is the nature of the physical facility where the equipment will be housed, used, and connected?
- What types of software packages will be run on the system?
- How much training will be required for existing central registry staff?
- Does the central registry's parent institution or agency have existing contracts for the purchase and maintenance of computer hardware? Existing contracts and agreements may dictate the types and brands of hardware that may be purchased. This may in turn limit software selections to those designed to run on the equipment the registry is required to use.

5.2.1. Peripherals

Besides the computer itself, some attention **SHOULD** be directed to peripheral equipment for backup of the registry and all operations, printing, and communications using e-mail, the Internet, and fax.

5.2.1.1. Printing and Graphics

Registries **SHOULD** be able to create high-quality reports and presentations. Careful analysis of printing needs is important. There is a sizable difference in cost between quality low-speed and high-speed printers. Printers that produce high-quality output also are more expensive. Depending on the central registry's particular needs, a combination of several types of printers may be appropriate. Color printing capabilities can be very useful when preparing graphs and charts for publications or presentations.

The central registry **SHOULD** consider adding software to develop presentations and/or graphics software to use in data analysis and for displaying data.

5.2.1.2. Communications

In a central registry operation, the quantity of information that needs backup is substantially greater than in a hospital registry, pathology laboratory, or radiation therapy center. In some instances, the responsibility for backup may be assigned to another organization, such as a data processing group responsible for a network server. Most often, the backup responsibility will be the central registry's. The optimum method for backup might include the purchase of additional hardware, such as a tape drive, CD burner, or communications hardware to allow transfer of data to another machine for backup. Central registries **MUST** have procedures in place to recover information from a backup.

The central registry also **MUST** carefully evaluate physical storage needs. A fireproof safe for storing backup files and off-site storage **MUST** be required.

A careful analysis of electronic communication needs will determine hardware requirements. In addition to the anticipated volume of information to be exchanged between the central office and the hospitals or laboratories, communication capabilities will be of value for other reasons. The central registry may benefit from implementation of internal e-mail, and may be able to communicate with a registry's website, bulletin boards, or external e-mail with other organizations around the globe. Basic, low-end data transfer can be accomplished by sending a diskette using mail or overnight service. A more flexible solution for data transfer involves Internet connectivity through high-speed Internet (DSL, cable, etc.). The use of networks can provide the capability for users at different places to be connected to the same system. A combination of these and other options also can be considered.

The selection and purchase of hardware usually **SHOULD** be one of the last decisions made, because the selection of operating systems, database management systems, and other commercial products will limit some of the hardware options. If hardware is selected first, the central registry may be severely limited in its software selections.

5.2.2. Life of the Hardware

The useful life of computer hardware is only a few years due to rapidly changing computer technology. The registry **SHOULD** build in hardware replacement costs every 3 to 5 years.

5.2.3. Standards

The registry **MUST** have computer hardware resources that are adequate in type and amount to support all of the central registry's required activities, including data collection, database management, quality control, analysis, and reports.

The central registry's hardware **MUST** adequately protect the accuracy of registry data and **MUST** have security features adequate to protect the confidentiality of data and security of the system.

5.3. SOFTWARE REQUIREMENTS

5.3.1. Database Overview

Database technology allows registry data to be processed as an integrated unit. It reduces the artificial barriers imposed by separate files for separate applications and permits users to access data more naturally. Database

processing provides several advantages. First, it eliminates or drastically reduces data duplication. Elimination of duplication saves storage space and frequently can reduce processing requirements. Because there is only one location for each piece of datum, update anomalies can be avoided and data integrity improved. Secondly, because all programs interface with the database through a DBMS, a uniform definition of a data item is used.

Database processing requires increased program and data overhead. Thus, database applications often require more powerful hardware in the form of more main memory, processing speed, and larger, higher-performance storage devices.

There are many DBMSs available today based on hierarchical, relational, and other conceptual designs. Many DBMSs are hardware-independent and will operate on a variety of hardware and operating system platforms. This allows the software to be moved to different types of computers with little or no reprogramming. Reduced dependence on a single vendor's hardware can have a large economic advantage when system changes are under consideration.

In general, the DBMS **SHOULD** be able to:

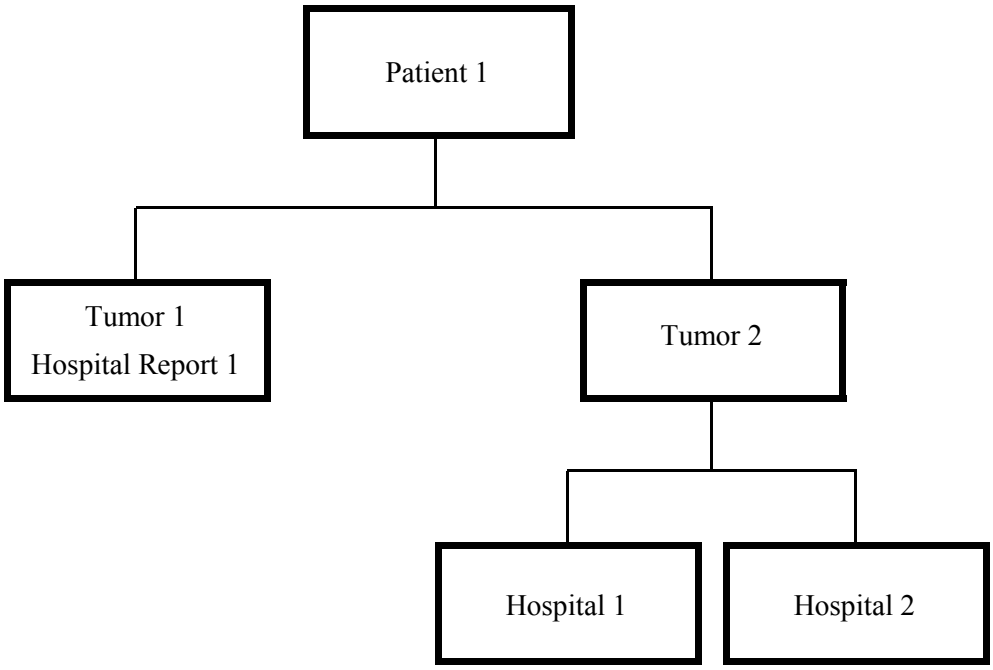
- Define and store specific information about the database structure.
- Provide a wide variety of methods for accessing data.
- Store and maintain data in a manner that maintains relational integrity.
- Provide security features to protect access to the data.
- Enable control over concurrent operations.
- Facilitate backup and recovery.

Appropriate database design and selection of a good DBMS are essential to providing efficient means for accessing the data and providing an adequate system throughout.

5.3.2. Hierarchical Model

Many central registries have used the hierarchical database model successfully. It allows a record type to be a member of only one relationship, which is referred to as parent-child. It is often diagrammed as a tree structure, as shown in Figure 1.

Figure 1. Database: Hierarchical Model



5.3.3. Relational Model

The relational database model views the database as a set of two-dimensional tables or relations. Each table represents an entity (person, place, or thing) and its relation to the other tables. The columns of a table correspond to data fields, and the rows correspond to record occurrences. Each table will have a unique primary key (PK). Relations between tables are accomplished by placing primary keys in other tables as foreign keys (FK). Figure 2 is a sample of a simple relational model. In this example, the Patient Table contains summary records for each person; the Tumor Table contains summary records for each tumor (there can be more than one tumor per person); and the Hospital Report table contains data from the individual abstracts submitted (there can be more than one abstract per tumor). This tabular view of a relational database, if done properly, clearly shows all of the entities, their attributes, and relation to all of the other entities. By testing this self-documenting model with sample data, it is possible to see if the database has been designed correctly. The tools for updating, inserting, deleting, and querying a relational DBMS vary widely by vendor. Because of its extreme importance for both the developer and the end user, careful consideration **SHOULD** be given to choosing a DBMS that has the capacity to fulfill current and, where they can be determined, future central registry needs.

Figure 2. Database: Relational Model

Patient

Patient ID Number	Name
PK	
1 2 3	Smith Jones Doe

Tumor

Patient ID Number	Tumor Record Number	Primary Site	Morphology
FK			
1 2 2	1 1 2	Breast Lung Larynx	850039 804239 807033

See also: Patient

Hospital Report

Patient ID Number	Tumor Record Number	Report Number	Hospital ID
FK	FK		FK
1 2 2 2	1 1 1 2	1 1 2 1	1 1 3 2

See also: Patient
Tumor
Hospitals

Hospitals

Hospital ID	Name
PK	
1 2 3	Mercy General VA

The central registry **MUST** carefully consider how the DBMS would interface with its application programs. Application programs are the primary software programs that have been designed to facilitate data entry, editing, error checking, reporting, and analysis. They are specific to the needs of the central registry and very often are written by registry support staff or the vendor providing the system. The ease with which application programs can access the database and make changes to it while maintaining data integrity is an important consideration when choosing DBMS software.

Another consideration is that most central cancer registry systems **SHOULD** be able to provide concurrent, multi-user access to the database. A key issue in a multi-user database application is data integrity. Ensuring that no data are lost and that data are logically consistent (i.e., the inter-data relationships that should exist do exist) usually is performed through three features:

- Some form of record and file locking.
- Transaction processing.
- Transaction logging and recovery.

A multi-user application **MUST** provide a mechanism whereby updates to shared data can be synchronized so that only one user can be updating the shared data at a time. This mechanism usually is some form of a lock that is used to serialize updates to data shared among multiple users. A lock is required to update shared data; it will prevent other users from updating the locked data. When considering various DBMS software, it is important to understand whether or not the DBMS software automatically handles the locking or if the developer is responsible for coding the locks.

Transaction processing is used to maintain the logical consistency of a database by allowing multiple, related updates to be grouped together and written to the database as a unit at the end of the transaction. Without transaction processing, changes are written to the database as they occur. If a failure occurs in the middle of a series of changes, the database could be left in an inconsistent state.

Transaction logging is used to provide data integrity protection in the event of a failure occurring while transaction changes are actually being written to a database. The process by which the changes are written to the database is called a transaction commit. Failures that occur during a commit are detected by the system, automatically initiating a recovery operation that ensures that changes to the database actually occur.

5.3.4. Utility Programs

Utility programs are software tools used to aid in the more general functions of an information system. Sorting, backup and recovery, and data import and export are just a few of the functions served by this class of software. Many utility programs will come with the operating system, and others will be part of a good DBMS system. Designers are encouraged to pay close attention to the utilities provided by these various sources.

Indexing utilities are provided either by the operating system or by the DBMS software. Effective use of database indexes and key fields can provide faster retrieval of computerized records identified by any of several attributes. There is processing overhead to maintain indexes, but the increased speed of retrieval, especially interactive retrieval, makes their use very attractive, if not mandatory.

5.3.5. Communications

The capacity for electronic communications within the central registry and with external sources is very important. Communications ability is dictated by the hardware platform and software selected. Good DBMS systems will enable database links to external computers. Most operating systems also will support connections to wide area networks and the Internet. Zip drives and CDs sent by mail, although slow, are capable of handling large volumes of data. However, large file transfers over broadband Internet connections are just as secure and faster than postal or courier services. File transfer protocol (FTP) transfers through the Internet can handle the larger files and can be transferred very quickly.

Data security **MUST** be maintained when any of these external links are used. This may include mailing or shipping of files or reports through a method with an audit trail (see Sections 2.2.8. and 5.3.6.3.).

5.3.6. Standards

5.3.6.1. General

The central registry **MUST** have software resources adequate in type and amount to support all of the registry required activities, including data collection, database management, quality control, analysis, and reports.

5.3.6.2. Use by All Registry Staff

The system **MUST** allow routine processing and database access by all registry personnel.

5.3.6.3. Security

The system's software **MUST** include features to adequately protect the accuracy of registry data and **MUST** have security features adequate to protect the confidentiality of data and security of the systems (see Sections 2.2.8. and 4.1.1.).

5.3.6.4. Standards

The software **SHOULD** incorporate standard edits, standard analysis categories, and other standard analysis tools (see Section 3.1.4., 4.2.1., 4.2.2., 4.2.3., and 4.2.4.).

5.3.6.5. Shared Software

Where possible, the system **SHOULD** incorporate subsystems and features that can be transported from systems already in use rather than requiring new development.

5.3.6.6. Graphics and Mapping

The central registry **MAY** have software for preparing graphs of the registry's data, and if not, then the registry **MUST** be able to export data in a format that can be imported into graph and map software (see Sections 4.2.3.3. and 5.5.2.7.1.).

5.3.6.7. Standard Statistical Software and Analytic Epidemiology Software

The central registry **SHOULD** have a statistical software package available to perform standard statistical calculations. A registry **SHOULD** have and use SEER*PREP and SEER*STAT (www.seer.cancer.gov), or other comparable software, for producing routine surveillance statistics. For more specialized epidemiologic analyses such as cluster analysis, cohort analysis, or modeling, the registry also **MAY** need specialized analysis software (see Section 4.2.2.).

5.4. STAFFING GUIDELINES

The computer and data management staff at the central registry are in a crucial position to influence the overall success of the registry. The lead computer staff person **SHOULD** be considered a part of the central registry's leadership and **SHOULD** be involved in planning and overall system design.

5.4.1. Standards for Number and Type of Staff

The central registry **MUST** provide data management staff sufficient in number and training to assure compliance with mandated reporting requirements, assure timely completion of all required tasks and reports,

and meet all other standards. It is desirable that the data management staff have a background in health applications as well as requisite technical knowledge.

Central registry personnel **MUST** be sufficiently trained and cross-trained in the operation of the system to protect against the possibility that the loss of a single person would cripple its operation.

5.4.2. Continuing Education

Continuing education **SHOULD** be provided to data management staff to assure that they have up-to-date knowledge about available technologies and cancer registries. Courses and workshops offered by NAACCR, NCRA, and other local, state, provincial, and national organizations can provide excellent training opportunities (see Appendix E for resources for education and training for providers and users of central registry data).

5.4.2.1. Access to Professional Literature, Online Services, and Other Activities

Data management staff **SHOULD** be supplied with appropriate references and literature to provide ongoing continuing education and to answer questions that arise. Current pertinent reference books, journals, and other periodicals **SHOULD** be available immediately. The central registry also **MAY** provide access to online services and bulletin board services so that staff have rapid access to the most current information.

5.4.2.2. Professional Associations and User Groups

Data management staff **SHOULD** be encouraged and funded to participate in local and national professional associations and user groups pertinent to their technical area, and also in registry-oriented scientific meetings. The central registry budget **SHOULD** include funds for participation by one or more persons at scheduled meetings. The central registry **SHOULD** fund data management staff to attend special symposia, conferences, and courses that may be offered from time to time (see Appendix E for addresses and organizations cited in this report).

5.5. PROCESS STANDARDS

5.5.1. Data Entry

Electronic reporting **SHOULD** be the method used for data collection. Data entry of tumor records is most often part of the process of abstracting directly onto a computer. Computerized data collection combines abstracting, coding, data entry, editing, and accessioning into one process. Some central registries provide software to reporting facilities to standardize this process. In addition, the central registry probably will employ a variety of data entry methods for some new tumor records; for corrections, deletions, or other transactions; or for physician and hospital data. These methods can include direct keying from source documents into the computer, key entry from data collection forms, the use of imaging software to scan abstracts, and other methods. Regardless of the methods used, some form of verification **SHOULD** be in place.

When electronic reporting is not possible, the central registry **SHOULD** implement some form of verification of keyed data to minimize entry errors. The method will vary with the data entry method, and may include visual comparisons, duplicate keying when manual forms are used, extensive editing and analysis of input data, or other quality reviews (see the discussion of edits in Section 5.8.; see Section 3.1.4.3. for a discussion of the importance of standardization of aspects of data entry to improve data quality).

5.5.2. Inputs

A central registry **MUST** be prepared to process cancer-related data collected in various forms from a variety of sources. These sources **MAY** include health care facilities; nursing homes; physician's offices; coroners' offices; state vital statistics departments; other local, state, and federal governmental agencies; other central registries; and outside vendors. Most of the data received by the central registry are of the following types:

- New tumor records to be added to the central database.
- Follow-up, correction, and deletion data from reporting facilities to be applied to previously collected tumor records.
- Data from sources other than reporting facilities to be applied to previously collected tumor records (death information and geocodes).
- Tumor records with limited information from sources such as physicians, outpatient surgery and radiation centers, pathology laboratories, or rapid case ascertainment reports from special studies.
- Other data to be applied to the central registry database include parameter file updates and population data for rate calculations.

5.5.2.1. Standards for General Input File Specifications

The central registry **SHOULD** adopt the following specifications for all input files to the central registry:

- As specified in *NAACCR Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*, the data files **MUST** be standardized in terms of data items, codes, and record layout.
- The data files **SHOULD** be submitted in machine-readable form and transmitted to the central registry from all reporting sources through network, modem, FTP link, Internet e-mail, or on diskettes. The central registry **SHOULD** be able to key in data from hardcopy forms when this is the only reporting source available from the reporting facility.
- The data **MUST** contain an appropriate level of patient and tumor identification, ranging from central registry case numbers or hospital chart numbers to personal identifiers.

5.5.2.2. Standards for New Case Data Input Files

5.5.2.2.1. Definition

These files include data pertaining to:

- Patient demographic characteristics.
- Patient confidentiality.
- Hospital specifics.
- Other confidential resources (i.e., linking with DMV, Medicare, etc.)

- Cancer identification.
- Stage.
- Prognostic factors.
- Treatment.
- Follow-up for each tumor.
- Supporting text.

5.5.2.2.2. Required Processing Functions

The central registry's data management system **SHOULD** have the capacity to perform the following functions regarding new case files:

- *Editing:* Apply standardized edits to new case files and provide the ability to reject individual records with errors and reject the submitted file if the error rate is above a threshold level and unacceptable for processing. (see Sections 3.1.4. and 5.8.)
- *Error Correction:* Produce indications of errors (printed or screen reports or other indications) to inform quality control staff and allow correction of case data.
- *Global Changes:* Provide the ability to mass-correct global errors in incoming case files.
- *Deletion:* Provide the ability to delete records from the input file.
- *Management Information:* Provide tracking information and appropriate management reports on the number of tumor records submitted by reporting facility, by time period, and by diagnosis year, as well as the number and types of errors (see Section 3.2.5. for a discussion of quality control activities).
- *Printed Abstracts:* Produce standardized printed abstracts with text and coded data presented in natural language as well as the coded data.
- *Phonetic Compression Index:* Provide a phonetic compression such as SOUNDEX or NYSIIS for last name, maiden name, and aliases. Indexes built on phonetic compression will facilitate record linkage.
- *Linkage and Accessioning:* Provide the ability to match incoming new tumor records with existing records in the database to identify duplicate or subsequent records or previously unreported tumors and to assign unique accession numbers to the new tumor records (see discussion of record linkage in Section 5.7.).
- *Reports to Reporting Facilities:* Provide the following to reporting facilities:
  - Reports detailing questions that arise during attempted correction of case data (e.g., edit failures that cannot be corrected at the central registry for lack of information).

- Reports indicating the status of data submissions, with items such as numbers of tumor records sent in this transmission and year-to-date, by diagnosis year, and estimated percent complete based on anticipated caseload.
- Reports analyzing number and types of edit errors (see Sections 3.1.4. and 3.2.5.).

5.5.2.3. Standards for Follow-Up Data Input Files

5.5.2.3.1. Definition

When a central registry collects patient follow-up from reporting facilities (e.g., hospital cancer registries following their hospitals' cancer patients) and a facility updates the date of last contact, vital status, or tumor status of a patient, that information **SHOULD** be sent to the central registry to update the central registry database (see Section 5.5.2.6. for a discussion of death information input file processing).

5.5.2.3.2. Required Processing Function

The central registry system **SHOULD** have the capacity to perform the following functions regarding follow-up input files:

- *Linkage:* Provide the ability to link an incoming follow-up record with the appropriate database tumor record.

- *Editing and Automatic Updating:* Provide the ability to automatically apply an incoming follow-up record to the database tumor record, when appropriate, after editing for compatibility and consistency (see Section 5.10. for updating guidelines).

- *Error Reports:* Produce error reports for incoming follow-up records failing edits.

- *Management Information:* Provide the means to identify database tumor records where follow-up information has been changed and provide appropriate management reports.

5.5.2.4. Standards for Correction Data Input Files

5.5.2.4.1. Definition

In addition to its own correction procedures for individual records, the central registry **MAY** receive files of corrections from reporting facilities that have made changes to previously reported tumor records. These files contain the changes made to required data items after the tumor information has been transmitted to the central registry.

5.5.2.4.2. Required Processing Functions

The central registry system **SHOULD** have the capacity to perform the following functions regarding correction input files:

- *Linkage:* Provide the ability to link an incoming correction record with the appropriate database tumor record.

- *Editing and Updating:* Provide the ability to either manually or automatically apply an incoming correction record to the corresponding database tumor record after editing for intrafield and interfield consistency (see Section 5.10. for updating guidelines).

- *Error Reports:* Produce error reports for incoming correction records failing edits.

- *Management Information:* Provide the means to identify database tumor records where information has been changed and provide appropriate management reports.

5.5.2.5. Standards for Deletion Data Input Files

5.5.2.5.1. Definition

This file contains information on previously reported tumor records that were deleted by the local registry.

5.5.2.5.2. Required Processing Functions

The central registry system **SHOULD** have the capacity to perform the following functions regarding deletion input files:

- *Linkage:* Provide the ability to link an incoming deletion record with the appropriate database tumor record.

- *Reports:* Produce reports from incoming deletion records containing patient identifiers and reason for deletion.

- *Manual Processing:* Provide the ability to manually delete a database tumor record.

- *Management Information:* Provide the means to identify deleted database tumor records and provide appropriate management reports.

- *Restore:* Provide the ability to restore a tumor record mistakenly deleted.

5.5.2.6. Standards for Death Clearance Input Files

5.5.2.6.1. Definition

Death clearance processing involves use of data about residents for whom death certificates were filed. The purpose is to provide new information about previously reported tumor records (follow-up) and to obtain new tumor record information for previously unreported patients or cancers (follow-back).

5.5.2.6.2. Required Processing Functions

The registry system **SHOULD** have the capacity to perform the following functions regarding death clearance input files:

- *Linkage:* Provide the ability to link an incoming death certificate record to the appropriate database case.

- *Editing and Updating:* For death certificate records that link to database tumor records, provide the ability to automatically apply the incoming death information to the database tumor record, when appropriate, after editing for compatibility and consistency; and to update other items coded in the death record, such as race and birthplace, when the database tumor record contains unknown or non-specific values and the death record is more specific (see the discussion of consolidation in Section 5.9.).

- *Error Reports:* For death certificate records that link to database tumor records, provide error reports on records failing edits.

- *Suspense:* For death certificate records that do not link to database tumor records but are tumor records that should have been reported, provide the ability to suspend the death records in the database for further follow-back investigation (see Sections 2.2.9. and 5.5.2.8.).

- *Management Information:* Provide the means to identify tumor records where death information has been applied to the tumor record or entered in a suspense file and provide appropriate management reports.

5.5.2.7. Standards for Geocoding Input Files

5.5.2.7.1. Definition

This file contains geographic data—usually census tract information, block, or other small area—for tumor records in the database. Polar coordinates also may be assigned for mapping use. The address at diagnosis of the patient is used to determine the appropriate census information, usually through an automated matching procedure, with some addresses requiring manual processing. Many central registries perform geocoding as a batch process relatively infrequently, sometimes using a commercial vendor (see Section 5.3.6.6 and the NAACCR document *Using Geographic Information Systems Technology in the Collection, Analysis, and Presentation of Cancer Registry Data*).

5.5.2.7.2. Required Processing Functions

The central registry system **SHOULD** have the capacity to perform the following functions regarding geocoding input files:

- *Linkage:* Provide the ability to link an incoming geocoded record with the appropriate database tumor record.

- *Editing and Updating:* Provide the ability to automatically apply the geocoded data to the database tumor record, when appropriate, after editing for compatibility and consistency.

- *Error Reports:* Produce error reports for incoming geocoded records failing consistency edits.

- *Management Information:* Provide the means to identify tumor records where geocoded information has been applied to the tumor records and provide appropriate management reports.

- *Canadian Geocoding Procedures:* In Canada, Statistics Canada provides each provincial registry with a user-friendly version of the Postal Code Conversion File, which enables registries to automatically convert most postal code information to census geographic units, including census tracts (neighborhood areas), census agglomerations, and census metropolitan areas (large urban centers), as

well as census divisions and census subdivisions. Special procedures are used, including generation of reports, for a small percentage of tumor records that cannot be directly converted, so that consistent and valid codes may be applied.

5.5.2.8. Standards for Limited-Tumor-Information Input Files

5.5.2.8.1. Definition

These files contain limited information about tumor records. The tumor may not have been reported because it is not yet complete (e.g., a tumor identified through rapid case ascertainment); the tumor record may have been ascertained from a source with limited information, necessitating follow-back to other sources (e.g., a tumor identified through a pathology laboratory); or the tumor may have been overlooked by the facility responsible for reporting it (e.g., a tumor identified through death clearance).

5.5.2.8.2. Required File Processing Functions

The central registry system **SHOULD** have the capacity to perform the following functions regarding limited-tumor-information input files:

- *Editing:* Edit the incoming data for very basic content.
- *Suspense:* Provide the ability to suspend the tumor records in the database for further investigation.
- *Reports:* Provide reports of the suspected tumors according to the source to which they need to be followed back and prepare inquiries to the appropriate sources.
- *Linkage:* Provide the ability to periodically link the limited information records with the database tumor records so that the limited information records can be deleted if the tumor records have been added to the database from another source.
- *Deletion:* Provide the ability to delete a limited information record if the tumor is found to be non-reportable.
- *Management Information:* Provide the means to identify disposition of limited information tumor records and provide appropriate management reports.

5.5.2.9. Standards for Parameter File Updates

5.5.2.9.1. Definition

These files contain changes or updates to parameter files used for batch and online editing and other system functions, including table variables and population denominator files (see Section 4.1.2.). Examples include tables of valid race codes with their natural language meanings, and tables of reporting facilities with their reference dates.

5.5.2.9.2. Required File Processing Functions

The system **MUST** provide the means to input files; update the appropriate edit tables; and receive online additions, changes, and deletions to parameter tables.

5.6. OUTPUTS

5.6.1. Introduction

In addition to analytical reporting covered in Chapter 4 and input processing covered in Section 5.5.2., the central registry's computer system **SHOULD** be able to provide several different types of outputs:

- Management reports that allow for monitoring of the database and central registry operations.
- Standard reports to give feedback to or request information from reporting sources.
- Output that responds to *ad hoc* queries from quality control operations, management staff, and others.

5.6.2. Standards for Management Reports

The central registry **SHOULD** produce management reports with a frequency that will enable monitoring the operations of the registry. Examples of possible reports include:

- A table presenting the number of tumor records reported for each reporting facility and for other sources of tumors (such as DCO cases, or physician-only cases) by month and year of admission (or, for DCO cases, month and year of death.)
- A table presenting the difference between the number of tumor records expected from each reporting facility and the number received. By ordering the table in descending order with the facility with the largest deficit on top, this report helps to allocate registry resources to the area with the greatest impact.
- A table presenting the tumors from all reporting sources by month and year of diagnosis.
- A table presenting the distribution of tumors by year of diagnosis by site for comparison with other registries.
- A table presenting the number of tumors by process completed (e.g., number inspected or visually reviewed, number in suspense, etc.), by date received in the central registry to monitor workflow.
- A table showing the interval between diagnosis date and date abstracted, and between diagnosis date and the date the tumor record was entered in the central registry system, by facility to show timeliness of abstracting.
- Tables showing the status of follow-up by facility and by diagnosis year, and for subpopulations of interest (e.g., specific age groups) for central registries collecting patient follow-up.

Other possible reports are described throughout Chapters 2 and 3.

5.6.3. Standards for Reports to Facilities

The central registry's data processing system **SHOULD** enable a variety of routine reports for all facilities submitting tumor records to the registry. These reports can be transmitted to the facilities electronically or in hardcopy form (see Section 4.2.3. for more detailed discussions of types of reports).

5.6.3.1. Reports for Monitoring Workflow and Completeness

To provide information to the reporting facilities about their caseload, or about their reporting completeness, reports such as the following are useful:

- Immediate or very rapid acknowledgment of the central registry's receipt of a tumor record submission (e.g., date received, number of tumor records received), so that the facility can verify that its tumor records were received and were readable.

- A table presenting the number of tumor records from that facility by month and year of admission.

5.6.3.2. Comparison Data

The central registry's system **SHOULD** have the capability to produce appropriate reports of comparison data (described in Section 4.2.3.) for facilities to use in their own registries' annual reports.

5.6.3.3. Requests for Information From Facilities and Physicians

The central registry computer system **SHOULD** facilitate requests for additional case-specific information from the reporting facilities by generating reports such as the following:

- Computer-generated letters addressed to the facilities or physicians requesting patient-specific information for death-certificate follow-back.

- Computer-generated letters addressed to physicians requesting information on tumors identified through screening of pathology laboratory reports where the patients were not seen in reporting facilities.

- Computer-generated letters to facilities, physicians, and patients requesting follow-up, and computer-generated letters including lists of patients to hospitals requesting follow-up information (when follow-up is performed by the central registry).

5.6.4. Standards for *Ad Hoc* Queries

The system **MUST** allow for easy routine querying of the database by management and quality control staff at the central registry, without programmer intervention.

The results from *ad hoc* queries may take the form of interactively displayed reports on the screen or printed output.

5.6.4.1. Listings

The system **SHOULD** be able to provide listings of records in the database that meet specified criteria and are sortable by the user. On a screen display, the user **SHOULD** have the ability to scroll through the rows. As an example, in resolving linkage problems manually, it often is necessary to query the database using alternate spellings, phonetic compression, or incomplete values for given fields and to review the records retrieved.

5.6.4.2. Patient-Tumor-Admission Displays

The system **MUST** be able to display all the data values that are stored for a specific patient, tumor, or admission.

5.6.4.3. Frequencies

The system **SHOULD** allow easy output of frequencies or counts by any variable or combination of variables. To prevent users who do not fully understand the organization of the data from obtaining misleading results, it is useful to require that the user provide answers to a series of questions before the count is generated, specifically:

- Should the results be limited to a certain time period?
- Should the results count patients, tumors, or hospital reports?
- Should the results include *in situ* diagnoses, invasive diagnoses, or both?
- Should the results be limited to residents of the registry's coverage area?
- Should DCO cases be included?

5.7. RECORD LINKAGE

When data are added to the central registry's database, whether adding data to an existing record, or adding new records, a suitable record linkage mechanism is needed to assure that the additional data are correctly associated with the existing data. If a record is added to the database without adequate checking for redundancy, case counts may be overestimated because multiple institutions may report a single tumor. However, if efforts to prevent duplicate records are overzealous, then truly distinct records can be linked together mistakenly, resulting in undercounting of cases.

In any of the above situations, the probabilities of falsely matching records increases, diminishing the quality of the database and resulting in incorrect incidence rates. Statistically speaking, an erroneous record linkage increases the type I and type II errors that are associated with it (the probability of accepting a match given it is the wrong match and the probability of rejecting a match given it is a true match, respectively).

5.7.1. Types of Record Linkage

A record linkage can be performed deterministically or probabilistically.

A **deterministic** record linkage involves the comparison of two records on several key fields (e.g., social security number, last name, first name, etc). A match is achieved if and only if all of the key fields coincide on both records. Any linkage is suitable for records with no errors or missing data.

A **probabilistic** record linkage also involves the comparison of two records on several key fields; however, a probability is associated with a correct and a false match. This usually is achieved by building a scoring algorithm based on the number of fields that coincide in both records and the degree of trust in these fields. In essence, this type of linkage assimilates an individual's thought process if the linkage were to be performed manually. At the same time, it allows assessing a degree of trust in the linkage.

5.7.2. Linking Patients Versus Linking Tumors

The key fields used for the record linkage should be analyzed before use to ensure that they are reliable. Items such as name, sex, social security number, phonetic comparison indices, date of birth, or county of residence can be used for record linkage at the patient level. Additional information, such as address, can be used for questionable linkages that need to be reviewed manually.

Multiple submitted tumor records for the same patient also need to be linked. Records that describe the same tumor must be identified so that they can be consolidated; records describing separate tumors for the same patient need to be stored as separate cases. The task of tumor consolidation is harder to fully automate; it involves comparisons of primary sites, histologies, and the dates of diagnoses. Complications that have to do with assigning morphology to a tumor and the ambiguous rules in determining the date of initial diagnosis can make this procedure cumbersome and may require more manual intervention (see Section 5.9.).

5.7.3. Software

Commercial record linkage software is available, and several registries have created their own software for that purpose. Some commercial packages provide a score that reflects the degree of certainty for a possible linkage and allow for the manual review of questionable linkages. The selection of key fields and compilation of the algorithm are determined by the user.

5.7.4. Standards

The central registry **MUST** have an effective record linkage system for linking patients and tumors. Record linkages can be done manually, by computer, or by a combination of both. Small and well-funded central registries can afford the employees necessary to manually link their tumor records. However, for large or under-funded central registries, this is an impossible task.

Although currently there are no standards established for linkages, reference should be made to reports of the NAACCR's Record Consolidation Committee (see current list of reports and tools on the NAACCR website, standards section, at www.naaccr.org).

5.8. EDITS

Computer edits are a key aspect of the central registry's overall computer system. Quality control edits are discussed in Section 3.1.4.4. The standard NAACCR edits are included in the electronic NAACCR EDITS metafile and can be downloaded from the NAACCR website (www.naaccr.org).

5.8.1. Standards

The central registry system **SHOULD** employ a complete set of standard edits (EDITS metafiles) to evaluate a registry database on file. Edits **SHOULD** be applied as physically close to the information source as possible, and as temporally close to the collection of the data as possible. In addition to a standard edit set, central registries participating in a call for data (e.g., NAACCR Call for Data) are required to employ a call for data edit set (e.g., NAACCR Call for Data metafile) prior to file submission.

Edits **MAY** be performed interactively, as a batch process, or both, and **MUST** be applied at several points in the data flow to:

- Tumor records, before submission to the central registry.
- Newly submitted tumor records before they are linked against the central registry database.
- Database tumor records after linkage.
- Database tumor records after any changes have been made.

The central registry's edits **MUST** allow for override flags for situations in which the edit identifies a rare condition that needs review but may be correct. The override flag prevents the condition from continuing to be identified as an error.

In error reports and discussions with abstractors and coders, it **MAY** be helpful to label data failing edits as "inconsistencies" rather than "errors," because the data are not necessarily incorrect.

5.9. RECORD CONSOLIDATION

Consolidation refers to the process of reconciling or compiling data obtained from more than one source on the same person or tumor. The sources can include multiple abstracts from hospitals, clinics, or other providers, or they can include information from the death records or from other registries. Values for the same data items for the same patient and tumor may be identical from each source, but they also may be contradictory or complementary. A large task of the central registry system is to prepare a composite set of values for each patient and tumor, incorporating information from a variety of sources. This composite set of values then can be stored and managed in a variety of ways, either as a separate consolidated record, or with the individual values in different records flagged as those to be used for the consolidated record. In any case, the original records always **SHOULD** be kept intact.

It is important to recognize the difference between record consolidation and the identification of multiple tumors for the same patient. Again, refer to the standards section on the NAACCR website for record consolidation resources (www.naaccr.org).

Examples

Hospital A:

> SMITH JOHN FITZGERALD
> 2/10/27
> Social Security Number: not recorded
> Carcinoma of colon, diagnosed 3/93 by biopsy elsewhere, treated at Hospital A by sigmoid colectomy on 4/15/93

Hospital B:

> SMITH JACK
> 10/2/27
> Social Security Number: 123-45-6789
> 2/6/93, biopsy, sigmoid colon, showing adenocarcinoma

Death Certificate:

SMITH F. JOHN
2/10/27
Social Security Number: 123-45-6789
Date of death: 9/12/93, Cause of death 153.9

Record From Neighboring State/Province/Territory Registry:

SMITH JOHN F
2/10/27
Social Security Number: not recorded
5/1/93, colon cancer with extensive node mets diagnosed 2 months ago. Seen at Major University Med. Center for chemotherapy, begun 5/5/93.

Once the linkage process has determined that the four records above are for the same person and tumor, the central registry needs a mechanism for categorizing this case, as follows:

Social Security Number:	123-45-6789
Date of Birth:	2/10/27
Date of Diagnosis:	2/6/93
Primary Site:	Sigmoid colon
Histologic Type:	Adenocarcinoma
Date Treatment Began:	4/15/93
First Course of Treatment:	Surgery, chemotherapy
Age at Diagnosis:	66
Survival Time:	7 months

The system also needs to determine the correct name and date of birth, or select a name and date of birth to be used in further linkage, analysis, and reporting. For some variables, especially those used in patient linkage, it is desirable to store all different values obtained for the patient, so that future linkage attempts are more likely to be successful. For other items, especially those related to tumor characteristics (primary site and histologic type) or those used for subsequent calculations (dates of birth and diagnosis for calculation of age and survival time), it is important to establish one value to be used in analysis. It has been helpful to some central registries to separately store all of the values that were submitted, so that the system can reproduce a record as originally submitted by a facility.

5.9.1. Standards

Standards for item-specific consolidation rules, either for computer application or manual application, have not been developed, but many existing systems can be used as models. Some general principles can be stated (see Section 5.10.):

- Where it can be ascertained (in a cost-effective manner), the best, or true, value for each item is the one that **SHOULD** be retained.

- The system **SHOULD** perform automatic consolidation whenever possible, and produce a report of the computer's actions for manual review, but also **SHOULD** be able to identify instances where the computer cannot determine the correct value.

- Known values are preferred over unknown values, and more specific values are preferred over less specific values. However, this rule should be applied with caution, because existing but non-valid

values are no better than missing values. For example, a social security number of 111-11-1111 is no better than a missing social security number and can create problems in linkage projects.

5.10. GUIDELINES FOR PROCESSING FOLLOW-UP, CORRECTION, AND DELETION TRANSACTIONS

The central cancer registry database is dynamic; the data are never final and a data set is never really closed or frozen in time. Tumor records continuously are added, changed, and deleted as long as the registry continues, even after patients have expired and the data have been included in reports. The central registry's system will need to process follow-up, correction, and deletion transactions. NAACCR has added two record layout types that can be used to transmit corrections for follow-up to data already submitted. The Update/Correction record (record type U) is a short format record, and the Modified record (record type M) transmits the entire tumor record (see *NAACCR Standards for Cancer Registries Volume I: Data Exchange Standards and Record Description*). Good data collection software for abstractors automatically will generate correction records for the central registry when changes are made to the local database. Ideally, the central system should handle these automatically; however, some problems arise when conducting automatic updating, especially when combining data from multiple hospitals and multiple software systems.

5.10.1 Potential Problems With Automated Updates

5.10.1.1. Keeping the Different Datasets Synchronized

If review of hospital data is performed centrally and some data items are accepted and others not accepted or modified, then the hospital needs to be notified about the changes and strongly encouraged to accept the modifications. Otherwise, as the two datasets (hospital registry and central registry) diverge, the quantity of information requiring review gets very large, and the central registry will repeatedly review information that already has been reviewed. Some software systems generate datasets of corrections to transmit to the central registry that include the entire tumor record but do not identify the specific field(s) that was (were) updated. When software generates updates that identify only the fields that have been updated recently, the quantity of reviewing can be greatly reduced. This requires extensive cooperation with all of the software suppliers.

In any case, it is very important to have a mechanism for reporting back to the hospitals any updated information or modifications to their data. This can be done through a printed report or automatic update file.

5.10.1.2. Software Differences

Different hospital software systems can create discrepancies that require review. When hospitals reporting to the central registry use the same software, the problems can be simplified, and the quantity of changed tumor records requiring review is smaller. In those instances where different software packages are being used, the central registry will have to develop procedures to reduce the reviewing workload.

5.10.1.3. Data Ownership

Proprietary ownership of data between the hospital and central registry is complex. See Sections 2.2.8.3. and 4.1.1. for policies and procedures for release of registry data.

5.10.1.4. Standards

NAACCR standards for correction and deletion transactions have not been established. The central registry needs to seek a balance among quality level, resources, money, and time to best reach the goals of the

registry. There are no simple answers, and there is no single solution to all problems. Procedures **MAY** vary by the type of data being changed.

5.10.1.5. Follow-Up Items

Items such as Date of Last Contact, Vital Status, Tumor Status, and Autopsy, can be handled easily by the computer and generally cause few problems for review and/or quality control. Occasionally, an inconsistency occurs, such as a death date reported as earlier than a reported date the patient was alive. The computer easily identifies these, and the transactions can be printed out for review and resolution by quality control staff.

5.10.1.6. Changing From an Unknown Value to a Known Value

Changing from an unknown to a known value for items such as Zip Code or Race easily can be handled by computer. There are a few items that **SHOULD NOT** be automatically updated, such as Cause of Death, because this represents the official cause of death as assigned by the vital statistics agency. These transactions easily can be printed out for review by quality control staff. These changes are a relatively small percentage of all updates.

Generally, the reverse is not allowed (i.e., automatically updating from a known value to an unknown value for items such as a Social Security Number). When such a change is necessary, it **SHOULD** be reviewed manually by quality control staff.

5.10.1.7. Changes to Variables Used for Linkage

When a hospital submits changes involving fields that are used for linkage, such as Patient's Last Name, these have to be handled carefully. Making these corrections manually is the safest method, allowing for review by quality control staff. If the changes are done automatically, there is a risk of getting information on two tumors or two individuals confused, unless the changes are made in the correct order, depending on the timing of the relinking procedures of the computer system.

5.10.1.8. Significant Analysis Variables

For critical items such as Primary Site, Morphology, and Collaborative Stage, manual review of the changes by quality control staff is recommended. Either the proposed changes and the current values can be listed for manual review and correction, or the changes can be applied automatically and listed for subsequent review. Many central registries require that documentation be submitted with updates of this type to justify the proposed changes.

5.10.1.9. Treatment and Physician Updates

Treatment and physician updates are minimal but present the biggest problems in automatic updating. When changes and additions of treatment come from different hospitals, it often is difficult to determine if the treatment update represents the same treatment as that already stored in the database. For example, information on surgery may be submitted with a different day from that currently in the database, or the date of treatment may be partly unknown or may be an estimate. When treatment information comes in from different hospitals and the treatment submitted is the same type or code and performed in the same month (from all of the hospitals), one of the cancer registry software vendors considers this to be the same treatment, otherwise the vendor considers it to be a different treatment. A report of the treatments before and after updating can be reviewed for treatments that probably are the same but did not match because of date differences or code differences.

Some registries do not allow any automatic updating of treatment. All treatment updates received are reviewed and central registry staff determine manually if an update of the tumor record is needed. Updating the Follow-Up Physician item can present problems. A software vendor can allow the hospital that is designated as the follow-up hospital to update the follow-up physician. Another method is to allow only the central office to designate the follow-up physician.

5.10.1.10. Other Data Items

Other data items usually are updated and then reviewed. Some central registries do not allow automatic updates except for basic follow-up items; all other changes are determined manually.

5.11. LINKAGES WITH EXTERNAL FILES

Linkage of the central registry database with non-registry files serves several purposes for the registry. For example, there may be external files that can provide follow-up for the central registry's cases, or there may be special research studies requiring the linking of a cohort against the registry database.

5.11.1 Standards

The central registry **MUST** develop the technical, procedural, and administrative capacity to perform linkages with external files.

5.11.1.1. Linkage With Death (Mortality) Files

Linkage with death files is a particular case of the general linkage problem, one that **MUST** be routinized in the central registry's processes. This procedure usually is a batch process that compares the annual and monthly or quarterly death files from the vital statistics agency to the registry database. For positive matched records, the process becomes one of updating the registry files with the death information (see the discussion of updating in Section 5.10.) For possible matches, the system **MUST** generate reports for quality control staff to resolve manually. Non-matched deaths due to cancer require manual processing and a tracking system as described in Section 2.2.9. Because both the death file and the registry file are dynamic, timing of the linkages is important.

5.11.1.2. Other Files

The system **SHOULD** be capable of linking other files to the registry, for the purposes of obtaining patient follow-up and for special studies:

- *Follow-Up:* The potential sources of follow-up data against which the central registry may be linked are listed in Section 2.2.13. Linkage generally will be a batch process of comparing the files and for positive matches, adding follow-up data to the registry.

- *Special Studies:* Some research studies involve linking an external file to the central registry. Examples would include linking a cohort, such as a roster from a place of employment, against the central registry to determine occurrence of cancer among the cohort; or linking another disease registry, such as an AIDS registry, against the central registry to ascertain the occurrence of cancer among the people with AIDS. Confidentiality precautions **MUST** be followed stringently in all such investigations (see Section 4.1.1.).

5.12. DOCUMENTATION

Good documentation is an essential aspect of a well-designed system. It is necessary for system maintenance, training, quality control, and security; yet it often is incomplete and out of date. Documentation **SHOULD** be high among the registry's priorities.

5.12.1. Standards

Adequate central registry staff and time **MUST** be provided to prepare and maintain high quality, up-to-date system documentation.

The system documentation **SHOULD** include a management-level, functional description of the system, including a comprehensive narrative and flow diagrams. In addition, manuals or subsets of the documentation **SHOULD** be produced for the system, as follows:

- *User Manual:* The user manual **SHOULD** describe the user interface with the input, processing, and output of the system.

- *Technical Manual:* The technical manual **SHOULD** provide information to computer-trained personnel about the design and software of the system. It **SHOULD** contain system flowcharts defining major components of the system, definitions of individual programs, numerical analyses defining special calculations, definition of inputs and outputs, and definitions of reports.

- *Operator Manual:* The operator manual **SHOULD** describe the database and security and recovery procedures for the system. It **SHOULD** contain error codes/messages and handling procedures, computer run instructions, definitions of file retention and backup procedures, and definitions of data security.

Documentation **SHOULD** be available online as well as in hardcopy form.

Appendix A: NAACCR Membership Standards

- ***Full:*** Full member organizations are central registries that are, or have the potential to become, population-based registries.

- ***Individual:*** Individual members are those persons who are not currently working in a member organization who have demonstrated career and professional commitments and interests that are consistent with or complementary to those of NAACCR. Candidates for individual membership must be able to demonstrate involvement or activity in one or more of the following areas: cancer epidemiology, patient care, cancer control, cancer registration, professional education, research, and biostatistics. Each candidate must make a commitment to support NAACCR through active participation in the activities of the Association. Individual members shall be entitled to participate and vote as a member of committees, subcommittees, or work groups. Individual members may chair subcommittees or work groups. Individual members may not chair a committee, vote on matters brought before the Membership at the Annual Meeting, vote for or hold an elected position in the Association.

- ***Sponsoring:*** Sponsoring member organizations are national organizations primarily involved in cancer control prevention and research. Each sponsoring member organization shall be entitled to one vote on each matter submitted to membership vote. No action taken by the Association shall be construed as committing any sponsoring member organization to a prescribed course of action. Each sponsoring member organization may designate one or more representatives from their organization to participate in the Association's affairs on behalf of such organization. Representatives of sponsoring member organizations may be a member of and chair a committee. Only one representative of a sponsoring member organization shall be entitled to cast that organization's vote.

- ***Sustaining:*** Sustaining member organizations are organizations interested in promoting the purposes of the Association. No action taken by the Association shall be construed as committing any sustaining member organization to a prescribed course of action. Each sustaining member organization may designate one or more representatives from such organization to participate in the Association's affairs on behalf of such organization. Sustaining member organizations shall not be entitled to vote, and their representatives shall not be entitled to hold office or to chair a committee, but they shall be entitled to serve as members of committees.

Appendix B: Sample Reporting Legislation From Louisiana

Regular Session, 1995

HOUSE BILL NO. 1991

BY REPRESENTATIVE THOMAS AND SENATORS JOHNSON AND LAMBERT

ACT No. 1197

AN ACT

To amend and reenact R.S. 40:1299.80(1) and (4), 1299.81, 1299.82(1) through (3), 1299.84, 1299.85, 1299.87, and 1299.89(A) and to enact R.S. 40:1299.80(6) and (7), relative to the operation of a statewide cancer registry; to clarify the cancer reporting responsibilities of medical care professionals and institutions; to provide for intervention in cases of non-compliance; to provide for confidentiality requirements; to protect program participants from civil liability; to authorize the exchange of cancer data with other states; and to provide for related matters.

Be it enacted by the Legislature of Louisiana:

Section 1. R.S. 40:1299.80(1) and (4), 1299.81, 1299.82(1) through (3), 1299.84, 1299.85, 1299.87, and 1299.89(A) are hereby amended and reenacted and R.S. 40:1299.80(6) and (7) are hereby enacted to read as follows:

§ 1299.80. Definitions

As used in this Part:

a. "President" shall mean president of the Louisiana State University System, or his designee.
b. "Participating hospital" shall mean every hospital operating as such in the state of Louisiana.
c. "Pathology laboratory" shall mean every pathology laboratory located or doing business in the state of Louisiana.
d. "Office" shall mean the office of president.
e. "Board" shall mean the Louisiana Cancer and Lung Trust Fund Board.
f. "Health Care provider" shall mean every licensed health care facility and licensed health care provider, as defined in R. S. 40:1299.41(A)(1), in the state of Louisiana.
g. "Radiation center" shall mean every freestanding radiation diagnostic and treatment facility in the state of Louisiana.

§ 1299.81. Cancer registry program; data: statewide

The president of the Louisiana State University System shall establish in the office of the president a statewide registry program for reporting cancer cases for the purpose of gathering statistical data to aid in the assessment of cancer incidence, survival rates, possible causes of specific cancers, and other related aspects of cancer in Louisiana. The program shall collect and disseminate cancer incidence data on a statewide level in accordance with the provisions of this Part.

§ 1299.82. Powers; duties

The president shall:

(1) Collaborate with each participating health care provider and radiation center in the state of Louisiana to establish a uniform statewide registry system for collecting cancer incidence data and shall promulgate rules and regulations therefore in accordance with policies established by the board.

(2) Establish quality control programs and a training program for health care providers and the personnel of the participating radiation centers.

(3) Cooperate with the National Cancer Institute and the Centers for Disease Control in providing cancer incidence data.

(4) Comply with reporting procedures and requirements established by the board for tumor registry.

(5) Collaborate in studies with clinicians and epidemiologists and publish reports on the results of such studies, and

(6) Establish, in accordance with policies of the board, rules and regulations to provide for confidentiality of a patient's records.

(7) Establish and promulgate, in accordance with policies established by the board, the rules and regulations necessary to effectuate the purposes of this Part.

(8) Contract with private tumor registries for the collection and furnishing of data to the statewide registry and for the necessary planning and coordination incident thereto.

§ 1299.83. Authority

In addition to other authority, the president may:

(1) Accept on behalf of the state any federal funds to assist in meeting the cost of carrying out purposes of the Part.

(2) Accept on behalf of the state funds from any private agency, such as the American Cancer Society, to assist in the cost of the carrying out the purposes of this Part.

§ 1299.84. Participation in program

A. Any health care provider or radiation center diagnosing or providing treatment to cancer patients shall report each case of cancer to the president in a format prescribed by the president within six months of admission or diagnosis. If the facility fails to report in a format prescribed by the president, the president may enter the facility, obtain the information, and report it in the appropriate format. In these cases, the facility shall reimburse the president for the cost of obtaining and reporting the information.

B. Any health care provider or radiation center diagnosing or providing treatment to cancer patients shall report each cancer case. In addition, health care providers shall furnish follow-up data on each cancer patient hen requested.

C. Any health care provider or radiation center which provides diagnostic or treatment services to patients with cancer shall report any additional demographic, diagnostic, or treatment information requested by the president concerning any person presently or previously receiving services who has or had a malignant tumor. Additionally, the president shall have physical access to all records which would identify cases of cancer or would establish characteristics of the cancer, treatment of the cancer, or medical status of any identified cancer patient.

§ 11299.85. Reports; liability for

A. No action for damages arising from the disclosure of confidential or privileged information may be maintained against any person, or the employer or employee of any person, who participates in good faith in the reporting of cancer registry data or data for cancer morbidity or mortality studies in accordance with this Part.

B. No license of a health care provider may be denied, suspended, or revoked for good faith disclosure of confidential or privileged information or the reporting of cancer registry data or data for cancer morbidity studies in accordance with this part.

C. Nothing in this Part shall be construed to apply to the unauthorized disclosure of confidential or privileged information when such disclosure is due to gross negligence or willful misconduct.

D. All information reported pursuant to this Part shall be a confidential arid privileged. The secretary shall take strict measures to ensure that all identifying information is kept confidential.

E. All information regarding case specific data, as distinguished from group, tabular, or aggregate data concerning patients or health care providers contained in records of interviews, written reports, and statements procured by the secretary or by any other person, agency, or organization acting in connection with cancer morbidity and mortality.

Studies shall be confidential and privileged and shall be used solely for the purposes of the study. Nothing in this Section shall prevent the secretary from publishing compilations relating to morbidity and mortality studies which do not identify case specific data or sources of information.

* * *

§ 1299.86. Advisory functions

a. The tumor registry shall be operated under policies developed by the board and administered by the president.

b. The board shall establish policies for the development, accumulation, and distribution of data obtained under this Part.

c. The board shall exercise its powers, duties, functions, and responsibilities in the manner provided for agencies transferred in accordance with R. S. 36:802. The terms “secretary” and “undersecretary” as used in such Section and as applicable to the board shall mean the president or the president’s designee.

§ 1299.87. Disclosure of medical records to cancer registries

A. Notwithstanding any other provision of law to the contrary, all health care providers and radiation centers shall release an abstract of the patient's record reflecting the past or present physical condition of a patient upon request of the Louisiana cancer registry program established pursuant to the provisions of this Part. The cancer registry shall take strict measures to assure that all identifying information contained in patient record abstracts will be kept confidential.

B. The president may enter into agreements to exchange confidential information with other cancer registries in order to obtain complete reports of Louisiana residents diagnosed or treated in other states and to provide information to other states regarding their residents diagnosed or treated in Louisiana. However, before releasing confidential information the president shall obtain from such state registries, agencies, or researchers an agreement in writing to keep non-aggregate, case-specific information confidential and privileged. In no event shall either cancer registry bear liability for loss, expense, attorney fees, or claims for injury or damages arising out of acts or omissions in the performance of this agreement on the part of the other registry.

C. The office of the president shall promulgate rules and regulations in accordance with the Administrative Procedure Act to specify the extent to which confidential data may be disclosed to other local, state, or federal public health or environmental agencies, or to corroborating medical researchers, when the confidential information is necessary to carry out the duties of the agency or researchers in the investigation, control, or surveillance of disease, as determined by the office of the president. Before releasing confidential information to the researchers, the president shall obtain an agreement in writing from the researchers that neither the office of the president nor the other entity shall bear liability for loss, expense, attorney fees, or claims for injury or damages arising out of acts or omissions in the performance of this agreement on the part of the other.

D. Any disclosure authorized by this Part shall include only the information necessary for the stated purpose of the requested disclosure, and shall be made only upon written agreement that the information will be kept confidential and will not be further disclosed without written authorization of the office of the president.

E. The furnishing of confidential data in accordance with this Part shall not expose any person, agency, or entity furnishing data to liability and shall not be considered to be in violation of any privileged or confidential relationship, provided the participant has acted in good faith in the reporting as required in this Part.

F. No case specific data shall be available for subpoena nor shall it be disclosed, discoverable, or compelled to be produced in any civil, criminal, administrative, or other proceeding, nor shall such records be deemed admissible as evidence in any civil, criminal, administrative, or other tribunal or court for any reason. Nothing in this section shall supersede the provisions of R.S. 40:3.1 (A) through (H).

G. Nothing in this Part shall be construed to apply to the unauthorized disclosure of confidential or privileged information when such disclosure is due to gross negligence or willful misconduct.

* * *

§ 1299.89. Annual cancer report

A. The office of the president shall annually publish a comprehensive report based on available information on the incidence of cancer in Louisiana and the progress made in reducing or eliminating the high cancer rates in Louisiana.

B. The report shall be submitted by March 31 of each year to the governor, the speaker of the House of Representatives, the president of the Senate, and the House and Senate Committees on Health and Welfare.

C. The Joint Subcommittee on Health of the Joint Committee on Health and Welfare shall oversee the compilation of the report during the year.

* * *

Section 2. This Act shall become effective upon signature by the governor or, if not signed by the governor, upon expiration of the time for bills to become law without signature by the governor, as provided in Article III, Section 18 of the Constitution of Louisiana. If vetoed by the governor and subsequently approved by the legislature, this Act shall become effective on the day following such approval.

Appendix C: Cancer Registries Amendment Act

The Cancer Registries Amendment Act, Public Law 102-515, is reproduced beginning on the next page.

106 STAT. 3372 PUBLIC LAW 102–515—OCT. 24, 1992

Public Law 102-515
102d Congress

An Act

Oct. 24, 1992
[S. 3312]

Entitled the "Cancer Registries Amendment Act."

Cancer Registries Amendment Act.
Diseases.
Health and health care.
42 USC 201 note.
42 USC 280e note.
42 USC 280e note.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "Cancer Registries Amendment Act."

SEC. 2. FINDINGS AND PURPOSE.

(a) FINDINGS.—Congress finds that—

(1) cancer control efforts, including prevention and early detection, are best addressed locally by State health departments that can identify unique needs;

(2) cancer control programs and existing statewide population-based cancer registries have identified cancer incidence and cancer mortality rates that indicate the burden of cancer for Americans is substantial and varies widely by geographic location and by ethnicity;

(3) statewide cancer incidence and cancer mortality data can be used to identify cancer trends, patterns, and variation for directing cancer control intervention;

(4) the American Association of Central Cancer Registries (AACCR) cites that of the 50 States, approximately 38 have established cancer registries, many are not statewide and 10 have no cancer registry; and

(5) AACCR also cites that of the 50 States, 39 collect data on less than 100 percent of their population, and less than half have adequate resources for insuring minimum standards for quality and for completeness of case information.

(b) PURPOSE.—It is the purpose of this Act to establish a national program of cancer registries.

SEC. 3. NATIONAL PROGRAM OF CANCER REGISTRIES.

Title III of the Public Health Service Act (42 U.S.C. 241 et seq.) is amended by adding at the end the following new part:

"PART M—NATIONAL PROGRAM OF CANCER REGISTRIES

42 USC 280 e.

"SEC. 399H. NATIONAL PROGRAM OF CANCER REGISTRIES.

"(a) IN GENERAL.—The Secretary, acting through the Director of the Centers for Disease Control, may make grants to States, or may make grants or enter into contracts with academic or nonprofit organizations designated by the State to operate the State's cancer registry in lieu of making a grant directly to the State, to support the operation of population-based, statewide cancer registries in order to collect, for each form of in-situ and invasive

cancer (with the exception of basal cell and squamous cell carcinoma of the skin), data concerning—

"(1) demographic information about each case of cancer;

"(2) information on the industrial or occupational history of the individuals with the cancers, to the extent such information is available from the same record;

"(3) administrative information, including date of diagnosis and source of information;

"(4) pathological data characterizing the cancer, including the cancer site, stage of disease (pursuant to Staging Guide), incidence, and type of treatment; and

"(5) other elements determined appropriate by the Secretary.

"(b) MATCHING FUNDS.—

"(1) IN GENERAL.—The Secretary may make a grant under subsection (a) only if the State, or the academic or nonprofit private organization designated by the State to operate the cancer registry of the State, involved agrees, with respect to the costs of the program, to make available (directly or through donations from public or private entities) non-Federal contributions toward such costs in an amount that is not less than 25 percent of such costs or $1 for every $3 of Federal funds provided in the grant.

"(2) DETERMINATION OF AMOUNT OF NON-FEDERAL CONTRIBUTION; MAINTENANCE OF EFFORT.—

"(A) Non-Federal contributions required in paragraph (1) may be in cash or in kind, fairly evaluated, including plant, equipment, or services. Amounts provided by the Federal Government, or services assisted or subsidized to any significant extent by the Federal Government, may not be included in determining the amount of such non-Federal contributions.

"(B) With respect to a State in which the purpose described in subsection (a) is to be carried out, the Secretary, in making a determination of the amount of non-Federal contributions provided under paragraph (1), may include only such contributions as are in excess of the amount of such contributions made by the State toward the collection of data on cancer for the fiscal year preceding the first year for which a grant under subsection (a) is made with respect to the State. The Secretary may decrease the amount of non-Federal contributions that otherwise would have been required by this subsection in those cases in which the State can demonstrate that decreasing such amount is appropriate because of financial hardship.

"(c) ELIGIBILITY FOR GRANTS.—

"(1) IN GENERAL.—No grant shall be made by the Secretary under subsection (a) unless an application has been submitted to, and approved by, the Secretary. Such application shall be in such form, submitted in such a manner, and be accompanied by such information, as the Secretary may specify. No such application may be approved unless it contains assurances that the applicant will use the funds provided only for the purposes specified in the approved application and in accordance with the

requirements of this section, that the application will establish such fiscal control and fund accounting procedures as may be necessary to assure proper disbursement and accounting of Federal funds paid to the applicant under subsection (a) of this section, and that the applicant will comply with the peer review requirements under Sections 491 and 492.

"(2) ASSURANCES.—Each applicant, prior to receiving Federal funds under subsection (a), shall provide assurances satisfactory to the Secretary that the applicant will—

"(A) provide for the establishment of a registry in accordance with subsection (a);

"(B) comply with appropriate standards of completeness, timeliness, and quality of population-based cancer registry data;

"(C) provide for the annual publication of reports of cancer data under subsection (a); and

"(D) provide for the authorization under State law of the statewide cancer registry, including promulgation of regulations providing—

"(i) a means to assure complete reporting of cancer cases (as described in subsection (a)) to the statewide cancer registry by hospitals or other facilities providing screening, diagnostic or therapeutic services to patients with respect to cancer;

"(ii) a means to assure the complete reporting of cancer cases (as defined in subsection (a)) to the statewide cancer registry by physicians, surgeons, and all other

health care practitioners diagnosing or providing treatment for cancer patients, except for cases directly referred to or previously admitted to a hospital or other facility providing screening, diagnostic or therapeutic services to patients in that State and reported by those facilities;

"(iii) a means for the statewide cancer registry to access all records of physicians and surgeons, hospitals, outpatient clinics, nursing homes, and all other facilities, individuals, or agencies providing such services to patients which would identify cases of cancer or would establish characteristics of the cancer, treatment of the cancer, or medical status of any identified patient;

"(iv) for the reporting of cancer case data to the statewide cancer registry in such a format, with such data elements, and in accordance with such standards of quality timeliness and completeness, as may be established by the Secretary;

"(v) for the protection of the confidentiality of all cancer case data reported to the statewide cancer registry, including a prohibition on disclosure to any person of information reported to the statewide cancer registry that identifies, or could lead to the identification of, an individual cancer patient, except for disclosure to other State cancer registries and local and State health officers;

"(vi) for a means by which confidential case data may in accordance with State law be disclosed to cancer researchers for the purposes of cancer prevention, control and research;

"(vii) for the authorization or the conduct, by the statewide cancer registry or other persons and organizations, of studies utilizing statewide cancer registry data, including studies of the sources and causes of cancer, evaluations of the cost, quality, efficacy, and appropriateness of diagnostic, therapeutic, rehabilitative, and preventative services and programs relating to cancer, and any other clinical, epidemiological, or other cancer research; and

"(viii) for protection for individuals complying with the law, including provisions specifying that no person shall be held liable in any civil action with respect to a cancer case report provided to the statewide cancer registry, or with respect to access to cancer case information provided to the statewide cancer registry.

"(d) RELATIONSHIP TO CERTAIN PROGRAMS.—

"(1) IN GENERAL.—This section may not be construed to act as a replacement for or diminishment of the program carried out by the Director of the National Cancer Institute and designated by such Director as the Surveillance, Epidemiology, and End Results Program (SEER).

"(2) SUPPLANTING OF ACTIVITIES.—In areas where both such programs exist, the Secretary shall ensure that SEER support is not supplanted and that any additional activities are consistent with the guidelines provided for in subsection (c)(2) (C) and (D) and are appropriately coordinated with the existing SEER program.

"(3) TRANSFER OF RESPONSIBILITY.—The Secretary may not transfer administration responsibility for such SEER program from such Director.

"(4) COORDINATION.—To encourage the greatest possible efficiency and effectiveness of Federally supported efforts with respect to the activities described in this subsection, the Secretary shall take steps to assure the appropriate coordination of programs supported under this part with existing Federally supported cancer registry programs.

"(e) REQUIREMENT REGARDING CERTAIN STUDY ON BREAST CANCER.—In the case of a grant under subsection (a) to any State specified in section 399K(b), the

Secretary may establish such conditions regarding the receipt of the grant as the Secretary determines are necessary to facilitate the collection of data for the study carried out under section 399C.

42 USC 280 e-1.

"SEC. 399I. PLANNING GRANTS REGARDING REGISTRIES.

"(a) IN GENERAL.—

"(1) STATES.—The Secretary, acting through the Director of the Centers for Disease Control, may make grants to States for the purpose of developing plans that meet the assurances required by the Secretary under section 399B(c)(2).

"(2) OTHER ENTITIES.—For the purpose described in paragraph (1), the Secretary may make grants to public entities other than States and to nonprofit private entities. Such a grant may be made to an entity only if

the State in which the purpose is to be carried out has certified that the State approves the entity as qualified to carry out the purpose.

"(b) APPLICATION.—The Secretary may make a grant under subsection (a) only if an application for the grant is submitted to the Secretary, the application contains the certification required in subsection (a)(2) (if the application is for a grant under such subsection), and the application is in such form, is made in such manner, and contains such agreements, assurances, and information as the Secretary determines to be necessary to carry out this section.

42 USC 280 e-2. "SEC. 399J. TECHNICAL ASSISTANCE IN OPERATIONS OF STATEWIDE CANCER REGISTRIES.

"The Secretary, acting through the Director of the Centers for Disease Control, may, directly or through grants and contracts, or both, provide technical assistance to the States in the establishment and operation of statewide registries, including assistance in the development of model legislation for statewide cancer registries and assistance in establishing a computerized reporting and data processing system.

42 USC 280 e-3. **"SEC. 399K. STUDY IN CERTAIN STATES TO DETERMINE THE FACTORS CONTRIBUTING TO THE ELEVATED BREAST CANCER MORTALITY RATES.**

"(a) IN GENERAL.—Subject to subsections (c) and (d), the Secretary, acting through the Director of the National Cancer Institute, shall conduct a study for the purpose of determining the factors contributing to the fact that breast cancer mortality rates in the States specified in subsection (b) are elevated compared to rates in other States.

"(b) RELEVANT STATES.—The States referred to in subsection (a) are Connecticut, Delaware, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, Vermont, and the District of Columbia.

"(c) COOPERATION OF STATE.—The Secretary may conduct the study required in subsection (a) in a State only if the State agrees to cooperate with the Secretary in the conduct of the study, including providing information from any registry operated by the State pursuant to section 399H(a).

"(d) PLANNING, COMMENCEMENT, AND DURATION.—The Secretary shall, during each of the fiscal years 1993 and 1994, develop a plan for conducting the study required in subsection (a). The study shall be initiated by the Secretary not later than fiscal year 1994, and

the collection of data under the study may continue through fiscal year 1998.

"(e) REPORT.—Not later than September 30, 1999, the Secretary shall complete the study required in subsection (a) and submit to the Committee on Energy and Commerce of the House of Representatives, and to the Committee on Labor and Human Resources of the Senate, a report describing the findings and recommendations made as a result of the study.

42 USC 280 e-4. "SEC. 399L. AUTHORIZATION OF APPROPRIATIONS.

"(a) REGISTRIES.—For the purpose of carrying out this part, the Secretary may use $30,000,000 for each of the fiscal years 1993 through 1997.

Out of any amounts used for any such fiscal year, the Secretary may obligate not more than 25 percent for carrying out section 399I, and not more than 10 percent may be expended for assessing the accuracy, completeness and quality of data collected, and not more than 10 percent of which is to be expended under subsection 399J.

"(b) BREAST CANCER STUDY.—Of the amounts appropriated for the National Cancer Institute under subpart 1 of part C of title IV for any fiscal year in which the study required in section 399K is being carried out, the Secretary shall expend not less than $1,000,000 for the study.".

Approved October 24, 1992.
Authorization extended through 1998.

LEGISLATIVE HISTORY–S. 3312:
CONGRESSIONAL RECORD, Vol. 138 (1992):
Oct. 2, considered and passed Senate.
Oct. 5, considered and passed House, amended.
Oct. 7, Senate concurred in House amendment.

Appendix D: Benign Brain Tumor Cancer Registries Amendment Act

The Benign Brain Tumor Cancer Registries Amendment Act, Public Law 107-260, is reproduced beginning on the next page.

PUBLIC LAW 107-260—OCT. 29, 2002 116 STAT. 1743

Public Law 107-260
107th Congress

An Act

To amend the Public Health Service Act to provide for the collection of data on benign brain-related tumor through the national program of cancer registries.

Oct. 29, 2002
[S. 2558]

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

Benign Brain Tumor Cancer Registries Amendment Act.
42 USC 201 note.

SECTION 1. SHORT TITLE.

This Act may be cited as the "Benign Brain Tumor Cancer Registries Amendment Act".

SEC. 2. NATIONAL PROGRAM OF CANCER REGISTRIES; BENIGN BRAIN-RELATED TUMORS AS ADDITIONAL CATEGORY OF DATA COLLECTED.

(a) In GENERAL—Section 399B of the Public Health Service Act (42 U.S.C. 280e), as redesignated by section 502 (2)(A) of Public Law 106-310 (114 Stat. 1115), is amended in subsection (a)—

(1) by redesignating paragraphs (1) through (5) as subparagraphs (A) through (3), respectively, and indenting appropriately;

(2) by striking "(a) IN GENERAL—The Secretary" and inserting the following:

(a) IN GENERAL—

"(1) STATEWIDE CANCER REGISTRIES—The Secretary";

(3) in the matter preceding subparagraph (A) (as so redesignated). By striking "population-based" and all that follows through "data" and inserting the following: "population-based, statewide registries to collect, for each condition specified in paragraph (2)(A), data"; and

(4) by adding at the end the following:

"(2) CANCER; BENIGN BRAIN-RELATED TUMORS—

"(A) IN GENERAL—For purposes of paragraph (1), the conditions referred to in this paragraph are the following:

"(i) Each form of in-situ and invasive cancer with the exception of basal cell and squamous cell carcinoma of the skin), including malignant brain-related tumors.

"(ii) Benign brain-related tumors

"(B) BRAIN-RELATED TUMOR—For purposes of subparagraph (A):

"(i) The term 'brain-related tumor' means a listed primary tumor (whether malignant or benign) occurring in any of the following sites:'

"(I) The brain, meninges, spinal cord, cauda equina, a cranial nerve or nerves, or any other part of the central nervous system.
"(II) The pituitary gland, pineal gland, or craniopharyngeal duct.

"(ii) The term 'listed', with respect to a primary tumor, means a primary tumor that is listed in the International Classification of Diseases for Oncology (commonly referred to as the ICD-O).

"(iii) The term 'International Classification of Diseases for Oncology' means a classification system that includes topography (site) information and histology (cell type information) developed by the World Health Organization, in collaboration with international centers, to promote international comparability in the collection, classification, processing, and presentation of cancer statistics. The ICD-O system is a supplement to the International Statistical Classification of Diseases and Related Health Problems (commonly known as the ICD) and is the standard coding system used by cancer registries worldwide. Such term includes any modification made to such system for purposes of the United States. Such term further includes any published classification system that is internationally recognized as a successor to the classification system referred to in the first sentence of this clause.

"(C) STATEWIDE CANCER REGISTRY—References in this section not cancer registries shall be considered to be references to registries described in this subsection.".

Grants.
42 USC 280e note.

(b) APPLICABILITY—The amendments made by subsection (a) apply to grants under section 399B of the Public Health Service Act for fiscal year 2002 and subsequent fiscal years, except that, in the case of a State that received such a grant for fiscal year 2000, the Secretary of Health and Human Services may delay the applicability of such amendments to the State for not more than 12 months if the Secretary determines that compliance with such amendments requires the enactment of a statute by the State or the issuance of State regulations.

Approved October 29, 2002.

LEGISLATIVE HISTORY—s. 2558:
Congressional record, Vol. 148 (2002):
Aug. 1. considered and passed Senate.
Oct 10. considered and passed House.

Appendix E: Resources for Education and Training for Providers and Users of Cancer Registry Data

The NAACCR Education and Training Committee maintains a resource list located on the NAACCR website (www.naaccr.org). Listed below is the contact information for standard-setting organizations.

American College of Surgeons (ACoS)
633 N. Saint Clair Street
Chicago, IL 60611-3211
Telephone: (312) 202-5000
Fax: (312) 202-5001
E-mail: postmaster@facs.org
Website: www.facs.org

American Joint Committee on Cancer (AJCC)
633 N. Saint Clair Street
Chicago, IL 60611-3211
Telephone: (312) 202-5290
E-mail: sburkhardt@facs.org
Website: cancerstaging.org

Centers for Disease Control and Prevention (CDC)
National Program of Cancer Registries (NPCR)
Division of Cancer Prevention and Control
National Center for Chronic Disease
Prevention and Health Promotion
4770 Buford Highway, NE
MS K53
Atlanta, GA 30341-3717
Telephone: (770) 488-4783
Fax: (770) 488-4759
Website: www.cdc.gov/cancer/npcr

Canadian Council of Cancer Registries
c/o Statistics Canada
Canadian Cancer Registry
Health Statistics Section
Health Statistics Division
Main Building, Room 22000, Section F
120 Parkdale Avenue
Ottawa, ON K1A OT6
Telephone: (613) 951-1630
Fax: (613) 951-0792
Website: www.statcan.ca

Commission on Cancer (CoC)
633 N. Saint Clair Street
Chicago, IL 60611-3211
Telephone: (312) 202-5085
E-mail: coc@facs.org
Website: www.facs.org

National Cancer Institute SEER Program
Cancer Surveillance Research Program
Division of Cancer Control and Population Sciences
6116 Executive Boulevard, MSC 8316
Suite 504
Bethesda, MD 20892-8316
Telephone: (301) 496-8510
Fax: (301) 496-9949
E-mail: cancer.gov_staff@mail.nih.gov
Website: www.seer.cancer.gov

National Cancer Registrars Association (NCRA)
1340 Braddock Place #203
Alexandria, VA 22314
Telephone: (703) 299-6640
Fax: (703) 299-6620
E-mail: info@ncra-usa.org
Website: www.ncra-usa.org

North American Association of Central Cancer Registries, Inc. (NAACCR)
2121 West White Oaks Drive
Springfield, IL 62704-6495
Telephone: (217) 698-0800
Fax: (217) 698-0188
E-mail: info@naaccr.org
Website: www.naaccr.org

Appendix F: Sample Case Sharing Agreement

Agreement for Exchange of Cancer Data
Between the

(name of submitting registry)

and

(name of receiving registry)

(1) Services:

By signing this agreement, the parties state their intention to exchange information concerning cancer patients who are residents of the other's state, province, or county. This exchange is based on the mutual assurance that the identifying information on the patient(s) exchanged are protected and shall be kept strictly confidential. This exchange does not pertain to any data collected as part of special morbidity or mortality studies or other research projects.

In addition, the parties agree to:

a) Provide the information electronically in the most recent NAACCR record layout.

b) Provide the full exchange record.

c) Provide the information within 20 months of the close of the diagnosis.

d) Carefully restrict use of the information. The information is intended to be used for registry administration and for aggregated statistical tabulations and analyses.

e) Restrict access to cancer incidence data or identifiable information on a cancer patient or health care provider that was supplied under the terms of the agreement from being released to anyone not employed in the direct operation of the recipient registry. Employees may include those involved in the processing, administration, quality control review and the statistical surveillance of cancer incidence data.

f) Notify the exchange registry if, in the conduct of approved research or other activities, there is release of a cancer patient's identifying information. Should such a release take place, the receiving registry will be notified in writing within 48 hours of the release of the data.

g) Terminate this agreement immediately upon the written notification of either party to terminate the agreement.

(2) Confidentiality:

a) The parties understand and agree that any and all data which may lead to the identification of any patient, research subject, physician, other person, or reporting facility is strictly privileged and confidential, and agree to keep all such data strictly confidential.

b) The parties further agree to require all officers, agents, and employees to keep all such data strictly confidential; to communicate the requirements of this section to all officers, agents, and employees; to discipline all persons who may violate the requirements of this section; and to notify the originating party in writing within 2 working days (48 hours) of any violation of this section, including full details of the violation and corrective actions to be taken.

c) The parties further agree that all data provided under the provisions of this agreement may only be used for the purposes named in this agreement.

d) In the event that either party receives a subpoena or other court order compelling disclosure of confidential data, the parties agree to notify the registry that initially provided the data within 2 working days (48 hours) of receipt of the subpoena or court order. Additionally, the parties agree that, should they receive such a subpoena, they shall take all legal steps reasonably necessary to oppose the subpoena.

(3) Amendments:

This agreement may not be amended without prior written approval of both parties to the agreement.

(4) Assignment:

The parties understand and agree that this agreement may not be sold, assigned, or transferred in any manner and that any actual or attempted sale, assignment, or transfer shall render this agreement null, void, and of no further effect.

(5) Term:

This agreement shall be in effect from the date of execution until terminated by either of the parties. Termination shall be in writing sent pursuant to Section (6).

(6) Notices:

All notices required or desired to be made by either party to this agreement shall be sent by certified mail to the following respective addresses:

(Provide address and contact for each party to this agreement.)

(7) Signatures:

(Provide name, title, agency, date, and appropriate signatures for each registry.)

Appendix G: Method To Measure Completeness

NAACCR uses the incidence-to-mortality rate ratio method to measure completeness of case ascertainment. The method assumes that cancer death data are complete, and that the ratio of age-adjusted cancer incidence rates to age-adjusted cancer death rates by sex, race, and site vary little by geographical area in the United States and Canada. Over time, the interpretation of the incidence-to-mortality rate ratio has become more refined. The following adjustments were made, either to the method itself or to the interpretation of the rate-ratios:

- It was assumed that 20 percent of any difference observed between analogous race-sex-site-specific, age-adjusted incidence-to-mortality rate ratios from two geographic areas could be attributed to differential case fatality, while 80% of the difference could be attributed to under-ascertainment of cases in one of the jurisdictions. Previously, it was assumed that 100% of the difference could be attributed to under-ascertainment.

- Breast cancer cases were included in the model. Previously, breast cancer cases were excluded from the calculations because geographically diverse increases in mammography utilization had destabilized breast cancer incidence-to-mortality rate ratios. Recent data suggest that mammography use, breast cancer incidence, and breast cancer incidence-to-mortality rate ratios have become more uniform in the United States.

- All 11 SEER (14% of the U.S. population) areas have been used to construct SEER-incidence-to-U.S. mortality rate ratios. SEER has added areas to its geographic base over the years to increase its representativeness of the United States population. Previously, NAACCR had used data from the nine "original" SEER areas (10% of the U.S. population), because much was known about the nature of these data, their stability, and their relation to NAACCR data. As more became known about data from the additional two SEER areas, it became desirable to use data from all 11 areas in the construction of SEER-incidence-to-U.S.-mortality rate ratios, to enhance the representativeness of the ratios for the United States population as a whole.

- For similar reasons, data for both whites and blacks (weighted in proportion to their share of the population) were used to construct incidence-to-mortality rate ratios. Previously, data for whites were used exclusively for this purpose. Whites-only ratios were used with 1996-2000 data from Canada and Hawaii, as race is not used to differentiate population groups in either of these jurisdictions.

Race-specific completeness of case ascertainment in jurisdiction *s* (C_{sk}) was computed by dividing the *observed* race-specific (white; black) age-adjusted (2000 U.S.) incidence rate for both sexes and all cancer sites combined ("Observed T") by the *expected* race-specific (white; black) age-adjusted (2000 U.S.) incidence rate for both sexes and all cancer sites combined ("Expected T"):

$$C_{sk} = \frac{ObservedT_{sk}}{ExpectedT_{sk}}$$

The *expected* incidence rate for jurisdiction *s* was computed from jurisdiction race-sex-site-specific age-adjusted (2000 U.S.) death rates and incidence-to-mortality rate ratios computed from SEER race-sex-site-specific age-adjusted (2000 U.S.) incidence rates and U.S. race-sex-site-specific age-adjusted (2000 U.S.) death rates, thus:

$$ExpectedI_{skij} = \left(M_{skij} \right) \left(\frac{I_{SEERkij}}{M_{U.S.kij}} \right)$$

$$ExpectedT_{sk} = \sum_{i=1}^{2} \sum_{j=1}^{N} ExpectedI_{skij}$$

where:

I =	Age-adjusted (2000 U.S.) incidence rate for race *k,* sex *i,* site *j,* 1996 to 2000
M =	Age-adjusted (2000 U.S.) mortality rate for race *k,* sex *i,* site *j,* 1996 to 2000
s =	State, SEER area, province, or territory
SEER =	Combined 11 SEER areas [1]
U.S. =	United States
T =	Age-adjusted (2000 U.S.) incidence rate for total sites [2]

Overall completeness of case ascertainment in jurisdiction *s* (C_s) was calculated by adding weighted estimates of race-specific completeness of case ascertainment in jurisdiction *s* (C_{sk}), using the proportion of the population in each of the race groups (P_{sk}) as weights:

$$C_s = \sum_{k=1}^{2} C_{sk} \times P_{sk}$$

This method of estimating completeness assumes that race-sex-site-specific incidence-to-mortality rate ratios are relatively stable (within 20% limits). The incidence-to-mortality rate ratio standard to which all registries were adjusted, using SEER incidence rates and U.S. death rates, is the current NAACCR standard for this purpose.

The same methods were applied to Hawaii and all Canadian registries, except that jurisdiction-specific data were not race specific, and SEER-incidence-to-U.S.-mortality rate ratios were computed for whites only.

[1] Includes Atlanta, Connecticut, Detroit, Greater Bay Area (San Francisco/Oakland and San Jose/Monterey), Hawaii, Iowa, Los Angeles, New Mexico, Seattle/Puget Sound, and Utah.

[2] The cancer sites included in this calculation were buccal cavity and pharynx, esophagus, stomach, colorectum, liver, pancreas, lung and bronchus, melanoma of the skin (white only), female breast (excl. *in situ*), cervix uteri, corpus uteri and uterus, NOS, ovary, urinary bladder (incl *in situ*), kidney and renal pelvis, brain and other nervous system, Hodgkin's disease, non-Hodgkin's lymphoma, multiple myeloma, and leukemia. Cancer of the prostate was not included because differential screening across regions has caused instability in prostate cancer incidence-to-mortality rate ratios.

C_s was adjusted for the presence of duplicate records in the data of jurisdiction *s* (CA_s) thus:

$$CA_s = C_s \times U_s$$

where:

CA =	Adjusted overall completeness of ascertainment
C =	Unadjusted overall completeness of ascertainment
s =	State, SEER area, province, or territory
U =	Proportion of unduplicated records, based on NAACCR's *Protocol for Assessing Duplicate Cases.*

Impact of the Modified Population Estimates on the NAACCR Completeness Estimates. Recently, the United States Bureau of the Census revised the U.S. population estimates for the 1990s by using 2000 decennial census data to adjust the original post-1990 census population projections. The revised population estimates have an effect on both the incidence and death rates differentially across cancer sites and regions. The completeness estimates for all cancer registries have also been affected. Despite this revision, the number of registries meeting the NAACCR combined inclusion criteria has increased compared to last year's monograph. The population represented by these registries has also increased this year from 55 percent to 68 percent of the United States population.

For more information on the completeness estimate method, consult the following reference: Holly L. Howe. Conclusions of the Work Group for High Quality Criteria for Data Use. NAACCR Narrative [serial online] 2001; Winter: 8 (On-line) Available: http://www.naaccr.org/News/index.html.

Appendix H: Major-Minor Discrepancy Definitions

ALL SITES					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Date of Diagnosis	Discrepancy last four digits	100	Discrepancy in first two digits	First two digits from 99 to 01-12 vice versa	90
Primary Site	Discrepancy in 2nd and/or 3rd digit	100	Discrepancy in fourth digit	Change from any digits to C809 vice versa	90
Histology	Discrepancy in first three digits	95	Discrepancy in fourth digit		
Behavior	Any	100			
Grade			Any	Change from 2 or 3 to 9 vice versa	
Date Initial RX	Discrepancy last four digits	97	Discrepancy in first two digits	First two digits from 99 to 01-12 vice versa	90
Reason No CA Dire Surg	0 to any in range of 1-8 vice versa 8 to any in range of 1-7 vice versa	90	Difference within range of 1-7	9 to any in range of 0-8 vice versa	80
RX Summ--Radiation	Changes from 0 or 8 to any in range of 1-7 vice versa	95	0 to 8 vice versa Difference between 1-7	9 to any in range of 0-7 vice versa	90
RX Summ--Chemo	Changes from 0 or 8 to any in range of 1-7 vice versa	95	0 to 8 vice versa	9 to any in range of 0-7 vice versa	90
RX Summ--Hormone	Changes from 0 or 8 to any in range of 1-7 vice versa	90	0 to 8 vice versa	9 to any in range of 0-7 vice versa	85
RX Summ--BRM	Changes from 0 or 8 to any in range of 1-7 vice versa	90	0 to 8 vice versa	9 to any in range of 0-7	85
RX Summ--Other	0 or 8 to any in range of 1-7 vice versa	80	0 to 8 vice versa	9 to any in range of 0-7 vice versa	75

BREAST ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	Code 0 to any in the range of 1-4 vice versa	90	1 to 2 vice versa 3 to 4 vice versa Change from 3 or 4 to 1 or 2 vice versa	9 to 1-4 vice versa	85
Reg Nodes Pos	00 to 01-96 vice versa 97 to 98 vice versa	95	Within range of 01-90 97 or 98 to 0-96 vice versa	99 to 01-98 vice versa	90
Reg Nodes Examined	00 to 01-98 95-98 to 01-90 vice versa	95	Within range of 95-98	99 to 01-98 vice versa	90
Tumor Marker 1 ERA	Codes 4, 5, 6 (invalid codes) 0 to 1-3 or 8 vice versa 1 to 2 vice versa	90	1 to 3 vice versa 2 to 3 vice versa	9 to 0-3 vice versa	85
Tumor Marker 2 PRA	Codes 4, 5, 6 (invalid codes) 0 to 1, 2, 3, 8 vice versa 1 to 2 vice versa	90	1 to 3 vice versa 2 to 3 vice versa	9 to 0-8 vice versa	85
EOD--Tumor Size	Change from 000 to 001-990 vice versa Changes from 001-020 to 021-050 or 051-990 vice versa Changes from 021-050 or 051-990 to 001-020 vice versa Changes from the range of 021-050 to the range of 051-990 Changes from the range of 051-990 to the range of 021-050 997 to any in range of 1-996 or 998 vice versa 998 to any in range of 1-997 vice versa 002 to 001 or any in range of 003-998 vice versa	95	Any changes in the range of 001 to 020 Changes in the range of 021 to 050 Changes in range of 051-996	999 to 000-998 vice versa	90

BREAST ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
EOD--Extension		95	Within range of 10-30 Any changes in range of 40-70 80 to 85 vice versa	99 to 00-85 vice versa	90
EOD--Lymph Node Involv	0 to 1-8 vice versa Any in range of 1-4 to 5-8 vice versa 5 to 7 or 8 vice versa 7 to any in range of 1-6 or 8-9 vice versa 8 to any 6, 7, or 9 vice versa	95	Any changes within range of 1-4 6 to1-4 vice versa	9 to 1-8 vice versa	90
RX Summ--Surg Prim Site	00 to range of 10-90 vice versa Range of 10-17 to 30-90 vice versa 30 to 10-17, 40-90 vice versa 40-42 to 10, 30, or 50-90 vice versa 50-52 to 10-17, 30, 40-42, or 60-90 vice versa 60-62 to 10-17, 30, 40-42, 50-52, or 70-90 vice versa 70-72 to 10-17, 30, 40-62, 80, or 90 vice versa 80 or 90 to 10-72 vice versa	95	Within range of 10-17 Within range of 40-42 Within range of 50-52 Within range of 60-62 Within range of 70-72 Within range of 80-90	99 to 00-90 vice versa	90
RX Summ--Scope Reg LN Sur	0 to 1-5 vice versa	90	Within range of 1-5	9 to 0-5 vice versa	85
RX Summ Reg LN Examined	00 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98	99, 95, 97, 98 to any 01-96 vice versa	90
RX Summ--Surg Oth Reg/Dist	0 to 1-6 vice versa	90	Changes within range of 1-6	9 to range of 0-6 vice versa	85
Reconstruction--First Course			2 to any in range of 3-8 vice versa 1 to any in range of 2-8 vice versa Within range 3-8	9 to any in range of 0-8 vice versa	80

COLON ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	Any code other than 0	100			
Reg Nodes Pos	00 to any in the range of 01-96 vice versa 01-03 to 04-90 vice versa 97 to 98 vice versa	95	Within range of 01-03 Within range of 04-90	99 to 01-98 vice versa	90
Reg Nodes Examined	00 to any in range of 01-90 vice versa 95-98 to 1-95 vice versa.	95	Within the range of 95-98	99 to any other #	90
Tumor Marker 1	NO ERROR	95			
Tumor Marker 2	NO ERROR	95			
EOD--Tumor Size	000 to any in range of 001 to 990 Range of 001-990 to 000	95	Changes within range of 001-990	999 to 000-990	90
EOD--Extension	00 to 10-85 vice versa 10-16 to 40-85 vice versa 40-45 to 10-16 or 40-80 vice versa 10-45 to 50-85 vice versa 85 to 10-80 vice versa	95	Within range of 10-16 40 to 45 vice versa 50-80 vice versa	99 to 00-85	90
EOD--Lymph Node Involv	0 to 1-8 vice versa 1-3 to 7-8 vice versa	95	1 to 2 vice versa	9 to 0-8 vice versa	90
RX Summ--Surg Prim Site	00 to 10-90 Range of 10-90 to 00 Range of 10-14 to any in range of 20-70 vice versa 20 to any in range of 10-14 or range of 30-70 vice versa 30 or 31 to any in range of 10-14 or range of 40-70 vice versa 40 to any in range of 10-31 or 50-70 vice versa 50 to any in range of 10-40 or 60-70 vice versa 60 to any in range of 10-50 or 70 vice versa 70 to any in range of 10-60 vice versa	95	Within the range of 10-14 Within range of 20-27 Within range of 30-31 80 to 90 90 to 80	99 to 00-90	90
RX Summ--Scope Reg LN Sur	0 to 1 1 to 0	95		9 to 0 or 1	90
RX Summ--Reg LN Examined	00 to 01-98 vice versa	95	Any changes within the range of 95-98	99 to 00-98 vice versa90	

COLON ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
RX Summ--Surg Oth Reg/Dist	0 to any in the range of 1-8 vice versa	90	3 to 4 vice versa 1 to any in range of 2-8 vice versa	9 to 0-8 vice versa	85
Reconstruction --First Cour	NO ERRORS – simply count # of blanks (correct) and # of other entries				

LUNG ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	Code 0	90	Change from 1 to 2 Change from 3 or 4 to 1 or 2	9 to 1 or 2	85
Reg Nodes Pos	00 to 01-96 vice versa Range of 01-03 to any in range of 04-90 vice versa 97 to 98 vice versa	95	Change numbers within the range of 01-03 Changes within the range of 04-90 97 or 98 to any in range of 1-96 vice versa		90
Reg Nodes Examined	00 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98 96-98 to any in range of 01-90 vice versa	99 to 00-98 vice versa	90
Tumor Marker 1	NO ERROR	95			
Tumor Marker 2	NO ERROR	95			
EOD--Tumor Size	000 to any in range of 001 to 990 001 to 030 to 031-990 vice versa Range of 001-990 to 000	95	Changes within range of 001-990	999 to 000-990	90
EOD--Extension	00 to any in range of 10-85 vice versa Range of 10-20 to any in range of 40-85 vice versa 25 or 30 to any in range of 10-20 or 40-85 vice versa 40 to any in range of 10-30 or 40-85 vice versa 50-60 or 73 to any in range of 10-30 or 40-72 or 75-85 vice versa 70-72, 75 to 10-60 or 73 or 78-85 78-85 to range of 00 to 75 vice versa	95	10-20 vice versa 20 to 40 vice versa Changes in range of 50-60 and 73 vice versa Changes in range of 70-72 or 75 Changes in range of 78-85	99 to any in range of 00-95	90

LUNG ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
EOD--Lymph Node Involv	0 to any in range of 1-8 vice versa 1 to any in range of 2-8 vice versa 2 to 1 or any in range of 5-8 vice versa 6 to any in range of 1-5 or 7, 8 vice versa 7 to any in range of 1-6 or 8	95		9 to any in range of 0-8	90
RX Summ--Surg Prim Site	00 to any in range of 10-90 Range of 10-14 to any in range of 20-90 vice versa Range of 20-22 to any in ranges of 10-14 or 30-90 vice versa Range of 30-32 to any in ranges of 10-22 or 40-90 vice versa 40 to any in range of 10-32 or 40-90 vice versa Range of 50-54 to any in range of 10-40 or 60-90 vice versa 60 to any in range of 10-54 or 70-90 vice versa 70 to any in range of 10-60 or 80, 90 90 to any in range of 10-70 vice versa 80 to any in range of 10-70 vice versa	95	Any changes within the range of 10-14 Any changes within the range of 20-22 Any changes within the range of 30-32 Any changes within the range of 50-54 80 to 90 vice versa	99 to any in range of 00-90	90
RX Summ--Scope Reg LN Sur	0 to any in range of 1-6 vice versa 2 to 1, or to any in the range of 3-6 vice versa 3 to 1 or 2 or any in the range of 4-6 vice versa 4 to any in the range of 1-3 or to 5,6 vice versa 5 to any in the range of 1-4 or 6 vice versa 6 to any in the range of 1-5 vice versa	95	9 to any in range of 0-6	Any in range of 0-6 to 9	90
# Reg LN Examined	00, 95, 96, 97, 98 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98	99 to any other #	90
RX Summ--Surg Oth Reg/Dist	0 to any in range of 1-9 2 or 3 to any in range of 4-7 vice versa	90	2 to 3 vice versa Changes in the range of 4-7	9 to any in range of 0-9	85
Reconstruction--First Cour	NO ERRORS – simply count # of blanks (correct) and # of other entries				

BLADDER ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	Any code other than 0	100			
Reg Nodes Pos/	00 to 01-96 vice versa 01-03 to 04-90 vice versa 97 to 98 vice versa	95	Within range of 01-03 Within range of 04-90	99 to 00-98 vice versa	90
Reg Nodes Examined	00 to any 01-98 vice versa 95, 96, 97, 98 to any in range of 01-90 vice versa	95	Within range of 95-98	99 to 00-98 vice versa	90
Reg Nodes Pos	00 to range of 01-96 vice versa 97 to 98 vice versa	95	Within range of 01-90 97 or 98 to 0-96 vice versa	99 to 01-98 vice versa	90
Reg Nodes Examined	95, 96, 97, or 98 to any in range of 01-90 vice versa	95	Within range of 95-98	99 to 01-98 vice versa	90
Tumor Marker 1	NO ERROR	95			
Tumor Marker 2	NO ERROR	95			
EOD--Tumor Size	000 to any in range of 001 to 990 vice versa	90	Changes within range of 001-990	999 to 000-990	85
EOD--Extension	05 (not valid code) 00 to any in range of 01-03 and 06-99 vice versa Range of 00-06 to any in range of 10-85 vice versa 10 or 15 to any in range of 00-06 or range of 20-85 20 or 21 to any in range of 00-15 or range of 23-85 22, 23, 40, 50 to any in range of 00-21 and range 60-85 Range of 40-50 to any in range of 00-30 or range of 60-85 vice versa Range of 60-75 to any in range of 00-50 or 80, 85 80 or 85 to any in range of 00-75 vice versa	95	01 to 03 vice versa 10 to 15 vice versa 20 to 21 vice versa Any changes between codes 22, 23, 40, 50 40 to 50 vice versa Changes within range of 60-75 80 to 85 vice versa	999 to 00-85 vice versa	90
EOD--Lymph Node Involv	0 to any in range of 1-8 vice versa 1 to any in range of 2-8 vice versa 2 to any in the range of 3-8 or 1 6, 7 to any in range of 1-3	95	8 to 9 vice versa	9 to any 0-7 vice versa	90

BLADDER ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	Any code other than 0	100			
RX Summ--Surg Prim Site	00 to any 10-90 30 to any 00-27 or 40-90 50 to any 00-30 or 60-90 60 to 00-50 or 70 to 90 70-74 to 00-60 or 80-85 vice versa 80 to 00-74 vice versa 90 to10-80	95	Changes within range of 10-14 Changes within range of 21 to 27 Changes within the range of 70-74 Range of 10-14 to range of 20-27 vice versa	99 to 00-80 vice versa	90
Scope Reg LN Sur	0 to any in range of 1-3 vice versa	90	2 or 3 to 1 vice versa	9 to any 1-3	85
# Reg LN Examined	00, 95, 96, 97, 98 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98	99 to any other #	90
RX Summ--Surg Oth Reg/Dist	0 to any 1-5 vice versa	90	Any in range of 2-5 to 1 3 to 4 vice versa 2 to 3 or 4 vice versa	9 to any in range of 0-5	85
Recons truction--First Cour	NO ERROR				

PROSTATE ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
Laterality	None	95	Any value other than 0	None	90
Reg Nodes Pos	00, 97, 98, 99 to any in the range of 01-96 vice versa 97 to 98 vice versa	95	Change numbers within the range of 01-95	99 to any other #	90
Reg Nodes Examined	00, 95, 96, 97, 98 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98	99 to any other #	90
Tumor Marker 1	Any other than 0, 1, 2, 3, 8, or 9 0 to any 1-3 or 8 vice versa 1 to 2, 3, or 8 vice versa 2 to 1, 3, or 8 vice versa 3 to 1, 2, or 8 vice versa	95	9 to any 0-3 or 8		
Tumor Marker 2	Any other than 0, 1, 2, 3, 8, or 9 0 to any 1-3 or 8 vice versa 1 to 2, 3, or 8 vice versa 2 to 1, 3, or 8 vice versa 3 to 1, 2, or 8 vice versa	95	9 to any 0-3 or 8		
EOD--Tumor Size	NONE	97	Any changes in range of 000-998	999 to any 000-998	90
EOD--Extension	Any in range of 10-15 00 to any in range of 20-85 vice versa Any in range of 20-40 to any in range of 41-85 vice versa Any in range of 41-48 to any in range of 20-40 or 50-85 vice versa 50-60 to any in range of 20-48 or 70 or 85 vice versa 00-85 to 90 or 99 95 to any 00-85 vice versa	95	Within range of 20-40 Within range of 41-48 50 to 60 vice versa 70 to 85 vice versa	99 or 90 to any 00-85	
EOD--Lymph Node Involv	0 to any 1-8 vice versa Changes in range of 1-5 to 6 or 7 vice versa 8 to 6 or 7 vice versa	95	Changes in range of 1-5 6-7 vice versa	8 to any 1-7 9 to any 1-8	90

PROSTATE ONLY					
Data Item	**Major**	**Accuracy Percentage**	**Minor**	**Unk to Known**	**Accuracy Percentage**
RX Summ--Surg Prim Site	00 to any 10-90 vice versa Any changes in the range of 30-70 Any in range of 10-17 or 20-70 vice versa 30 to any in range of 10-20 or 40-70 vice versa 40 to any in range of 10-30 or 50 or 70 vice versa 50 to any in range of 10-40 or 70 vice versa 70 to any in range of 10-50 00 to 80, 90, or 99	95	Any in range of 10-17	80, 90 or 99 to any in range of 00-70	90
RX Summ--Scope Reg LN Sur	0 to 1 vice versa 0 or 1 to 9	90		9 to 0 or 1	85
# Reg LN Examined	00, 95, 96, 97, 98 to any in range of 01-90 vice versa	95	Any changes within the range of 95-98	99 to any other #	90
RX Summ--Surg Oth Reg/Dist	00, 95, 96, 97, 98 to any in range of 01-90 vice versa	90			85
Reconstruction--First Course	NO ERROR		NO ERROR		

Appendix I: NAACCR Policy Statement 99-01: Confidentiality

Whereas:

The burden of cancer on U.S. and Canadian populations is enormous. More than 1.2 million Americans will be newly diagnosed with cancer in 1999, and more than 560,000 Americans will die from the disease in the same year. In Canada, most recent statistics suggest that more than 129,000 Canadians will be newly diagnosed with cancer and more than 63,000 Canadians will die from the disease in 1999. The lifetime probability of being diagnosed with cancer is one in two for males and one in three for females. Nearly all persons in the United States and Canada are affected by the diagnosis, treatment, or care of a family member with cancer;

Population-based cancer surveillance and research are basic and fundamental activities in cancer control, reducing the disparities among populations in early detection, access to care, and receipt of state-of-the-art treatment. Cancer research is a requisite to the discovery of new prevention and treatment strategies, the very activities that will enable success in the war on cancer;

In nearly all states and provinces, a newly diagnosed case of cancer is a reportable condition and cancer registration is required by law. Cancer patients may not choose not to be registered and may not remove their personal identities from cancer registry records. Facilities that service patients in the diagnosis or treatment of cancer may not choose not to participate in reporting. However, both patients and facilities are assured that their confidentiality will be protected. This must include the prevention of the release of their identities for legal purposes without their permission. Without this protection, compliance with cancer reporting statutes will diminish and the quality of the information reported about cancer patients will be adversely affected;

Without complete and accurate cancer surveillance data, local health authorities will not have basic information to use for defining target populations for cancer control efforts, for identifying populations most likely to benefit from cancer screening and other early detection modalities, for developing sound public health policy that is derived from scientific fact, for prioritizing public health activities based on need or community burden, for responding to citizen concerns about disparate cancer burden, and for generating questions and hypotheses to be used in prioritizing and determining appropriate directions in research;

Successful research cannot be achieved without participation of the public, both cancer patients and non-cancer patients (controls). Cancer patients must have the assurance that their voluntary participation will not result in violation of their privacy, protecting both the fact and details of their disease as well as additional information that they may be asked to divulge for research purposes;

Information entered into evidence in legal proceedings becomes public record. The principle of protection of confidentiality is violated if the information is released without the patient's consent;

For individual types of cancer, specific characteristics can be used or triangulated to produce unique records describing cases, even when the more obvious identifiers, such as name, address, or social security number (personal health number in Canada) are not part of the record. Many people, including judges and attorneys, are unfamiliar with how seemingly anonymous data items can be combined to deduce an individual's identity, especially in combination with other legally accessible data sources. Redacting name, address, telephone number and social security number (personal health number in Canada) can still allow identification of individuals under certain circumstances;

Population-based cancer registries primarily are funded through public dollars—these dollars are scarce and leave little resources for purposes other than registry operations;

Legal proceedings involving cancer registries require substantial time and expense to produce information, to respond to repeated requests for the same information by multiple parties in the legal proceedings, to educate the legal professionals in the epidemiologic perspective, to correct misinterpretations of the data, and to ensure that promises made in court are actually upheld;

Experience by at least one NAACCR member, the American Cancer Society, demonstrated that in one case data were conditionally released, and the recipients used the data beyond their original, permissible purpose, which was to use the information in a legal defense; and

The uses, in the aforementioned instance, expanded into data reanalyses that did not follow the principles or guidelines for scientific inquiry, including sound scientific method, and appropriate dialogue within the scientific community to maximize the validity of the data results and interpretation, but rather released erroneous information directly to the lay public. This action required enormous resources by the American Cancer Society to reanalyze and to correct misrepresentation of the study findings.

Therefore, it is resolved by NAACCR that:

- The integrity of population-based central cancer registries must be maintained as a key resource to protect the public's health and a key component of the public health surveillance system

- The public health surveillance system must be exempted from restrictions on collection and retention of personal identifying information in medical privacy legislation

- Personal identifiers for all cancer reports must be collected and retained in cancer registries without individual consent

- Data from cancer registries that would allow for the identification of individuals must be protected from disclosure in any legal proceedings.

Position approved by the Board of Directors on **November 17, 1999.**

Appendix J: Inventory of Best Practices Assurance of Confidentiality and Security

Name of Organization: _____

Date: _____

Effectively protecting the confidentiality of individually identifiable data requires uniform and comprehensive practices. Please indicate whether _____ (*firm name or registry)* meets the following best practices guidelines for security and data confidentiality.

General Confidentiality Practices

YES	*NO*	
❍	❍	Employees sign confidentiality agreements.
❍	❍	Confidentiality agreements with staff are signed on a routine basis at a _____ -month interval.
❍	❍	The security practices of the organization have been audited with no material findings.
❍	❍	If material findings were noted, they have been corrected.
❍	❍	Written and explicit institutional policies and procedures are in place to deal with breaches of confidentiality.
❍	❍	Methods are proactive and in place to monitor and detect the adherence to confidentiality protection procedures.
❍	❍	Data submissions are fully protected against legal discovery, including subpoena and freedom of information inquiries.
❍	❍	Organizational or institutional penalties for misuse of confidential data and breach of confidentiality by staff exist, are available in writing, and are enforced.
❍	❍	Access to data files are restricted to specific project staff and access by non-project staff is not permitted.
❍	❍	An individual is formally designated to assure compliance with established institutional standards.
❍	❍	Specific sanctions for confidentiality violation can be imposed that include employee disciplinary action and any of the following: remedial training in confidentiality, loss of certification of competency in confidentiality, prohibition from future work with confidential data at the institution, discharge.

Education

_____ *(Firm or registry)* can assure _____ *(Registry)* that it:

YES	***NO***	
❍	❍	Has developed and implemented education programs regarding confidentiality that includes information about the lack of security inherent in faxing, e-mailing, and other electronic data transfer; reminders about not using names or other personal identifiers in conversations in public areas such as open labs, elevators, or hallways; and reminders to employees of their special duty to maintain confidentiality when research involves individuals they know personally.
❍	❍	Formally credentials staff who have received confidentiality training.
❍	❍	Conducts a routine evaluation of skill and performance with regard to protection of confidentiality an identifies re-training needs based on performance.
❍	❍	Routine evaluation of employees' skill and performance is conducted.
❍	❍	Re-training needs are based on performance indicators, either for individuals or groups.

Electronic Security

_____ *(Firm or Registry)* has the following *technical practices* in place:

YES	***NO***	
❍	❍	Authentication of users by means of passwords or digital ID.
❍	❍	Access control by means of role-based authentication/access, locked server room, and an internal firewall.
❍	❍	An audit trail that documents who, when, and for what purpose data (including paper) was accessed.
❍	❍	A disaster prevention and recovery plan including adequate fire and entry alarms where data are stored; a fireproof file space for paper, routine backups of electronic data at intervals appropriate for the rate of data accrual; and offsite storage of backups (e.g., a safe deposit box).
❍	❍	External firewalls in places to prevent remote access by unauthorized users.
❍	❍	Virus checking is routine as are updates to the data files and engines to provide maximum protection of data files.
❍	❍	System assessment including diagnostics runs and external audits conducted regularly to insure the integrity of the system.
❍	❍	Data that are sent and received in conjunction with _____ *(Registry)* activities are electronically encrypted.

YES	NO	
❍	❍	A data retention schedule is defined which includes a notation of the date when files are destroyed.
❍	❍	Data file owners are notified when their file is destroyed.
		The *transfer of data* is accompanied by:
❍	❍	A data-transfer agreement incorporating confidentiality standards to ensure data security at the recipient site and set standards for the data use at the recipient site.
❍	❍	A paste (electronic) or stamp (paper) on all records containing identifiable data as a reminder of the need for special handling.
❍	❍	Telecommuting and the use of home offices maintains the same level of security and procedures to address special issues, including data-transfer agreements, secure transmission procedures, and encryption. Additional safeguards are also followed, including: maintenance of minimal data on home computer, use of electronic screen savers, and password control at home.

Paper Record Security

_____ *(Firm or Registry)* maintains the confidentiality of paper records by:

YES	NO	
❍	❍	Restricting access to data-storage areas, the use of locked file rooms or cabinets in limited-access areas, a forms tracking log for any external disclosures, and a sign-out system for internal use of data.
❍	❍	Development and implementation of policies by institutions for the secure transport of information from one physical location to another.
❍	❍	Assuring confidentiality of written evidence that a patient is on a specific research study; for example, logs or lists of screened individuals or participants should not be left out on desks or in other open-access areas.
❍	❍	Safeguarding of ancillary records, e.g., pharmacy records, data on patients screened for clinical trials participation, etc.
❍	❍	Situating FAX machines in secure or limited-access areas; use of pre-coded phone number to eliminate dialing errors; cover sheets so data are not physically exposed; testing FAX machines to insure correct number and function; and de-programming FAX memory storage after use to prevent recovery of confidential information.
❍	❍	Employing established shredding procedures for disposal of documents after use.
❍	❍	Hardcopy information of sensitive information sent outside of the department is protected.

Re-release of _____ *(Registry)* Data Files

_____ *(Firm or Registry)* does not release any _____ *(Registry)* data files to anyone without written consent of the Registry Director or designee.

A written consent is required every time a data request is received, even if the requester has obtained previous approval or if new data are added to a data file that was previously approved for release.

_____

Signature

_____

Typed Name

_____

Title

_____

Date

Appendix K: Data Use Agreements

NAACCR DATA USE AGREEMENT

DATA CONFIDENTIALITY AGREEMENT FOR NAACCR RESEARCHERS

Agreement executed this _____ day of _____, 200___, by and between

_____ ("Researcher") of _____,
(Name) *(City)*

_____
(State/Province)

and **NORTH AMERICAN CENTRAL CANCER REGISTRIES, INC.** ("NAACCR, Inc."), a California corporation. Researcher is engaged in research into the causes, control, or prevention of cancer, specifically described as follows:

_____

_____

_____

NAACCR, Inc. collects and maintains certain research data (the "Data") that will or may assist Researcher in this regard. Researcher agrees and acknowledges that patient confidentiality is of the utmost importance in the use of the Data and in the manner in which all research results are presented and/or published. Accordingly, in consideration of his/her receipt of the Data from NAACCR, Researcher agrees as follows:

1. Researcher agrees to treat the Data received from NAACCR, Inc. as private, non-public health information. The Data will be used solely for the specified research described hereinabove and not for any other purpose. The Data will never be used as a basis for legal, administrative or other adverse actions that can directly affect any individual about whom personal and/or medical information is included in the Data.

2. Researcher understands and agrees that any and all Data which may lead to the identity of any patient, research subject, physician, other person, or reporting facility is strictly privileged and confidential and agrees to keep all Data strictly confidential at all times.

3. If, in the course of his/her research, Researcher believes it necessary to provide access to the Data to any other individual, Researcher will **NOT** do so unless and until such individual has properly executed a Data Confidentiality Agreement which has been accepted, in writing, by NAACCR, Inc. And, Researcher agrees to notify NAACCR in writing within forty-eight (48) hours of his/her becoming aware of any violation of this Confidentiality Agreement or any Confidentiality Agreement executed by any other individual, including full details of the violation and corrective actions to be taken by Researcher.

4. Researcher further agrees that all data provided under the provisions of this Data Confidentiality Agreement may only be used for the purposes described hereinabove, and that any other or additional use of the data may result in immediate termination of this Confidentiality Agreement by NAACCR, Inc.

5. Researcher agrees that (i) any and all reports or analyses of the Data prepared by Researcher shall contain only aggregate data. Researcher further agrees that (ii) at no time will he/she ever publish any

individual names or other personally identifying information or information which could lead to the identification of any Data subject, and (iii) no report of the Data containing statistical cells with less than six (6) subjects shall be released without the prior written authorization of NAACCR's Executive Director, who has received written authorization from contributing registries.

6. Researcher further agrees that all data provided under the provisions of this Confidentiality Agreement shall remain the sole property of NAACCR, Inc. and may not be copied or reproduced in any form or manner without NAACCR's prior written consent.

7. Researcher shall indemnify NAACCR, Inc. from any and all liability, loss, or damage (including attorneys' fees) suffered as a result of claims, demands, costs or judgments arising out of the failure of Researcher or those acting in connection with Researcher to conform to and obey the provisions of this Data Confidentiality Agreement. In the event a claim should be brought or an action filed against NAACCR, Inc. in connection with any such failure, Researcher agrees that NAACCR, Inc. may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR, Inc., at the expense of Researcher. NAACCR, Inc., at its option, shall have the sole authority for the direction of the defense and shall be the sole judge of the acceptability of any compromise or settlement of any claims or action against NAACCR, Inc.

8. Researcher will not take any action that will provide any Data furnished by NAACCR, Inc. to any unauthorized individual or agency without the prior written consent of NAACCR, Inc.

9. Researcher will not discuss in any manner, with any unauthorized person, information that would lead to identification of individuals described in the Data furnished by NAACCR, Inc. Also, Researcher will not provide any computer password or file access codes which protect the Data to any unauthorized person.

10. Should Researcher become aware of any unauthorized access or disclosure of the Data to other persons, Researcher will report it immediately to NAACCR's Executive Director. Researcher understands that failure to report violations of confidentiality by others shall be considered as Researcher's own violation and may result in civil or criminal penalties and termination of current and future access to confidential data.

11. In the event that any attempt is made to obtain from Researcher any or all of the Data provided to Researcher by NAACCR, Inc. by subpoena or other legal means, Researcher will notify NAACCR, Inc. immediately. Researcher agrees that NAACCR, Inc. may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR, Inc. NAACCR, Inc. at its option, shall have the sole authority for the direction of the defense and shall be the sole judge of the acceptability of any compromise or settlement of any claims or action against NAACCR, Inc.

12. Researcher's obligations hereunder shall remain in full force and effect and survive the completion of Researcher's research project described hereinabove.

13. The terms of this Confidentiality Agreement shall be binding upon Researcher, his/her agents, assistants and employees.

14. Notwithstanding any contrary language in this Confidentiality Agreement, Researcher acknowledges and agrees that Researcher's access to the Data maintained by NAACCR, Inc. shall at all times be in the sole discretion of NAACCR, Inc.

15. NAACCR, Inc. reserves the right to review any and all of Researcher's reports prior to dissemination or Researcher's manuscripts before submission for publication to ensure that confidentiality is not violated and the Data are used appropriately.

16. Researcher understands that access to the Data will be terminated when the report is submitted to the NAACCR Scientific Editorial Board or on May 1, the release date of an updated NAACCR analytic file, whichever is sooner. However, the researcher may request in writing an extension to access the Data.

17. If Researcher is required by any other party or parties, including the state or a state agency, to execute any additional confidentiality agreement(s) as a condition of access to the Data, in the event of a conflict between the provisions of such agreement and this Agreement, Researcher agrees that the most restrictive agreement shall prevail.

18. This Confidentiality Agreement shall be governed by and interpreted under the laws of the State of Illinois.

Dated this _____ day of _____, 200__.

Researcher _____ ("Researcher" Signature)

_____(Print Name)

Address: _____

_____

_____

E-mail address:_____

Phone: ( ) _____ ext. _____

Received and accepted this _____ day of _____, 200__.

North American Association of Central Cancer Registries, Inc.

By:_____

Its_____

SEER DATA USE AGREEMENT
SAMPLE PUBLIC USE FILE AGREEMENT

NAME: HOLLY HOWE
004399 SEER Public-Use CD-ROM, 1973-1997

SURVEILLANCE, EPIDEMIOLOGY, AND END RESULTS PROGRAM
Public-Use Data Agreement

It is of utmost importance to ensure the confidentiality of patients who have been diagnosed with cancer. Every effort has been made to exclude identifying information on individual patients from the computer files. Certain demographic information such as sex, race, etc. have been included for research purposes. It is mandatory that all research results be presented/published in a manner which ensures that no individual can be identified. In addition, there should be no attempt to identify individuals from any computer file nor to link with a computer file containing patient identifiers.

In order for the Surveillance, Epidemiology, and End Results Program to provide a public-use or another version of data to you, it is necessary that you agree to the following provisions.

1. You will not use nor permit others to use the data in any way other than for statistical reporting and analysis.

2. You will not present/publish data in which an individual can be identified.

3. You will not attempt to link nor permit others to link the data with individually identified records in another database.

4. You will not attempt to learn the identity of any person whose cancer data is contained in the supplied file(s).

5. If the identity of any person is discovered inadvertently, then the following should be done:
   a) no use will be made of this knowledge,
   b) the SEER Program will be notified of the incident,
   c) no one else will be informed of the discovered identity.

6. You will not release nor permit others to release the data in full or in part to any person except with the written approval of the SEER Program.

7. If accessing the data from a centralized location on a time sharing computer system or LAN with SEER*Stat or another statistical package, you will not share your logon name and password with any other individuals. You will also not allow any other individuals to use your computer account after you have logged on with your logon name and password.

8. For all software provided by the SEER Program, you will not copy, distribute, reverse engineer, profit from its sale or use, or incorporate it in any other software system.

My signature indicates that I agree to comply with the above-stated provisions.

_____

Signature Date

Please fax this *signed and dated* agreement to: The SEER Program, 301-496-9949.

References

ACoS. Cancer Program Standards. Standards of the Commission on Cancer. Volume I. Chicago: American College of Surgeons Commission on Cancer; 1996.

ACoS COC. Facility Oncology Registry Data Standards Revised for 2004 (FORDS). Standards of the Commission on Cancer. Volume II. Chicago: American College of Surgeons Commission on Cancer; 2004.

AJCC. AJCC Cancer Staging Manual. Fleming ID, et al., editors. Sixth Edition. Philadelphia: Lippincott-Raven; 2002.

Band PR, Gaudette LA, Hill GB, Holowaty EJ, Huchcroft SA, Makomaski Illing EM, et al. The Making of the Canadian Cancer Registry: Cancer Incidence in Canada and Its Regions, 1969-1988. Ottawa: Canadian Council of Cancer Registries, Health and Welfare Canada, Statistics Canada; 1993.

Cancer Registries Amendment Act, Pub. L. No. 102-515, 106 Stat. 3372; Oct. 24, 1992.

Fleiss, JL. Statistical Methods for Rates and Proportions. Second Edition. New York: John Wiley & Sons; 1981.

Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, et al, editors. International Classification of Diseases for Oncology. Third Edition. Geneva: World Health Organization; 2000.

Hernandex J, Estrada L, Alvirez D. Census Data and the Problem of Conceptually Defining the Mexican American Population. *Soc Sci Q* 1973;53:671-87.

Hilsenbeck SG, Glaaefke GS, Feigl P, Lane WW, Golenzer H, Ames C, et al. Quality Control for Cancer Registries. Bethesda, MD: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health; 1985. Third printing, U.S. Government Printing Office No. 1987-793-588; May 1987.

Hillsenbeck SG, Kurucz C, Duncan RC. Estimation of Completeness and Adjustment of Age-Specific and Age-Standardized Incidence Rates. *Biometrics* 1992;48:1249-62.

IACR. Guidelines on Confidentiality in the Cancer Registry. Lyon: International Agency for Research on Cancer. IARC Internal Report No. 92/003; March 1992.

Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, editors. Cancer Registration: Principles and Methods. Lyon: International Agency for Research on Cancer. IARC Scientific Pub. No. 95; 1991.

Kramarova E, Stiller CA, Ferlay J, et al. International Classification of Childhood Cancer. Lyon: International Agency for Research on Cancer. IARC Technical Report No. 29; 1996.

Menck HR, Smart CR, editors. Central Cancer Registries: Design, Management and Use. Langhorne, PA: Harwood Academic Publishers; 1994.

Muir CS, Demaret E. Cancer Registration: Legal Aspects and Confidentiality. Chapter 15. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, editors. Cancer Registration: Principles and Methods. Lyon: International Agency for Research on Cancer. IARC Scientific Pub. No. 95; 1991.

NAACCR. Report From the Non-Hospital Audit Protocol Work Group. Sacramento, CA: North American Association of Central Cancer Registries; 1999.

NAACCR. Central Cancer Registry Record Consolidation: Principles and Processes. A Report of the Record Consolidation Committee to the North American Association of Central Cancer Registries. Sacramento, CA: North American Association of Central Cancer Registries; July 1999.

NAACCR. Death Clearance: Procedures for Central Registries. A NAACCR Instructional Module for Cancer Registries. Sacramento, CA: North American Association of Central Cancer Registries; September 1998.

NAACCR. Cancer Registry Management Reports: Design and Implementation. A NAACCR Instructional Module for Cancer Registries. Sacramento, CA: North American Association of Central Cancer Registries; June 1998.

NAACCR. Data Standards and Data Dictionary. Standards for Cancer Registries. Volume II. Version 10.2. In: Havener L, Hultstrom D, editors. Standards for Cancer Registries Volume II. Springfield, IL; North American Association of Central Cancer Registries; March 2004.

NAACCR. Data Exchange Standards and Record Description. Version 10.1. In: Havener L, Abe T, Bushhouse S, Gordon B, Hill K, Hurlbut A, Seiffert J editors. Standards for Cancer Registries. Volume I. Springfield, IL: North American Association of Central Cancer Registries; July 2003.

NAACCR. Standard Data Edits. In: Seiffert JE, Capron S, Tebbell J, editors. Standards for Cancer Registries. Volume IV. Sacramento, CA: North American Association of Central Cancer Registries; April 4, 1996. (*NOTE: Updated metafiles corresponding to changes in data standards are issued periodically. These are available from the NAACCR website, at www.naaccr.org.)*

Percy C, VanHolten V, Muir C, editors. International Classification of Disease for Oncology. Second Edition. Geneva: World Health Organization; 1990.

Pollisar L, Feigl P, Lane WW, Glaefke G, Dahlberg S. Accuracy of Basic Cancer Patient Data: Results From an Extensive Recoding Survey. *J Natl Cancer Inst* 1984;72:1007-14.

SEER. The SEER Program Code Manual. Third Edition, Revision 1. Bethesda: National Institutes of Health, National Cancer Institute. NIH Pub. No. 98-1999; January 1998 (updated January 2003).

SEER. Comparative Staging Guide for Cancer. Major Cancer Sites. Version 1.1. Bethesda, MD: National Institutes of Health, National Cancer Institute; June 1993.

Shambaugh EM, Young JL, Jr, Zippin C, Lum D, Akers C, Weiss MA. Self-Instructional Manual for Cancer Registries: Book 7: Statistics and Epidemiology for Cancer Registrars. Bethesda, MD: National Institutes of Health. National Cancer Institute. NIH Pub. No. 94-3766; September 1994.

Statistics Canada, Standards Division. Standard Geographical Classification. Volume 1: The Classification. Ottawa: Statistics Canada. Catalog No. 12-572-XPB; 2001.

Statistics Canada. Cancer in Canada 1989. Ottawa: Statistics Canada. Catalog No. 82-218; August 1993.

WHO. ICD-10: International Statistical Classification of Disease and Related Health problems. 10th Revision, Volume 1 [of 3]. Geneva: World Health Organization; 1992.

WHO. International Classification of Diseases for Oncology. First Edition. Geneva: World Health Organization; 1976.

Index